Vol. 6 No. 1
June 2016
ISSN 2047-2978

journal of global health

journal of global health
www.jogh.org

Journal of Global Health **(JoGH)** is a peer-reviewed general medical journal focusing on issues relevant for global health. The mission of the journal is to serve the community of researchers, funding agencies, international organizations, policy-makers and other stakeholders by providing an independent critical assessment of the key issues that dominated the global health community.
JoGH publishes original articles, viewpoints, research protocols and other publication items, making them available in full text online as soon as they are ready for publication.
The articles are collected in two annual print issues, at the end of June and December.
Guidelines for authors are available at http://jogh.org/contributors.htm.
All submissions and editorial inquiries should be directed to editor@jogh.org.

Publisher: Edinburgh University Global Health Society (EUGHS).

Print ISSN: 2047-2978
Online ISSN: 2047-2986
Paper version of the journal is printed on acid-free paper.

JoGH editors and editorial board members are independent researchers based at academic institutions, thus being well placed to provide professional evaluation of the key topics and ongoing activities and programs.

Journal of Global Health: The Mission Statement

Human population growth and six important pathogens that interfered with the growth (from the National Museum of Natural History, Washington, DC, USA). Photo by Igor Rudan.

The *Journal of Global Health* is a peer-reviewed journal published by the Edinburgh University Global Health Society, a not-for-profit organization registered in the UK. The *Journal* publishes editorials, news, viewpoints, original research and review articles in two issues per year.

The *Journal*'s mission is to serve the community of researchers, funding agencies, international organizations, policymakers and other stakeholders in the field of international health by:

- presenting important news from all world regions, key organizations and resources for global health and development;
- providing an independent assessment of the key issues that dominated the previous semester in the field of global health and development;
- publishing high-quality peer-reviewed original research and providing objective reviews of global health and development issues;
- allowing independent authors and stakeholders to voice their personal opinions on issues in global health.

Each issue is dedicated to a specific theme, which is introduced in the editorial and in one or more viewpoints and related articles. The news section brings up to five news items, selected by the *Journal*'s editorial team, relevant to seven regions of the world, seven international agencies and seven key resources important to human population health and development.

We particularly welcome submissions addressing persisting inequities in human health and development globally and within regions. We encourage content that could assist international organizations to align their investments in health research and development with objective measurements or estimates the disease burden or health problems that they aim to address. Finally, we promote submissions that highlight or analyse particularly successful or harmful practices in management of the key resources important for human population health and development.

All editors and editorial board members of the *Journal* are independent health professionals based at academic institutions or international public organisations and so are well placed to provide objective professional evaluation of key topics and ongoing activities and programs. We aim to stay true to principles of not-for-profit work, open knowledge and free publishing, and independence of academic thought from commercial or political constraints and influences. Join us in this publishing effort to provide evidence base for global health!

March 7, 2011 — The Editors, *Journal of Global Health*

EDITORIAL

The legacy of the Child Health and Nutrition Research Initiative (CHNRI)

Robert E Black

Institute for International Programs, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Under the Global Forum for Health Research, the Child Health and Nutrition Research Initiative (CHNRI) began its operations in 1999 and became a Swiss foundation in 2006. The vision of CHNRI was to improve child health and nutrition of all children in low– and middle–income countries (LMIC) through research that informs health policy and practice. Specific objectives included expanding global knowledge on childhood disease burden and cost-effectiveness of interventions, promoting priority setting in research, ensuring inclusion of institutions and scientists in LMIC in setting priorities, promoting capacity development in LMIC and stimulating donors and countries to increase resources for research. CHNRI created a knowledge network, funded research through multiple rounds of a global competitive process and published research papers and policy briefs. A signature effort was to develop a systematic methodology for prioritizing health and nutrition research investments. The "CHNRI method" has been extensively applied to global health problems and is now the most commonly used method for prioritizing health research questions.

In the early 1990s there was growing recognition that low– and middle–income countries (LMIC) continued to have longstanding threats from infectious diseases, malnutrition and maternal and perinatal conditions, but were also increasingly facing non–communicable diseases and injuries. Research was considered essential to address these diverse problems, but given limited resources and capacity it was thought that priorities must be set. In 1994 the World Health Organization named an Ad Hoc Committee on Health Research Relating to Future Investment Options. The report [1] issued by this Committee provided cogent arguments for better aligning research priorities with the global disease burden and building capacity for research, especially in LMIC. The report proposed a five step process to inform research and development resource allocation: 1) How big is the health problem?; 2)Why does the disease burden persist?; 3) Is enough known about the problem now to consider possible interventions?; 4) How cost-effective will these interventions be?; and 5) how much is already being done about the problem? These questions were usually asked broadly about a disease such as malaria or problem area such as emerging microbial threats. Others built upon that for research topics within these broad areas, but methods were not proposed to more systematically prioritize specific research questions. The call in this report for a focus on operational research to make existing interventions more efficient and responsive to the needs of households was largely unheard, possibly in part because the report itself named as "best buys" the development of new drugs, vaccines, tests and other technologies, rather than studies of how to enable health systems to deliver existing services more effectively and equitably.

As a follow–up to the Investing in Health Research and Development Report, The Global Forum for Health Research began as an international foundation headquartered in Geneva, Switzerland in 1997. Its aim was to increase the amount of research on global health issues. In its advocacy

it pointed to the "10/90 gap", identifying that only 10% of the world's health research spending is targeted at 90% of global health problems. The Forum continued to promote the five step process to advocate for research and held international meetings on research. As part of its mandate, the Forum facilitated the creation of more specific research initiatives, one of which was the Child Health and Nutrition Research Initiative (CNHRI). Begun under the Forum in 1999, CHNRI became a Swiss foundation in 2006. The vision of CHNRI was to improve child health and nutrition of all children in LMIC through research that informs health policy and practice. Specific objectives included expanding global knowledge on childhood disease burden and cost-effectiveness of interventions, promoting priority setting in research, ensuring inclusion of institutions and scientists in LMIC in setting priorities, promoting capacity development in LMIC and stimulating donors and countries to increase resources for research. With an international foundation Board and a Secretariat, based sequentially in Geneva, Dhaka and New Delhi, CHNRI played an active role in Global Forum annual conferences, created a knowledge network, funded research through multiple rounds of a global competitive process and published research papers and policy briefs.

A signature effort of CHNRI was to develop a systematic methodology for prioritizing health and nutrition research investments. This method included asking a wide selection of stakeholders and experts for specific research questions addressing a topic area. These questions were then curated and scored for priority using criteria such as the question's answerability and the resulting intervention's effectiveness, impact on disease, contribution to equity and deliverability. The "CHNRI method" has been extensively applied to global health problems and is now the most commonly used method for prioritizing health research questions [2,3]

In the 15 years that CHNRI operated before the foundation was dissolved in 2015, there have been substantial increases in child health and nutrition research and more reliance on sound evidence for policy and programs. The capacity for research in LMIC has improved; much more capacity building is needed, especially because research funding for global problems has improved. There has been much greater use of systematic and transparent methods involving multiple stakeholders in prioritizing and focusing research funding. The CHNRI method may be a lasting legacy of the foundation and the efforts of its Board, Secretariat and many contributors.

REFERENCES

1 World Health Organization. Investing in health, research and development. Report of the Ad Hoc Committee on Health Research Relating to Future Intervention Options. TDR/Gen/96.1. Geneva: WHO, 1996.

2 Rudan I. Setting health research priorities using the CHNRI method: IV. Revised guidelines for implementation. J Glob Health. 2016;6:010501. doi:10.7189/jogh.06.010501.

3 Yoshida S. Approaches, tools and methods used for setting priorities in health research in the 21st century. J Glob Health. 2016;6:010507. Medline:26401271 doi:10.7189/jogh.06.010101.

Correspondence to:
rblack1@jhu.edu

Africa

In the wake of Ebola and the resultant 11 300 deaths, world leaders are reflecting on its lessons that could prevent, or at least lessen the impact of, the next epidemic. To date, the focus has been on treatment and vaccines, with an announcement at the World Economic Forum on an experimental Ebola vaccine. However, experts warn that this policy – albeit vital, and in line with trends towards supporting targeted health initiatives –risks undermining more crucial, and less funded, efforts to improve basic health infrastructure. Whilst it is easy to demonstrate the impact of targeted initiatives on specific diseases, it is harder to measure the impact of systematic and more low–key initiatives on basic health care, eg, more clinics, better supplies of basic equipment, improved training and diagnosis – and donors like impact. Moreover, more people died of other causes during the Ebola outbreak, but the frightening nature of Ebola gives it a higher profile than more insidious diseases like malaria and diarrhoea. (*International Business Times*, 23 Jan 2016)

Mr Jakaya Kitwete, Africa United Global Ambassador and former president of Tanzania, has called on all African governments to prioritise access to vaccination across the continent. Africa United is an African–wide innovation led by GAVI, the Confederation of African Football, the African Union, the World Bank and the CDC Foundation, and will use sport as a catalyst on critical health issues facing Africa. Mr Kikwete spoke at his formal acceptance as Champion for Immunisation and Global Ambassador, and ahead of the launch of Africa United's universal immunisation initiative in Kigali, Rwanda. "In Kigali, my esteemed partner in this fight, CAF, will remind everyone that 'every shot counts' as we look to achieve our goal of universal immunisation," he said. (*allafrica.com*, 2 Feb 2016)

15% of Zimbabwe's population is HIV–positive, giving it one of the world's highest prevalence rates, and more than 1.3 million people are living with the virus. In 2013, the Global Fund gave a US$ 555 million grant to enable antiretroviral treatment for Zimbabweans, and currently 700 000 people access treatment. However, these efforts are being seriously undermined by Zimbabwe's drought, which has reduced its maize harvest from the required 1.2 million tonnes to 200 000 tonnes, and up to 2.4 million people are suffering from food insecurity. Nutrition and HIV are closely linked, as poor nutrition can damage the immune system and accelerate the development of full–blown AIDS. Antiretroviral drugs should be taken on a full stomach, so if people forego meals they could be forced to stop drug therapy – risking the virus mutating into a drug–resistant version. Poor nutrition also renders people more vulnerable to opportunistic infections like tuberculosis. The UN's Food and Agricultural Organization calls for "vigorous efforts to achieve and maintain good nutrition among HIV–infected people." Zimbabwe's National AIDS Council is not tasked with providing additional nutrition to HIV–positive people; and in an attempt to avert catastrophe, the government has acquired 650 000 tonnes of maize for national consumption. (*IPS*, 1 Mar 2016)

In March 2016, South Africa became the first middle–income country to fund social impact bonds (SIBs) for maternal and early childhood development. SIBs are a funding mechanism that draws on upfront capital from investors for services, and a government agency repays the investors, based on outcome performance. These impact bonds were also a first for Africa. Studies of SIBs in other countries have found that they are suited for interventions that have high potential returns (but learning, adaptability and service combinations are needed to realise them), and for interventions that are not core government–funded. Linking payment to outcomes can encourage more government investment in early childhood development; it supports performance management and adaptability; and helps develop the knowledge base of the most effective early childhood development interventions. More countries will launch SIBs in the coming years, and it is important to share lessons to improve the efficiency and quality of their implementation. (*Brookings*, 6 Apr 2016)

In a blog published in the *Financial Times*, Tom Kariuki of the African Academy of Sciences argues that although most African countries face similar health and developmental challenges, their researchers work separately, thus wasting limited human resources and infrastructure. He condemns this lack of collaboration, which meant that eg, lessons from earlier Ebola outbreaks in Uganda and the DR of the Congo could not be shared with Guinea, Liberia and Sierra Leone, as well as competing for a small pool of funding. He calls for the pooling of human resources and more career opportunities for researchers, to halt Africa's loss of 20 000 scientists each year to developed countries. Funders should promote pan–African collaboration, such as the Developing Excellence in Leadership, Training and Science Africa Initiative (DELTAS). DELTAS supports large networks and consortiums, which address emerging, infectious and non–communicable diseases. The African Union should lobby for more government funding and improve Africa's R&D expenditure – currently 1.3% of the global total. This should provide the resources for Africa to deal with its problems before they spiral out of control or spread globally. (*Financial Times*, 9 May 2016)

Asia

Michael Gideon Marmot, the President of the World Medical Association, highlighted how Thailand's universal health programme is a model for other emerging economies in Asia in its provision for low–income workers. The programme was introduced in 2002, and was rapidly extended to 18 million uninsured people, with an additional 29 million people covered by less comprehensive schemes. In Thailand, 11% of health care is met by out–of–pocket payments, compared to 63% in India – depriving Indian people of access to health care. However, there are cost constraints within Thailand's health system, resulting in the government facing pressure to abolish the US$ 0.80 co–payment programme, although there is a great deal of popular support for Thailand's health system. (*Voice of America*, 9 Feb 2016)

Increasing rates of HIV infection in the Philippines are going against the global trend of decline, with rates of new infection increasing by more than 25% from 2001 to 2011. In 2015, there were 30 356 recorded cases, with more than 80% occurring since 2010, leading to an estimated number of 133 000 recorded cases by 2022. Higher infections amongst injecting drug users (partially due to banning the provision of clean needles without a prescription), coupled with low condom use and high fertility rates have raised fears of "downstream" HIV infections to groups who are not generally at risk, such as children being infected through mother–to–child transmission. Women are often infected through drug–using partners, and are deterred from seeking follow–up exams or ART – which could prevent mother–to–child transmission – due to the stigma surrounding an HIV diagnosis. To offset this, the Sotto Treatment Hub has a designated nurse who tracks women from when they test positive to when they give birth and their babies are tested for HIV. (*Thomson Reuters Foundation*, 15 Feb 2016)

In 2014, an unlicensed doctor in a remote area of Cambodia infected 300 people with HIV, due to contaminated needles. So far, 14 people have died, and others remain weakened from the infection. Cambodia has one of the world's lowest doctor/patients ratios – 1 doctor per 5000 people – so unlicensed doctors are common. They are particularly common in remote areas with poor health infrastructure, and where it can be very expensive and difficult to travel to the nearest hospital. However, it can be argued that unlicensed doctors can help fill a yawning gap in provision, as they are on hand, can offer immediate help and are much more affordable – an important consideration when people may have to choose between food and medical care. Since the HIV outbreak, there has been a government clampdown on unlicensed doctors, many of whom do have medical training. Some NGOs are stepping in, by providing mobile clinics, training for unregistered midwives, and using local volunteers to share knowledge on health and nutrition. However, this does not resolve the fundamental problems of Cambodia's health system being under–funded and under–resourced, which was highlighted by the HIV outbreak. (*Al Jazeera*, 11 Mar 2016)

In March, Myanmar swore in its first civilian president for more than 50 years, and his party, Aung San Suu Kyi's National League for Democracy won large majorities in both houses of parliament in November. Mr Thein Sein, the outgoing president, handed over power peacefully, and Mr Min Aung Hlaing, head of the army, openly supported Myanmar's democratic transition. Whilst an apparent triumph for democracy, the situation is more complex, as Miss Suu Kyi (the people's preferred choice for president) was barred from office by the constitution which disallows anyone with a foreign spouse or children from this role. Instead, she is in charge of the foreign ministry, and intends to run the country via the nominal president – installing a puppet president is an inauspicious start. However, the larger democratic threat is the power reserved for the army, which still controls defence, border affairs and home affairs – giving it control over the country's entire administrative structure. It also dominates the National Defence and Security Council, which can impose martial law and run the country, and can potentially veto any constitutional changes. The country's future success and prosperity depends on the civilian and army sides working together, despite controlling different parts of government. (*The Economist*, 3 Apr 2016)

More than 57 000 Indonesian people with mental health conditions have been subjected to *pasung* – ie, being shackled or locked up in a confined space at least once, and currently 18 000 people are shackled. The practice was banned in 1977, but remains widespread due to a lack of awareness, superstition and inadequate mental health services (a law requiring mental health medication to be provided in primary health centres is not being implemented). In their report *Living in Hell*, the campaigning organisation Human Rights Watch calls on the government to enforce the ban on shackling and ensure access to mental health care. They recently met with Indonesia's health minister, Nila Moeloek, and were encour-

aged by her commitment to providing mental health medication in 9 500 community health centres across Indonesia – which could help end shackling. They now call on her to demonstrate her willingness to deliver by giving a full plan for providing community health centres with the necessary medication – no longer should a lack of medication be used as an excuse to shackle people. (*Human Rights Watch*, 11 Apr 2016)

Australia and Western Pacific

Tonga is the world's most obese country, with up to 40% of the population estimated to have type 2 diabetes and other non–communicable diseases and life expectancies have fallen by 10 years (from the mid–70s to mid–60s). The traditional Tongan diet consisted of fish, root vegetables and coconuts, but gradually was replaced by cheap offcuts of meat from the USA and New Zealand and tinned produce. Tonga's obesity problem may partly be genetic, as islanders had to survive long periods without food so their bodies are programmed to hold onto fat stores. However, Tongan society traditionally views overweight people as attractive, and there is a culture of over–eating at celebrations. There are government efforts to raise awareness of the risks of diabetes, but there is little evidence of changing lifestyles or diet. According to one doctor, this will take generations and things will worsen before they improve. (*BBC*, 18 Jan 2016)

Australia's annual *Closing the Gap* report shows that indigenous Australians continue to die younger, are more likely to be unemployed and with lower educational levels than other Australians. The government has seven targets across health, education and employment to close this gap, and there are five failings in these areas. First indigenous people die 10 years younger than non–indigenous people, although recent declines in mortality rates from chronic and circulatory diseases may improve this in the future. Second, the infant mortality rate for indigenous people is 6.4 per 1000, compared to 3.6 across all Australians. Third, 60% of indigenous students finish Year 12 of school, compared to 86.5% for other Australians. Fourthly, the employment rate amongst indigenous people is 47.5%, compared to 72.1% amongst non–indigenous Australians. Indigenous people are more vulnerable to economic shocks, and unlike non–indigenous Australians, their employment levels have fallen from 2008 levels. Finally, an indigenous man is 15 times more likely to be jailed than a non–indigenous man – indigenous people comprise 3% of the Australian population, but are 25% of the prison population. (*Sydney Morning Herald*, 10 Feb 2016)

The Marshall Islands are the latest Pacific Island country to confirm the arrival of the zika virus. Although there is one sole case to date, the government has declared an "outbreak" and a state of health emergency due to the islands being extremely vulnerable to the disease. It is the islands' dry season, and zika could spread further and affect more pregnant women when the rains start – the state of emergency aims to prevent this happening. Efforts are currently focused on pregnant women, with volunteers distributing leaflets house–to–house on cleaning, mosquito control, symptom recognition and zika kits (containing mosquito repellent, treated bed nets and condoms). The government is considering importing blood supplies to avoid transmission by contaminated blood. With scarce health facilities on the outlying islands, health officials are emphasising zika prevention in these areas. (*Radio New Zealand*, 8 Mar 2016)

More than 200 000 women and girls in New Zealand have been vaccinated against the human papilloma virus (HPV), which is linked to several cancers, including cancer of the cervix and throat. However, in contrast to Australia where the vaccine is available to both males and females, New Zealand restricts state–funded HPV vaccinations to females, and males aged 9–26 who identify as gay. However, the prevalence of HPV–linked oropharyngeal cancers is higher amongst men – 4 per 100 000 compared to 1 per 100 000 for women – and is approaching the incidence of cervical cancer. This had led to campaigners to call for government funding for HPV vaccination to be extended to boys. "Men are more exposed to the virus because the route of exposure is understood to be oral sex and that the concentration of virus in the female genital tract is much higher than in the male tract," says ENT surgeon Dr John Chaplin. (*New Zealand Herald*, 20 Mar 2016)

A young woman – a refugee from Africa – was raped in Nauru's refugee detention centre, and was sent to Papua New Guinea to terminate her pregnancy. Abortion is illegal in Papua New Guinea, and Australia's Federal Court ruled that Australia owes her a duty of care to ensure she receives a safe and legal abortion – Australia's government had previously argued that it owes no duty of care towards her. The woman – known only as S99 – suffers from violent epileptic seizures, severe mental health issues and specialist medical needs arising from a medical procedure she was subjected to as a young girl. Her barrister told the Court

that sending S99 to Papua New Guinea for an abortion was tantamount to "procuring illegal conduct," and the court ruled that it was unreasonable to consider an abortion procured in Papua New Guinea as a safe or legal procedure. However, the court did not require her to be brought to Australia for the procedure, and it is expected that a third country will be found for her termination. (*Sydney Morning Herald*, 6 May 2016)

China

China's government estimates that 13% of China's mainland population has a "psychiatric handicap", and the World Health Organization estimates it as almost 10%. Despite these large numbers, there are only 20 000 mental health specialists in the entire country, and 1.7 psychiatric beds per 10 000 people – the global average is 4.4 beds. This gap has led to Dr Guan Weili to set the Wenzhou Kangning Hospital, China's largest private mental health facility. The hospital owns five centres, manages four others and plans to expand into Hong Kong. Although psychiatric hospitals have existed since 1949, the first mental health laws – including the right of patients not to be detained against their wishes – were only implemented in 2013. Mr Guan believes that the rise in psychiatric problems is linked to China's rapid economic growth – indeed, some large corporations offer 24–hour helplines and suicide prevention counselling – comparing it to Japan at a similar stage of development. (*Forbes*, 26 Jan 2016)

Shares in Alibaba Health Information Technology Ltd fell as China suspended a drug–coding system (PIATS) developed and owned by the company. PIATS identifies counterfeit medicines, and generated nearly 50% of the company's revenue in the previous financial year. The company received no advance warning of the Chinese Food and Drug Administration's decision, who have made draft amendments to existing rules to allow other methods of tracking drugs to their source. The company plans to continue operating PIATS, and will work closely with regulators to continue such operations under their direction. (*Bloomberg*, 22 Feb 2016)

27 people have been arrested for the illegal sale and distribution of incorrectly stored vaccines across China; and four hospital chiefs have been charged with buying vaccines from illegal sources. The vaccines are allegedly still safe, and people who have been vaccinated do not need to be re–vaccinated. However, neither the vaccines' batch numbers, nor the clinics which have used them, have been disclosed, leading to parents and experts demanding more transparency. China's State Council is reforming the legal framework for privately–produced vaccines by making provincial disease control centres responsible for purchasing vaccines. This would block the current loophole that allows vaccines to be sold privately. (*South China Morning Post*, 14 Apr 2016)

500 school students at the Changzhou Foreign Languages School in Jiangsu suffered serious health problems, including cancer, after their school was relocated to a site near former chemical plants. 493 students developed blood abnormalities, leukemia, lymphoma, dermatitis, eczema and bronchitis. A report on China's state broadcaster China Central Television showed that soil and groundwater in the area contained toxic compounds and heavy metals, and one carcinogen was 100 000 times above the safety limit. It appears that the environmental assessment prior to the school's construction did not look for pesticides, and builders used heavily–polluted groundwater during construction. China increasingly views environmental pollution as a destabilising social factor rather than the inevitable result of economic expansion, and has recently increased efforts to combat it. (*Irish Times*, 18 Apr 2016)

Following the end of China's one–child policy, the role of the country's army of family planning officers is changing. Over the past 35 years, family planning officers covered cities, towns and villages across China, and uncovered families suspected of breaking China's one–child legislation. They were widely distrusted – confiscating property if families couldn't afford to pay fines levied for breaking the one–child rule, and pressurising women into abortions, even when women were six months' pregnant. However, some are being retrained in child welfare, and are teaching parents and grandparents how to develop and stimulate toddlers' minds by talking, singing and reading to them. Overall, China's one–child policy has led to a shrinking labour–force that could undermine economic growth and innovation within the workplace, and investing in young children's development could be essential for the country's future. (*BBC*, 4 May 2016)

Europe

According to a report compiled by Health Consumer Powerhouse, waiting times for emergency treatment in Irish hospitals are the worst in Europe, and frequently exceed 3 hours. Overcrowding in emergency departments places the health system under strain, with a snapshot of 391 patients on trolleys showing that 210 had waited more than nine hours. In addition, waiting times for minor operations and CT scans also rank amongst Europe's longest. Ireland also ranks poorly for direct access to a specialist and physical activity in schools, and worst for binge drinking. However, Ireland's smoking rates are amongst Europe's lowest, and is rated highly for access to essential drugs. Other problems include the high percentage of people purchasing duplicate health insurance, although the report praised Ireland's "dedicated efforts" in halving the rate of MRSA infection between 2008 and 2015. (*Irish Times*, 26 Jan 2016)

NGOs warn that HIV infections are rising in eastern Ukraine, where treatment and prevention programmes, condom distribution to high–risk groups, and needle exchange programmes have been affected by two years of conflict between pro–Russian separatists and government forces. Even before the conflict, Ukraine had one of the highest rates of HIV infection in Europe – 1.2% of Ukrainians aged between 15–49 years were HIV–positive, and this may have doubled. 30% of new infections are in the partially rebel–controlled areas of Donetsk and Luhansk, and HIV–related tuberculosis infections are also common. There are shortages of antiretroviral treatment – although this is slowly improving – and diagnostic and treatment equipment. Mikhail, an HIV–positive man, says "it's hard to live in such conditions. Half of the pharmacies are closed [and those that are open] don't even have the most basic medication." (*IRIN*, 23 Feb 2016)

Despite high overall spending on welfare, Finland's health system is under–resourced, with long waiting times – a 2012 report found that 80% of people had to wait 2 weeks to see a family doctor – and severe cost pressures. Finland spends 7% of its GDP on health, compared to 8% in the UK, and has a high proportion of private primary care doctors. Some Finnish people use private doctors in nearby Estonia to save time and money. The health service is locally funded, so that poorer areas have correspondingly lower–quality health care. Users of the public health system face high charges for drugs and consultations, and many people are forced to use private health care. However, reforms are planned, with 301 municipalities being merged into 19 larger, more efficient, organisations. (*The Guardian*, 23 Feb 2016)

In the country's 2016 Budget, the UK's Chancellor, Mr George Osbourne, announced a sugar tax on soft drinks. The tax, aimed at reducing the rise in childhood obesity, will begin in 2018 and money raised will fund sports activities in primary schools. It will be levied on drinks companies, and assessed on the volume of sugar–sweetened drinks they produce or export. The tax will be graduated, with one level for drinks with sugar above 5 g per 100 ml, and a second level for drinks with more than 8 g per 100ml, leading to price increases of GBP 0.08 (US$ 0.11) per drink. Mr Osbourne noted that a can of cola typically contains 9 teaspoons of sugar, and others have up to 13 teaspoons – more than double a child's recommended added sugar intake. A 2015 report by Public Health England recommended a 10–20% tax on high–sugar products. The move was welcomed by the British Medical Association, the Labour Party and Jamie Oliver, the celebrity chef and healthy eating campaigner, although the political party UKIP opposed it. (*Huffington Post*, 16 Mar 2016)

Romania's health minister Patriciu Achimas–Cardariu has resigned following public protests over the use of sub–standard disinfectant in dozens of Romanian hospitals. Authorities conducted searches at hospitals and at the drug company, Hexi, which reportedly supplied the disinfectant for use on surfaces and hands. Police also took away documents and samples from 25 hospitals. This is the latest problem to hit Romania's underfunded health system, which faces an outflow of staff and systemic bribery and informal payments. (*Washington Post*, 9 May 2016)

India

A 2010 World Bank estimate showed that premature mortality, lost productivity, health care provision and other losses due to inadequate sanitation costs India US$ 53.8 billion each year. However, Shamika Ravi and Rahul Ahluwalia argue that large gains in India's public health, especially the health of its poorest people, can be made if the

government prioritises the expansion and effective delivery of "public goods", such as vaccination, health education, sanitation, public health, primary care and screening, and reproductive and child health. These gains could be made economically, as neighbouring Sri Lanka and Bangladesh spend less on health as a percentage of GDP, yet have better outcomes. India should therefore focus on setting appropriate goals, and reforming the public health sector's governance and management systems so it can deliver these goals. There are also severe shortages of qualified health professionals, especially in rural areas, and the government must be creative in addressing this. Health financing requires reform, and Shamika Ravi and Rahul Ahluwalia also urge the adoption of Medical Savings Accounts with tax deductions for medical expenses, and direct payments for those who cannot pay themselves. (*Brookings*, 26 Jan 2016)

▸▸ New Delhi's pollution levels exceeded Beijing and Shanghai on 24 Dec 2015, when they reached levels of 295mg/m^3 for PM 2.5 and 470 mg/m^3 for PM 10 (against recommended upper limit of 60 and 100 respectively). The city's pollution crisis has led to a sharp rise in respiratory illnesses, skin and eye allergies, cardiac arrest, memory loss, depression and lung damage, and 4–in–10 children suffer from severe lung problems. Pollution is responsible for 10 000–30 000 deaths in New Delhi each year, and is the 5th largest cause of death in India. New Delhi's massive population growth is fuelled by polluting industries, sharp rises in the number of vehicles, and spiralling energy consumption supplied by polluting power stations. Thanks to public pressure, the Delhi government has introduced a number of measures, including road rationing to reduce vehicle pollution. Environmentalists state that this ignores the role of industry, and calls on Delhi to copy China's lead in issuing pollution–related alerts. (*The Diplomat*, 8 Jan 2016)

▸▸ India is conducting its first survey of the prevalence of drug–resistant tuberculosis (TB), and plans to release the results by December 2016. With an estimated 2.1 million cases, India has the world's highest number of TB patients, and is believed to have the largest number of drug–resistant TB cases after China. The World Health Organization describes drug–resistant TB as a global threat to TB treatment. The survey would improve the detection of drug–resistant TB by highlighting high–prevalence areas, and inform India's future TB–control strategy – and improved detection is vital to prevent transmission of the disease. (*Reuters*, 7 Mar 2016)

▸▸ Diabetes is rapidly increasing in India, with 70 million cases amongst the adult population in 2015, and prevalence has risen by 80% amongst women between 1980 and 2014. According to research published in the *Lancet*, India has the second–highest prevalence of diabetes worldwide, and until it was recently overtaken by China its prevalence rate was the highest. To help alleviate the problem, India's National Programme for Prevention and Control of Cancer, Diabetes, Cardiovascular Diseases and Stroke is focusing on awareness raising, behaviour and lifestyle changes, screening, and early diagnosis and referral for at–risk people. (*Times of India*, 27 Apr 2016)

▸▸ In 2015, India inaugurated its Mission Indrahanush, which aims to immunise every child against 7 vaccine–preventable diseases by 2020, and achieve above 90% coverage – compared to the current 65%. This target is additionally ambitious, in light of India's Universal Immunisation Programme (UIP) introduction of four new vaccines (polio, rotavirus, rubella and Japanese encephalitis). However, there are 25 vaccine–preventable diseases, which will require the expansion of the UIP over time, raising the issue of funding. This is particularly complex given India's large population, and it will be ineligible for GAVI support by 2021. A recent report published by the IKP Trust and Global Health Strategies outlined the possible costs and finance options for expanded immunisation coverage. Amongst other options, it considered a "National Trust Fund for Health and Immunisation." Self–financing trust funds can potentially protect against volatility in drug prices and donor support. The report authors also call for all expenditure on maternal and child health and preventative health care to be re–classified as capital expenditure rather than revenue expenditure; and for the US$ 471 million needed for India to meet its immunisation targets to be allocated additionally and separately from the health ministry budget and ring–fenced as an annual recurring outlay, indexed for inflation. (*Asian Age*, 6 May 2016)

The Americas

▸▸ Barbados was one of seven countries to take part in the first WHO/UN Development Programme global project on adapting public health systems to climate change. In Barbados, one of the key aims was to improve water storage facilities to eliminate mosquitoes, give technical advice on building and maintaining water tanks, and raise awareness about safe ways to harvest rainwater. This is especially crucial for Barbados, which has a high rate of dengue fever and

some recently–detected cases of zika. As zika spreads, there is more pressure to analyse the health impacts of climate change and extreme weather, and to understand how climate stresses can shape health risks. Although any link between climate change, powerful El Niño weather phenomenon and the rise of zika is unproven, it is certainly plausible that unusual weather conditions make it easier to transmit the virus. Understanding these linkable could lead to targeted preventative public health measures in areas at high risk of an outbreak. (*Thomson Reuters Foundation*, 29 Feb 2016)

▸▸ According to a report from Amnesty International, access to contraception and abortion can be a lottery in Latin America, and is often dependent on the woman's ability to pay or the personal and religious views of health workers. In most countries in the region, abortion is only allowed in cases of rape, incest or if the mother's life is in danger, giving the region some of the world's strictest abortion laws. When combined with the low availability of contraception – according to the UN Population Fund, 1–in–3 women of child–bearing age who would like to use birth control has no access – women can be forced to undergo dangerous backstreet abortions, which causes at least 10% of maternal deaths in Latin America. Each year, around 760 000 women in Latin America receive hospital treatment for complications related to unsafe abortions. This spread of zika has led to some countries recommending that women delay pregnancy. However, Amnesty states "this recommendation is not just absurd, it is insulting in a region where more than half of pregnancies are unwanted or unplanned, where there are extremely high rates of sexual violence, where the demand for contraception far outstrips availability." (*Thomson Reuters Foundation*, 7 Mar 2016)

▸▸ Almost 10% – US$ 619 billion – of the US government's health spending is on Medicare, the country's health insurance scheme for elderly people. Medicare covers the average market price for a drug, plus a 6% premium and separate compensation for administering the drug. This incentivises doctors to prescribe expensive drugs over cheaper, similar, drugs. To reduce costs, the federal government has proposed reducing the 6% premium to 2.5%, plus a flat fee for treatment. There are concerns that this proposal will stifle innovation from smaller providers, as lower margins could mean that treatment could only be provided at scale, eg, within hospitals. Doctors and drugs companies, the main beneficiaries of the current arrangements, are also strongly opposed to the proposal, but it is difficult to disentangle justifiable concerns from scaremongering. (*The Economist*, 16 Apr 2016)

▸▸ Following the 7.8 magnitude earthquake which struck Ecuador in May, killing more than 600 people, causing billions of dollars' worth of damage and leaving 720 000 people in need of humanitarian assistance, the government and the UN have launched an urgent appeal for US$ 72.7 million from the international community. However, response has been slow with donors only raising US$8.6 million, with an additional US$ 6.5 million pledged outwith this appeal. This brings the total to US$ 15.1 million – far short of what is required. Homes, roads and public infrastructure have been razed by the earthquake, and Ecuador's already–struggling economy will slip back into recession as GDP will fall by an estimated 2–3%. Lenders such as the World Bank have opened US$ 600 million of credit to help the Ecuador response, and the government announced emergency finance measures (one–off taxes, asset sales etc.) to fund reconstruction. "We thank those countries who have responded to our appeal and call on others to do the same," said Mr Jens Laerke, a UN spokesperson. (*Public Finance International*, 9 May 2016)

▸▸ Undocumented immigrants in Canada often avoid mainstream health services because they fear being reported to the country's border services – some patients have said they could rather die than be deported – and the US$ 695 hospital consultation fee acts as a further deterrent. Bryon Cruz, an outreach worker with the migrants' rights group Sanctuary City, works to connect undocumented immigrants with the health care they need, and notes that people have been deported for accessing medical services. He receives 25 calls a week from undocumented immigrants too afraid to access mainstream services, and he has a network of social and health care workers who will treat these people without reporting them. Examples of this work include an off–duty doctor replacing a patient's dislocated shoulder behind a café, and a veterinarian was on the brink of stitching another patient's wound until a nurse came forward. Mr Cruz has seen progress from Vancouver Coastal Health and Fraser Health on not reporting immigrants, and calls for the children of undocumented immigrants to get free health care. (*CBC news*, 5 May 2016)

The Bill and Melinda Gates Foundation

Bill Gates and George Osbourne, the UK Chancellor of the Exchequer, announced funding of US$ 4.3 billion over the next five years in efforts to eliminate malaria. The fund comprises US$ 345 million from the UK's overseas budget for the next five years, with US$ 200 million in 2016 from the BMGF, with more donations to follow. In a joint letter to *The Times* newspaper, Mr Gates and Mr Osbourne highlight the global scale of malaria – 200 million cases each year, and the economic cost – mainly borne by Africa and running into billions of dollars – of lost productivity and public health expenditure. According to the World Health Organization, there were 438 000 malaria deaths in 2015, mostly in children aged under 5, and mostly in Africa. Progress is controlling malaria is being undermined in the spread of resistance to antimalarial drugs and to insecticide. (*The Guardian*, 25 Jan 2016)

In November 2015, the BMGF announced an additional US$ 120 million investment in family planning over the next three years. This is a 25% increase on its current funding level, and aims to meet the Family Planning 2020 goal of giving a further 120 million girls and women voluntary access to birth control. The BMGF will continue to invest in new forms of birth control to expand the range available to women, eg, injectable methods than can be easily delivered by community health workers, or self–administered at home. (*Devex*, 11 Feb 2016)

In an interview ahead of the publication of the BMGF's annual letter, Bill and Melinda Gates highlighted how, in the wake of the Ebola tragedy, the zika virus has spurred brought a faster and more united response. Bill Gates noted that the BMGF has invested in modifying mosquitoes not to carry viruses, and in reducing their numbers, and that the same breed of mosquitoes carries dengue and zika pathogens. This year's letter calls for young people's involvement in tackling inequity, focusing on energy and time. It highlights the need for cheap carbon–free energy that would benefit people in developing countries, and the gap in the amount of time spent on unpaid work between men and women. This gap hampers people rising out of poverty; bringing labour–saving devices to developing countries would help free women to earn money for their families and improve health care and nutrition. (*Reuters*, 23 Feb 2016)

PATH, the non–profit global health organisation, is opening a Center for Vaccine Innovation and Access, with initial funding of US$ 11 million from the BMGF. PATH will use the funding to accelerate the development and distribution of vaccines to halt deaths from preventable diseases. Currently, PATH has more than 20 vaccines at different stages of development and use, which target the world's leading causes of child mortality – pneumonia, diarrhoea, malaria, plus diseases like polio and meningitis. PATH also plans to use the new Center to address the new threats from diseases like ebola and zika. "The new Center will bring together PATH's expertise across the entire vaccine development and introduction process, from pre–clinical trials on novel vaccine concepts to regulatory approval and policy review, from design and conduct of field trials to innovative approaches for new vaccine development," says Mr David Kaslow, the Center's head. (*GeekWire*, 16 Mar 2016)

The BMGF has made a US$ 5 million equity investment in Amyris Inc., a US–based bioscience company. The investment will fund work on further reducing the cost of a leading malaria treatment, focusing on the continued production of high–quality and secure supplies of artemisinic acid and amorphadiene for use in artemisinin combination therapies (ACTs), which are recommended by the WHO as the primary first–line treatment for malaria. Amyris made its artemisinic acid–producing strains available to Sanofi in 2008 on a royalty–free basis. Sanofi scaled–up this technology to produce artemisinin for ACT treatments, intending to produce enough semi–synthetic artemisinin for up to 150 million treatments by 2014, and ensure distribution on a "no profit, no loss" principle. (*Business Standard*, 12 Apr 2016)

The GAVI Alliance

GAVI has signed a US$ 5 million deal with the pharmaceutical company Merck to keep 300 000 Ebola vaccine doses ready for emergency use or further clinical trials. Merck will submit a licensing application by the end of 2017, which would help GAVI prepare a global stockpile. Early trials of the VSB–EBOV vaccine – which combines a fragment of the Ebola virus with another safer virus – suggest it may give 100% protection, although this is still preliminary. The deal was announced at the World Economic Forum at Davos, Switzerland. Isolated flare–ups of Ebola

are still anticipated, and the vaccine could be important with dealing with these, as well as heading off any future epidemics. (*BBC*, 20 Jan 2016)

▸▸ Médecins Sans Frontières (MSF) expressed grave concern that the high cost of vaccines is not being given higher priority at the Ministerial Conference on Immunization in Africa. MSF say that the high cost of vaccines affects its ability to provide health care in developing countries, and call for pharmaceutical companies to reduce the cost of three–dose pneumonia vaccinations to US$ 5 per child, amongst others. MSF argue that lack of immunisation progress in some countries since 2013 is due to high prices – eg, vaccinations for pneumonia, diarrhoea and HPV have increased 68 times between 2001 and 2014. MSF are concerned that countries that are not poor enough to qualify for GAVI support have to negotiate prices on their own, risking their coverage rates. They call for GAVI to negotiate better deals with pharmaceutical companies. "From 2001–2014, the US has given GAVI US$ 1.2 billion in direct funding, and has pledged US$ 1 billion for 2015–18. This money can go much further if the vaccines, like Pfizer's pneumonia vaccine, are cheaper." (*Humanosphere*, 26 Feb 2016)

▸▸ Nepal passed the country's immunisation bill in January, which aims to improve oversight of immunisation services, set higher standards for vaccine testing and usage, and change how Nepal finances its immunisation programme. Nepal currently relies on financial support from GAVI to fund 60–70% of its purchases. However, Nepal will no longer be eligible for GAVI support when it transitions from low– to middle–income status – expected by 2022, thus giving a few years to establish its own domestic financing arrangements. The new law sets out two methods for financing immunisation. First, the law commits the government to allocating funds to the National Immunization Fund, levied through general taxation. Second, health partners can contribute to a separate Sustainable Immunization Support Fund – although this will probably requires incentives such as tax exemptions to be effective. The law highlights Nepal's commitment to immunisation, and the Chairperson of the Parliamentary Committee on Women, Children, Senior Citizens and Social Welfare, Hon. Ranju Kumari Jha, calls it "a milestone to protect child rights of getting quality immunisation service, increase country ownership and sustain the national immunisation programme by securing adequate funding." (*healthaffairs.org*, 7 Mar 2016)

▸▸ Seth Berkley, GAVI's CEO, believes that Ebola and zika have removed attention from measles. He argues that measles should be prioritised, partly because it is highly infectious and kills 115 000 people each year, and also because measles outbreaks act as an early warning system against other threats to global health security. Measles' highly infectious nature means that outbreaks are a useful measure for gauging a health system's ability to cope with potential global epidemics, as 90% immunisation coverage is needed to reach herd immunity, compared to 80–85% for other common diseases. If populations are under–immunised, it is likely that other vital health interventions are lacking, rendering people even more vulnerable to disease outbreaks. Mr Berkeley calls for more resources on routine immunisation, supplemented with catch–up campaigns as required. (*Devex*, 27 Apr 2016)

▸▸ GAVI is backing a new national drone delivery network to distribute blood supplies in Rwanda. There will be further tests planned of its suitability for a wider range of drugs, including vaccines, HIV treatments and treatments for malaria and tuberculosis. This phase will see the drones making up to 150 deliveries of blood to 21 transfusion facilities in western Rwanda – crucial for Africa, which has the world's highest rate of maternal deaths from postpartum haemorrhaging. If successful, Rwanda's drone network could save thousands of lives and be a model for other countries to duplicate. The project uses drones from the Californian robotics company Zipline, and the "global citizenship" art of the delivery and logistics giant UPS. "It is a totally different way of delivering vaccines to remote communities and we are extremely interested to learn if UAVs [unmanned aerial vehicles] can provide a safe, effective way to make vaccines available for some of the hardest–to–reach children," says Seth Berkley, CEO of GAVI. (*Pharma Market Live*, 9 May 2016)

The World Bank

▸▸ Ahead of a World Bank conference on how improved land management can reduce poverty and foster development, Mr Klaus Deininger, the conference organiser, argues that women's right to land creates other benefits. These include improved health and education for children, increased household resources, and fewer child marriages as daughters are less likely to be married for financial reasons. Women with land rights tend to have bank accounts, and their financial resources can render them less vulnerable to domestic violence. In sub–Saharan Africa, women com-

prise more than 50% of the agricultural labour force, but fewer than 20% own farms. According to the UN World Food Programme, if women farmers had the same access to land as men farmers, global hunger could be substantially reduced. The conference will focus on women and property, with particular emphasis on gender equality and land rights – key to achieving the Sustainable Development Goals. (*Thomson Reuters Foundation*, 13 Mar 2016)

▸▸ Costa Rica, already considered to have one of the best health care systems in Latin America, has been granted a US$ 420 million loan to further strengthen the financial sustainability of its universal health insurance system, and the management, organisation and delivery of its services. This is in line with Costa Rica's strategic health agenda, which was developed by the Costa Rican Social Security Administration to modernise primary health care networks. It will include the expansion of e–health and a 40% increase in screening in areas with high incidence of colon cancer. The programme will include a 25% advance to ensure progress. (*Public Finance International*, 21 Mar 2016)

▸▸ The World Bank announced a US$ 5 billion loan to Tunisia to support its democratic transition and economic development. Tunisia has been hit with falling tourist revenues – which account for 7% of GPD – after the Islamic militant attacks in 2015, unrest over unemployment and limited economic reforms despite wider political advances following its 2011 uprising. Economic growth was 0.8% in 2015, and is forecast to increase to 2.5% in 2016, but unemployment is 15.1% and is much higher amongst the country's youngsters – who comprise more than 50% of the population. The loans will be used to stimulate investment and job creation, and to intensify development in disadvantaged areas. The World Bank agreed that Tunisia's economic reforms to date are headed in the right direction, but more reforms are needed in the financial sector and to increase transparency. The International Monetary Fund and Tunisia are also in talks over a US$ 2.8 billion credit to support economic reform. (*Al Arabiya*, 25 Mar 2016)

▸▸ Following on from the annual spring meeting of the World Bank and IMF, 5 key themes have emerged. First, the World Bank's track record of involuntary resettlement – which has faced severe criticism from human rights groups – was put under the spotlight, as the bank's Inspection Panel made recommendations for better practice. Second, there were moves towards closer collaboration with other development banks, which could help close the US$ 60–70 billion infrastructure gap in Africa, amongst others. Third, in the wake of the "Panama papers", World Bank President Jim Yong Kim emphasised how tax avoidance hinders ending poverty, and that world leaders wish to work with the Bank to track down illicit revenue flows. Fourthly, the President of the African Development Bank, Akinwumi Adesina, wants Africa's leaders to focus more on nutrition. And finally, UN Secretary General Ban Ki–moon called for more work on addressing the root causes of conflict behind the global refugee crisis, and for the world to mobilise to ensure the safety and well–being of those crossing borders. (*Devex*, 20 Apr 2016)

▸▸ According to a World Bank study, South Asia could create millions of new jobs in the clothing industry by taking advantage of rising manufacturing costs in China, boosting both economic growth and job opportunities for women. With low labour costs and a growing young, working–class population, South Asia is strongly positioned to increase its share of this labour–intensive industry. Women's participation in South Asia's labour market is low, and increasing job opportunities for women is vital for raising marriage ages, reducing birth rates, improving nutrition and school enrollments, and stronger economic growth. However, the industry has a track record of poor working conditions – highlighted by the collapse of the Rana Plaza building in Bangladesh in 2013 – and growth opportunities will not be fully realised without closer attention to safety and improved conditions, due in part to increased scrutiny from global brands and retailers. (*Voice of America*, 29 Apr 2016)

▸ United Nations (UN)

▸▸ The World Food Programme (WFP), a UN agency, is leading a project to boost incomes and improve food security in developing countries. It will help 1.5 million small–scale farmers across Africa, Asia and Latin America with contracts to buy their crops, signed before they are planted, to a value of US$ 750 million. It aims to enable marginalised farmers to access reliable markets – 50% of the world's 795 million people are farmers, and in some African countries up to 90% of the population are smallholder farmers – so that farmers could move from subsistence to market–oriented production. However, critics warn that the project could fail if it does not prioritise helping poor farmers to adapt to climate change, by promoting crops which are more resilient to drought. In addition, there are concerns that farmers will be encouraged to buy hybrid seeds which require chemical fertilisers which deplete soil,

and rely on regular rainfall. The WFP responded by noting that participating farmers can buy seeds from any supplier, but will receive recommendations on which seeds are best for their soil and water conditions. (*Thomson Reuters Foundation*, 10 Feb 2016)

▸▸ 175 world leaders gathered in New York to ratify the Paris climate deal on the world Earth Day, marking the first steps towards binding the countries to the promises they made to cut greenhouse gases. It will come into effect when the 55 countries responsible for 55% of greenhouse gases have ratified the accord, and is set to begin in 2020. China and the USA have agreed to ratify in 2016, and the EU's 28 member countries are expected to ratify within 18 months. The agreement comes as the 2016 El Niño is believed to have caused droughts, floods, severe storms and other extreme weather patterns, and 2016 is set to break global temperature records. In welcoming the agreement, the UN Secretary–General Ban Ki–moon said "the era of consumption without consequence is over. We must intensify efforts to decarbonise our economies. And we must support developing countries in making this transition." (*Al Jazeera*, 22 Apr 2016)

▸▸ April's UN General Assembly Special Session on the World Drug Problem (UNGASS) did not lead to any radical shifts in drug policy. The central goal of the UN global drug policy is the elimination in the sale and use of illegal narcotics. The hard–line interpretation of this policy – used by most countries – does not lead to harm reduction, which underpins the UN conventions on drugs. Countries such as Mexico, Guatemala and Colombia have agreed that this approach has failed, and benefits criminals. Whilst UNGASS has not shifted from this policy, there are changes in the language around drug use, which reflects a greater focus on prevention and treatment – albeit falling short of what is required to address the estimated 400 000 drugs–related deaths each year. In moves widely seen as significant, countries such as Mexico and Canada quietly announced at the session that they are moving away from UNGASS policy by introducing their own reforms (eg, on cannabis use and legalisation). Many campaigners call for full decriminalisation of drug use, although this would not ensure the elimination of violence and corruption around black markets. In summary, whilst there was no radical changes, the session heralds some important first steps in the evolution of global drugs policies. (*Huffington Post Australia*, 2 May 2016)

▸▸ The UN has convened the first World Humanitarian Summit (WHS), to take place on 23–24 May, in response to the worst humanitarian crisis since World War II. One UN report found that the average length of displacement is 17 years, and another UN/World Bank report found that 90% of Syrian refugees in Jordan and Lebanon live below the national poverty line, and many are unable to legally earn money, and many children cannot access education. However, commentators have noted it is unclear what outcomes or actions the summit will produce. Médecins Sans Frontières (MSF) have pulled out of the summit, expressing concerns that the summit will not improve emergency response and reinforce impartial humanitarian aid; and nor will it make states accountable or responsible. This decision has added to the debate over creating "better aid", and Care International emphasises the importance of addressing the demand side, as well as reactive humanitarian aid. Mr Gareth Price–Jones of Care International notes the nexus between humanitarian aid and development aid, and addressing them together could more effectively address complex, long–term crises. (*IPS*, 9 May 2016)

▸▸ Turkish security forces have been accused by the UN and the group Human Rights Watch of committing serious human rights violations against Turkish civilians and Syrian refugees. Turkish security forces may have deliberately shot civilians, destroyed infrastructure, carried out arbitrary arrests, and caused displacements in an ongoing military campaign against ethnic Kurdish separatists in the country's southeast. A separate report from Human Rights Watch claim that Turkish border guards have shot and beaten Syrian asylum seekers. The UN said that many Kurdish–majority towns in the southeast have been sealed off "for weeks" and are almost impossible to access, and that there are reports of ambulances and medical staff being prevented from reaching the wounded. This is in the wake of a deal between the EU and Turkey to halt the flow of migrants to Europe, in exchange for aid and visa–free travel for Turkish citizens. (*Washington Post*, 10 May 2016)

UN AIDS and The Global Fund

▸▸ Cambodia's dispute with the Global Fund over travel expenses is now resolved, allowing the country to access millions of dollars in aid money to fight malaria – amidst fears over the rise of drug–resistant malaria. The dispute appears to have arisen over receipts for travel payments – these are difficult to obtain in rural Cambodia and are therefore not a requirement for government officials – and the Global Fund has agreed not to ask for these. However,

travel plans must be submitted in advance, spot checks can be carried out to verify travellers' locations, and staff will have to reimburse any "irregular" funds. The agreement means that Cambodia's National Malaria Centre (CNM) can now access a new US$ 12 million grant, plus another (almost untouched) grant of US$ 9 million. Although welcoming the resolution, CNM's director Dr Huy Rekol noted that Village Malaria Workers were not paid during the dispute and stopped alerting the authorities on local malaria cases, and this may be linked to some deaths from malaria. (*Phnom Penh Post*, 28 Jan 2016)

▸▸ The Global Fund plans to send an advance supply of antiretroviral drugs to Uganda, after the country ran out of supplies at the end of 2015. In Uganda, 1.5 million people – 1.5% of the population – are HIV positive. The shortages, which began in September 2015, affected 240 000 patients on publicly–funded treatment programmes, forcing them to modify treatment or stop outright. Private–sector clinics were unaffected. The government claimed that a weak currency and insufficient foreign exchange hindered its ability to finance drug imports. However, critics blame high election spending for the financial shortfall. The Global Fund acknowledged that the advance supply is a "short–term solution" and called for the government to mobilise resources to fill the gaps and find a long–term solution. (*Yahoo*, 25 Jan 2016)

▸▸ UNAIDS and Xinhua News Agency have signed an agreement to enhance global co–operation towards ending HIV–AIDS by 2030. The deal builds on an existing agreement from 2011, and new measures include strengthening collaboration in areas such as social media. The two sides will work towards this goal through in–depth co–operation, consultation and information exchange. Mr Michel Sidibé, the Executive Director of UNAIDS, said "combined with the power of media and communication, we could work together to build a legacy in promoting ending AIDS." Xinhua President Cai Mingzhao also stated that "to end AIDS requires the joint efforts from the international community." (*Xinhua*, 18 Mar 2016)

▸▸ Ahead of the UN General Assembly Special Session on the World Drug Problem, UNAIDS has released a report which shows that countries which do not adopt health– and rights–based approaches for drug–users experience no falls in HIV infections in people who inject drugs. Countries have implemented health– and rights–based approaches to drugs have reduced new HIV infections in these groups. Examples of successful programmes include the free voluntary methadone programme in China, Iran's integrated services for the treatment of sexually–transmitted infections, injecting drug users and HIV, and a peer–to–peer outreach programme in Kenya on using sterile equipment. A key part of ending the HIV epidemic by 2020 is reaching 90% of injecting drug–users with HIV prevention and harm reduction services. This would require an annual investment of US$ 1.5 billion in outreach, needle–syringe distribution and opioid–substitution therapy in low– and middle–income countries. However, these programmes are cost–effective and deliver wider benefits, such as lower crime rates and reduced pressure on health services. (*Merh*, 17 Apr 2016)

▸▸ Médecins Sans Frontières (MSF) has called for governments, UN and European agencies, PEPFAR and the Global Fund to develop and implement a fast–track plan to scale–up antiretroviral treatment (ART) for countries where coverage reaches less than 33% of the population, particularly in West and Central Africa. MSF warn that globally–agreed goals to halt the HIV epidemic by 2020 will not be met without this plan. In West and Central Africa–a region of 25 countries, 75% of people who require them cannot access HIV care–equivalent to 5 million people. "The converging trend of international agencies to focus on high–burden countries and HIV 'hotspots' in sub–Saharan Africa risks overlooking the importance of closing the treatment gap in regions with low antiretroviral coverage. The continuous neglect of the region is a tragic, strategic mistake: leaving the virus unchecked to do its deadly work in West and Central Africa jeopardises the goal of curbing HIV/AIDS worldwide", says Dr Eric Goermaere, MSF's HIV referent. (*Health24.com*, 20 Apr 2016)

UNICEF

▸▸ UNICEF has warned that 25 000 children are suffering from acute severe malnutrition in North Korea, and are in need to urgent treatment. It calls for US$ 18 million to support this, as part of a wider US$ 2.8 billion appeal to help 43 million children in humanitarian emergencies. In the wake of severe droughts that causes a 20% reduction in North Korea's crop production, UNICEF needs US $8.5 million for nutrition, US$ 5 million for water and sanitation, and US$ 4.5 million for health care to help these children. There are often shortfalls in humanitarian funding for North Korea – 70% of North Koreans suffer from food insecurity, funding fell from US$ 300 million in 2004 to under US$ 50 million in 2014 – due to its restrictions on humanitarian workers and concerns over its nuclear capa-

bilities. According to Ghulam Isaczai, the UN's resident co–ordinator for North Korea, "[North] Korea is both a silent and underfunded humanitarian situation. Protracted and serious needs for millions of people are persistent and require sustained funding." (*International Business Times*, 26 Jan 2016)

▸▸ On the eve of the International Day for Zero Tolerance of Female Genital Mutilation (FGM), UNICEF warned that growing populations in high–prevalence countries are undermining efforts to tackle the practice, which is widely regarded as a serious abuse of human rights. 50% of girls and women subjected to FGM live in Egypt, Ethiopia and Indonesia, and if current trends continue, the number of cases will increase over the next 15 years. Previously Indonesia has been excluded from UNICEF's FGM statistics due to a lack of reliable data – its recent inclusion has led to a sharp upwards revision in the number of global FGM victims – and other countries were FGM is reported are also omitted, such as India, Oman and the United Arab Emirates. However, countries such as Liberia, Burkina Faso and Kenya have experienced steep falls in FGM cases and condemnation is growing, and UNICEF calls for accelerated efforts to eliminate the practice. (*Thomson Reuters Foundation*, 5 Feb 2016)

▸▸ UNICEF estimates that one–third of combatants in Yemeni's civil war are children, on both the rebel side and troops fighting for President Abdullah Mansour Hadi. UNICEF believes that children as young as 14 are front–line fighters, despite pledges from both sides to end the practice. The massive destruction of schools and infrastructure encourages children to fight, and the rise of terrorist groups such as Isis and al–Shabaab makes negotiations over child combatants impossible. The situation in South Sudan is graver still, with 16 000 children being recruited into both sides in the country's civil war. Anthony Nolan, one of UNICEF's child protection specialists, says many children are driven to join by a lack of resources or a desire to seek revenge for their families, and that their recruitment threatens to prolong the conflict for future generations. (*The Independent*, 8 Feb 2016)

▸▸ On the 5th anniversary of the Syrian civil war, a UNICEF report shows the resultant refugee crisis, with over 2.4 million Syrian children living as refugees outside their country, 200 000 live as refugees within Syria – and a further 306 000 children were born as refugees. Over 250 000 people have died in the conflict – at least 400 children were killed in 2014. And now, twice as many people live in areas under siege or otherwise hard–to–reach compared to 2013, and 2 million of those cut off from help are children, with UNICEF reporting children suffering from extreme malnutrition or death from starvation. There are concerns over the increases in recruitment of child soldiers – both boys and girls, and children have reported being beaten, indoctrinated and forced to commit violence. The psychological effects of living under siege are also devastating. "Children living under siege almost have to re–learn what it's like to be a human being," says Mr David Nott, a trauma surgeon who has worked in Syria. (*Business Insider*, 14 Mar 2016)

▸▸ According to UNICEF, more than 700 million women were married before their 18th birthday, and Bangladesh has the world's second highest rate of marriage of girls aged under 15, after Niger. However, a study by the New York–based Population Council shows that child marriage fell by 31% when girls are educated or took classes in critical thinking and decision–making, with further falls when girls received job skills training. In Bangladesh, 75% of girls marry before they are 18 years old. "In Bangladesh, limited evidence exists on what works to delay child marriage. These results are a major leap forward," said Ann Blanc, Vice President of the Population Council. (*Thomson Reuters Foundation*, 23 Mar 2016)

World Health Organization (WHO)

▸▸ A WHO report published ahead of the first Ministerial Conference on Immunisation in Africa shows that Rwanda's immunisation coverage is 99%. This success is attributed to improving routine immunisation and the introduction of new vaccines. Dr Matshidiso Moeti, the WHO regional director for Africa, noted that Africa has increased vaccination coverage from 64% in 2004 to 79% in 2014. However, she urged further action from governments at the conference, because only 9 countries have immunisation coverage of 80% or higher, and 1–in–5 children in Africa do not receive basic vaccinations. "We have the tools, we need to save children's lives, and all we need is the political will and financial support to deliver," she said. Currently, GAVI funds immunisation in 70% of African countries, but as more African countries move from low– to middle–income status they will be ineligible for GAVI support, so must prepare to meet immunisation costs from their own budgets. (*allafrica.com*, 26 Feb 2016)

The WHO, in its role as pharmaceutical watchdog in markets with inadequate regulation, has suspended its approval of tuberculosis drugs made by India's Svizera Laboratories. The company is a major supplier to developing countries, and the move follows concerns over its manufacturing and quality standards. The WHO also recommended that batches of medicine already on the market should be retested by independent experts, and that supplies may need to be recalled. The company disagreed with the WHO's decision (which follows earlier warnings on standards in Svizera Laboratories, including dirty surfaces, black mould in a cleaning area, poor hygiene and inadequate record keeping), claiming the WHO had ignored evidence that Svizera's operations were up to standard. India's pharmaceuticals industry supplies cheap copies of generic drugs, but in recent years it has been beset by problems over the quality of its products. (*medicaldaily.com*, 19 Mar 2016)

The WHO's zika response differs markedly from its 2014 response to the Ebola outbreak. The WHO quickly flagged zika as a public health emergency – despite significantly fewer deaths – it took 5 months and nearly 1000 deaths before it declared Ebola a "public health emergency of international concern." Although the faster response – intended to jump–start scientific research, vaccine and treatment development, and mosquito control, may partly be caused by a wish to act quickly after criticisms over Ebola – the overall picture is more nuanced. For example, the WHO's regional office for the Americas (PAHO) had more expertise on zika compared to the Ebola expertise within the WHO's regional office in Africa, and PAHO came under pressure from the USA – which is more likely to be affected by zika than Ebola – to act decisively. Finally, the impact of zika can be presented in distressing images of newborn babies with microcephaly, whereas Ebola affected wider swathes of society, thus making it harder to press for action for a single group. (*Chicago Tribune*, 5 Apr 2016)

A WHO–led analysis published in *The Lancet Psychiatry* shows how the global failure to tackle depression and anxiety costs US$ 1 trillion each year in lost productivity and causes "an enormous amount of human misery." It found that without scaled–up treatment, 12 billion working days – 50 million years of work – will be lost to depression and anxiety disorders each year up to 2030. Scaling–up treatment would cost US$ 147 billion, meaning that every US$ 1 invested in treatment would lead to a US$ 4 return in better health and ability to work, and the authors argue that both developing and developed countries should improve mental health care. The study notes that common mental health conditions are increasing – the number of people suffering with depression and/or anxiety rose from 416 million in 1990 to 615 million in 2013. 10% of the world's population is affected, and mental disorders account for 30% of the global burden of non–fatal diseases. Treating these disorders would help the world meet the SDG of reducing premature deaths from non–communicable diseases by 33% by 2030. War and humanitarian crises increase this urgency, as the WHO estimates that up to 20% of people suffer from depression and anxiety during emergencies. (*The Guardian*, 12 Apr 2016)

In April, the WHO launched its global strategy to combat leprosy, with the overall aim of reducing to zero the number of children diagnosed with leprosy. Although the disease prevalence rates for leprosy fell to below 1 per 10 000 population in 2000, worldwide there are still 213 899 new cases a year – and India, Brazil and Indonesia account for 81% of these cases. Key interventions to combat leprosy include targeting detection amongst higher–risk groups via campaigns in highly endemic areas, and improving health care coverage for marginalised groups. Early detection, especially amongst children, is essential for reducing disabilities and transmission. (*livemint.com*, 21 Apr 2016)

Demography

In 2003, China relaxed its restriction on divorce, and divorce rates have risen rapidly–a remarkable transformation in a country where marriage was universal and permanent. This reflects the underlying transformation of Chinese society, with mass urban migration dividing couples, the improving status of women, and increased prosperity making it more feasible to live alone. Divorce is easier and cheaper in China compared to most other countries, and divorce rates are approaching levels in the USA. Often, falling divorce rates are a reflection of falling marriage rates and rising births outside wedlock, but neither hold true in China where marriage rates, including remarriage, remain high. Marital tensions can be worsened by the pressure on young Chinese people to marry early – and they often lack opportunities to meet suitable potential partners, and cohabitation before marriage is still rare (although increasing). Moreover, although women initiate over 50% of divorces, settlements still tend to favour men. (*The Economist*, 23 Jan 2016)

According to the UN Population Fund, 25% of the world's population is youthful (aged 10–24 years), with the vast majority living in the developing world. In India, home to the world's largest number of young, working–age people, 1 million people reach 18 years every month–and at 422 million, the number of people aged 15–34 years is equivalent to the combined populations of the USA, Canada and the UK. This is partly due to global successes in reducing child mortality, and more children being enrolled in school. This generation of young people is more likely to be educated, connected and ambitious – leading them to more mobile. Globally, the youth bulge is felt unevenly – in Germany, the median age is over 46 years, compared to 18 years in Nigeria. Governments face the staggering challenge of job creation to absorb these new workers – 40% of whom are either unemployed or working in insecure jobs at poverty pay. Youth unemployment is a strong predictor for social unrest, and to combat the uneven generation gap, countries with older populations need more migration from younger countries, alongside job creation in the global south. (*New York Times*, 5 Mar 2016)

The transformation of Medellin, Colombia's capital city, has been recognized by its winning the Lee Kuan Yew World City Prize, which is awarded by the Singapore government. This award honors outstanding urban achievements and solutions, and the Nominating Committee praised Medellin's transformation from uncontrolled urban expansion and violence, to its status as model for urban innovation. The committee made special mention of the political will, leadership, and long–term plans shown by the city's past three mayors, which have tackled security problems, aided economic development and improved its citizens' quality of life and employability. Medellin's current mayor, Federico Guitierez Zuluaga said, "this is an important recognition that we feel proud of for our city. We thank you for the encouragement to continue working for our city, a spectacular city that has come a long way but also has a long way to go." (*Cities Today*, 18 Mar 2016)

Japan's extreme demographic challenges follow on from 20 difficult years for the country, which has been beset with deflation, budget deficits and high public debt. The first challenge is Japan's rapidly shrinking and aging population, with the share of people aged over 65 years rising from 5 in 1950 to 25% in 2012 – the highest figure in the world, and its median age is 45.9 years, compared to the global median of 29 years, and 38.7 years for other OECD countries. This trend is set to continue, while the country's population is expected to decrease by 22–23%. This is compounded by Japan's low fertility rates, partly as long working hours and high population density in urban areas discourage women from having children. However, Japan's overseas–born population was 1.7%, significantly below other OECD countries, so migration is an untapped potential for population growth. Also, carefully managing longevity gains (eg, healthier aging, longer careers, efficient health care) can help offset the economic impacts of aging. (*OECD*, 11 Apr 2016)

Life expectancy continues to increase in the UK, and men's life expectancy is slowly catching up with women's. However, underlying these very welcome developments is a new, and worrying trend – for the first time since the 19th century, the narrowing of the gap in life expectancy between rich and poor people has reversed. Increases in life expectancy before 1940 were due to clean water, improved sanitation, affordable housing and cleaner air, among other things–and this benefitted poorer people as wealthier people could afford healthier environments and lifestyles. Post–1950, advances in health care tended to benefit all sections of society equally, so the life–expectancy gap remained constant. However, this began to change in the 1990s, mainly due to lifestyle factors (eg, diet, exercise, alcohol and tobacco use) which increase the risk of non–communicable diseases. For example, previously smoking was evenly spread throughout all social groups, and 82% of UK men smoked in 1948. This has now fallen to 21%, and poorer men are more likely to smoke. The rich–poor divide in unhealthy lifestyles is widening the life – expectancy gap, but tackling individual lifestyles may be more difficult than wider public health measures like cleaning up water supplies. (*New Scientist*, 3 May 2016)

Economy

Small states such as Botswana, the Seychelles and Mauritius rank the highest among African states in most human development indices. This is highly beneficial to their own citizens in terms of health, prosperity, safety and good governance, but has little impact beyond their own borders. States which have larger impact – eg, Nigeria, Kenya and South Africa – do not fare so well. If Kenya's economic dynamism translated into GDP levels equivalent to Mauritius, oil–rich Nigeria's governance matched Botswana's, and South Africa's post–apartheid moral stature continued to endure, this could have profound impact on their neighbors and Africa in terms of development, peace and security. Although each country falls short of its own potential, some indicators (eg, Nigeria's free elections in 2014, Kenya's entrepreneurism and South Africa's international standing) hint at their potential to lift the entire continent. Improving linkages between these countries and their neighbors – politically, economically and culturally–is vital to ensuring the success of the pan–African Union, and the East African Community (Africa's most integrated regional bloc) showcases the benefits of integration. (*dailymaverick.co.za*, 22 Jan 2016)

Underneath the public health concerns raised over the zika virus as it spreads among 26 countries in the Americas, lies a quieter question of economic loss and hardship caused by the virus. While it is too early to give a definitive answer on the potential financial havoc wrecked by the illness, estimates can be gleaned from the impact of another illness spread by the Aedes mosquito–dengue fever. Donald Shepard of Brandeis University estimates that the 2013 outbreak of dengue fever cost the global economy US$ 8.9 billion, with the largest burden shouldered by developing countries. This does not include the impact on tourism, and already shares in travel companies are falling as pregnant women and those planning pregnancies are advised against visiting affected countries. This will hit Brazil especially hard, as it hosts its annual Carnival festivities and anticipates 500 000 tourists for the August 2016 Olympic Games. (*Bloomberg View*, 5 Feb 2016)

The International Monetary Fund (IMF) has signaled that it will downgrade its outlook for the global economy in April. This follows recent warnings from the OECD that global economic growth will slow within the next few months. However, there are few signs that governments are heeding these warnings – indeed, a senior US Treasury official commented that it is not reasonable to expect a crisis response over economic uncertainty. The IMF is concerned that any announcement of joint action by the G20 group of leading economies could risk the global economic expansion. Part of the problem is a shortage of tools to deal with the next economic crisis, and the IMF wants the G20 to boost spending, delay interest rate rises and for the European Central Bank to boost stimulus efforts. (*Wall Street Journal*, 8 Mar 2016)

Mr John Mangudya, the governor of Zimbabwe's central bank, confirmed that the country expects its first International Monetary Fund (IMF) loan since 1999 later in 2016, after paying off foreign lenders. The exact amount has yet to be agreed, but the IMF has agreed to double the amount available to Zimbabwe to US$ 984 million. Zimbabwe is trying to emerge from international isolation, which are largely blamed on its government's policies. Its worst drought since 1994 has left 4 million Zimbabweans facing hunger, and forced the government to lower its economic growth forecast from 2.7% to below 2% – the IMF and World Bank forecast growth of 1.4% and 1.5% respectively. As part of the loan agreement, the government agreed to major reforms, including compensation for evicted farmers and a reduction in public sector wages. Once Zimbabwe's arrears are cleared, it will be ready for rating by international ratings agencies with the aim of issuing bonds. (*Reuters*, 16 Mar 2016)

China's foreign capital reserves fell by US$ 28.57 billion in February, marking the fourth consecutive monthly decline, albeit at a decelerating rate. The government is examining a range of measures to curb speculative foreign transactions. These measures include a levy designed to penalise short–term currency speculation (the so–called "Tobin tax"), and imposing fees on the sales of forward positions. "We are considering policies to increase the costs of short–term speculation as long as they don't affect normal capital flows," says Mr Wang Yungui, the head of regulation at China's State Administration of Foreign Exchange. (*WSJ*, 22 Mar 2016)

Energy

1.1 billion people – mainly in southern Asia and sub–Saharan Africa – lack electricity, and electrification is barely keeping pace with population growth in the latter. Some entrepreneurs are using technology to provide ac-

cess to clean, cheap systems which are metered and paid for by mobile technology. The user by–passes electricity grids, instead harvesting solar energy from rooftop panels that are connected to batteries that store energy until nightfall. This is more suitable for rural dwellers with low energy needs and high grid connection costs. Urban areas could benefit from pre–paid meters topped up by mobile phone – and higher revenues would encourage more investment. There is tension between countries electrifying using cheaper fossil fuels and risking severe pollution problems akin to Beijing and New Delhi, against clean energy that may be more expensive and intermittent. However, regional transmission networks that share power, alongside "baseload" energy systems (geothermal, hydro, natural gas) that operate constantly can support business needs, while smaller–scale solar power to cover households' energy needs could form part of the solution. (*The Economist*, 27 Feb 2016)

▸▸ According to the news agency, Xinhua, China's gas consumption rose by 3.5% in 2015–the smallest increase in a decade, and lower than the official 5.7% forecast. Gas for domestic usage was the largest component (69%, or 131.8 billion m^3) of overall consumption. This is the second year of sluggish growth in gas consumption (partly caused by weaker economic growth), despite government efforts to promote gas over coal as a cleaner energy source. These figures bring China's gas market to below the International Energy Agency's long–range growth forecast of 4.7%. The downturn is adding to a glut of gas supplies, with current prices lower than contract prices, and firms are attempting to postpone gas shipments. Australia, which has invested heavily in gas projects in China, has been badly affected by the downturn, leading to Western Australia's credit rating being downgraded by Moody's. The Chinese government has attempted to raise demand by slashing prices, but this has deterred domestic production and led to even lower output. (*Radio Free Asia*, 29 Feb 2016)

▸▸ Since industrialisation 250 years ago, humans have released 500 billion tonnes of CO_2 into the atmosphere from fossil fuels and deforestation–and is set to release another 500 billion tonnes over the next 40 years. 50% of a CO_2 increase is removed from the atmosphere within 30 years, 30% is removed within a few centuries, and the remaining 20% may remain for millennia. Carbon capture and storage (CCS)–a technology which captures CO_2 at emission and stores it underground, eg, in depleted oil and gas reservoirs–could be a solution. The International Energy Agency (IEA) estimate that CCS is the most effective way of reducing CO_2 emissions by 13% by 2050, but note that progress in CCS implementation is slower than hoped. There is a current wave of 22 CCS projects worldwide, and they will collectively capture 48 million tonnes of CO_2 each year from coal–fired power stations, gas processing and other industrial processes. However, there are few further CCS projects in the pipeline, and the world risks losing momentum without policy intervention. The IEA believes that the world has more than enough CO_2 storage resources, but more investment in exploration and development is required. (*OECD*, 20 Apr 2016)

▸▸ Venezuela has been forced to cut its national power supply for 4 hours daily, to last over the next 40 days – due to reduced rainfall that drives electricity–generating turbines in hydroelectric dams. The cuts will affect 10 out of 23 states, including major cities. This is an additional blow to the country's citizens, who already face shortages of food and medicine. However, the oil sector is unlikely to be included in the cuts thanks to its importance to the Venezuelan economy. The economy is already struggling with falling oil prices, and is set to shrink by 8% in 2016, and the power shortages will cause further economic damage. (*Newsweek*, 22 Apr 2016)

▸▸ In an interview shortly before his death, Prof Sir David MacKay–the UK's former chief scientific adviser–called the idea of renewable energy powering the UK "an appalling delusion." He believes that solar, wind and biomass energy would require too much land, huge battery back–ups and cost too much to be viable options for the UK, although he believes that solar power has great potential in hot, sunnier countries. He notes that renewable energy – which produces 1% of the UK's electricity–cannot sufficiently be scaled up. Instead, he calls for the UK to focus on nuclear energy and carbon capture storage technology to reach zero carbon emissions. He believes that carbon capture and storage is essential in tackling climate change, and is disappointed by the UK's lack of progress with the technology. (*The Guardian*, 3 May 2016)

Environment

▸▸ As President Obama's last term draws to a close, the US Supreme Court put on hold his administration's Clean Air Act. This Act – seen as a key legacy of Obama's administration – is designed to reduce emissions from power plants by 32% by 2030, and is the main tool for the USA to meet its emissions targets agreed at December's UN climate talks.

Various business groups and 27 states (led by West Virginia and Texas) launched the bid to block the Act, arguing that it would devastate their economies. The decision means that the regulations will not take effect while court battles continue over their legality, and raises doubts over the long–term future of the Environmental Protection Agency. "We are thrilled that the Supreme Court realised the rule's immediate impact and froze its implementation, protecting workers and saving countless dollars as our fight against it continues," said Mr Patrick Morrisey, the Attorney General for West Virginia. (*Scientific American*, 9 Feb 2016)

▸▸ Mongolia is currently experiencing a *dzud* – a natural disaster which occurs when a summer drought is followed by heavy winter snowfall that makes already scarce pastures inaccessible to livestock. Previously, dzuds occurred once a decade, but they have recently been occurring every few years. It is believed that the increasing frequency could be due to a combination of climate change – the average temperature in Mongolia has risen by 2.1 °C since 1940 and Mongolia is ranked the 8th most vulnerable country to the impact of climate change – and human activity. In a country where 50% of people rely on livestock production, oversupply of animal products has led to falling prices – and increased animal numbers to maintain incomes. When combined with climate change, this has had a devastating effect on Mongolia's pastoral land, with over 70% being degraded, and increasing forest fires has reduced forest area by 0.46% each year – and threatening Mongolia's ancient way of life. The Ministry of Foreign Affairs has asked for international help to deal with the dzud – an estimated US$ 4.4 million is needed for emergency vehicles, medicine, food and livestock supplies – but the government has stopped short of declaring a state of emergency. (*IRIN*, 7 Mar 2016)

▸▸ A study by the Frankfurt School of Finance and Management for the UN Environment Programme shows that a record US$ 286 billion was invested in renewable energy in 2015 – more than double the investment in fossil fuels. In another milestone, developing countries' investment in renewable energy – US$ 156 billion – outstripped the US$ 130 billion investment by developed countries for the first time. China and India have led the way in developing countries' investment in renewables, and the USA increased its investment by 19%. However, Europe's investment in renewables fell by 21% in 2015, despite being a previous trailblazer. These investments are beginning to have an impact on climate change, and the International Energy Agency has pinpointed their growth as the main reason why global CO_2 emissions have been stable for 2 years, despite 6% economic growth. Without them, annual CO_2 emissions would be 5% – or 1.5 billion tonnes – higher. There are concerns that in the short–term, low prices for fossil fuels could spark a surge in their use, although pledges made at December's climate summit in Paris should limit this, and it is not expected to have a lasting impact. (*New Scientist*, 24 Mar 2016)

▸▸ Hokeng Metal Processing Co.'s industrial plant, located in Nonthong Village in Lao, appears to be pumping polluted waste water into the neighboring area. The plant extracts valuable copper, lead and other valuable minerals from discarded electronics, and re–sells them worldwide. Water contamination is 16 times higher than normal levels. Although Lao's Ministry for Natural Resources claim they are working on the plant's pollution management, no formal action has yet been taken against the company. The company may also have failed to comply with international regulations on transferring hazardous materials from developed to less developed countries. (*Radio Free Asia*, 7 Apr 2016)

▸▸ Cycling and walking are normally healthy activities, but a study published in *Preventative Medicine* found that air pollution in heavily polluted cities – such as Delhi in India, Karachi in Pakistan and Doha in Qatar – means that the harms can outweigh the benefits. In Delhi, cycling is only beneficial if people cycle for less than 5 hours a week, so if people's daily commute is longer than 30 minutes each way, cycling will damage their health. People who cycle for longer each day, eg, bike couriers, will experience even more harm. Air pollution is strongly linked to heart disease and lung cancer, and causes thousands of death a year – and exercise can intensify its harmful effects as heavy breathing causes more dirty air to be drawn into the lungs. (*New Scientist*, 5 May 2016)

Food, Water and Sanitation

▸▸ Following floods in 2015, the northern state of Arakan in Myanmar has seen a sharp rise in the number of severely and moderately malnourished children. The flooding – caused by heavy rain and Cyclone Komen – has destroyed crops and rice paddies and has contaminated water sources. The numbers of malnourished children aged under 5 years seen by a European Commission food program in Maungdaw district rose to 1500 in October, compared to 500 in July. The children are eating fewer and less diverse foodstuffs, sometimes relying on rice and water only. The

real number is likely to be much higher, as the food programmes only see the minority of affected children. 90% of Arakan's population belong to the Muslim minority Rohingya, who face violence and discrimination with no legal recognition of their citizenship, and in 2015, 14 000 of the state's children were admitted to this program, including 10 900 children aged under 5 years. (*Irrawaddy*, 28 Jan 2016)

▸▸ In 2015, Papua New Guinea experienced a severe drought, arising from the El Niño weather system, followed by floods and mudslides after heavy rains in February 2016. According to the UN Office for the Co–ordination of Humanitarian Affairs, 480 000 people are facing critical food shortages and are in need of food aid. The extreme weather conditions has also caused health care facilities to close, or operate at lower capacity due to water shortages or problems with storage. Aid agencies are working with the government to distribute food and monitor dengue outbreaks in Daru, and possible cases in Kiunga. (*Reuters*, 4 Mar 2016)

▸▸ About 60% of Africa's farms are less than one hectare, and agriculture employs more than 50% of people in sub–Saharan Africa. Therefore, improving agricultural productivity is one of the best ways to raise Africa's living standards – and its farms are less productive than Latin American and Asian farms. The value of Africa's agriculture has increased by 400% since 1961, albeit by bringing more land into cultivation rather than improving yields, so output per person actually fell. There are inherent difficulties in increasing yields due to infertile soils, and varying climate patterns make crops more heterogeneous and less amenable to a "green revolution" – unlike Asia's staple crops of rice and wheat. Inadequate roads, price controls and corruption over subsidies to poor farmers also hinder agricultural development. However, hybrid seeds are improving yields, the gradual lowering of tariffs encourage exports, and land reforms give more control to women farmers. This gradual brightening is underpinned by better governance and fewer conflicts across the continent. Improved roads, information on market prices, better storage and food processing to aid diversification and job creation, better husbandry to boost livestock production would all further support Africa's agricultural evolution. (*The Economist*, 12 Mar 2016)

▸▸ According to the charity Water Aid, India has the world's highest number of people without access to clean water, with 75.8 billion people – or 5% of India's population – being forced to buy expensive water or use contaminated supplies. Buying water costs up to US$ 0.72 a day – nearly 20% of a poor person's income – and diarrhea kills 140 000 children in India each year. India's water problems are set to worsen, as rivers become more polluted, groundwater reserves diminish, and global warming causes more erratic rainfall. It is predicted that India can only meet 50% of its water needs within 15 years. Indian states are experimenting with measures to improve water supply and management, including privatization, water filtration units and water kiosks in drought–prone areas. Water shortages could cause tensions, and Satya Tripathia, an advocate in India's Supreme Court says "the government really has to pay attention. Water is the one thing that can tear this country apart." (*Yahoo*, 22 Mar 2016)

▸▸ On the first anniversary of the earthquake which struck Nepal on 25 April 2015, thousands of people across the country still face problems of water, sanitation and hygiene. Swift action by the government, community health workers and aid organisations immediately after the earthquake prevented disease outbreak, but more aid and delivery is urgently required to ensure that rebuilding happens as quickly as possible. Basic infrastructure was badly damaged by the earthquake, and natural springs – a major source of water for rural Nepalese people – produce less water, or have dried up completely. Some communities that previously had continuous access to water only have access for 1–2 hours a day, and many people are having to wash in rivers or queue for hours at taps. The aid organization, WaterAid, is working with local partners and communities to find new spring sources and build water supply infrastructure. (*The Himalayan*, 26 Apr 2016)

Peace and Human Rights

▸▸ Crime rates in Japan are exceptionally low by international standards, and those arrested for minor crimes are treated with leniency – less than 5% of those found guilty of a penal offense are sent to prison. Japan emphasizes rehabilitation, and has extremely low rates of re–offending. However, 99.8% of prosecutions end in a guilty verdict, and the system relies heavily on confessions – ie, 90% of criminal prosecutions. There are few safeguards for suspects being questioned, as they can be held for 23 days without charge, often with little contact with a lawyer. Few interrogations are recorded, and although physical torture is rare other methods such as sleep deprivation are not. Moral blackmail (eg, citing shame brought on family), and fabricating confessions and pressurising the suspect to sign

them can happen. Once in court, the non–adversarial nature of the trial system means that judges seldom question whether confessions are voluntary. According to one estimate, 10% of all convictions leading to a prison sentence are based on false confession. There have been recent miscarriages of justice – a mother convicted of murdering her daughter was released after her innocence was proved by a crime reconstruction; and Iwao Hakamada was freed after 46 years on death row after his conviction was declared unsafe – he appears to have been tortured when arrested. (*The Economist*, 5 Dec 2015)

▸▸ Latin America has consistently been home to 86% of the world's most violent cities, according to data published since 2011 by the Mexican Citizens' Council for Public Security – an NGO whose annual survey assesses the world's 50 most violent cities. It found that not only does Latin America have the highest number of violent cities, but violence is much more widespread. Of the 43 Latin American cities on the 2014 list, 40% have homicide rates greater than 50 per 100 000 people, and 46% have rates of 30 per 100 000 people, or more – the global average is 7 per 100 000 people. Some cities such as Juarez (Mexico), San Juan (Puerto Rico) and Medellin (Columbia) have seen sharp falls in violence, and the number of Mexican cities in the list has fallen from 25 to 2. However, the number of Brazilian cities has increased from 14 to 19, and violence has been resurgent in El Salvador after the gang truce broke down. (*insightcrime.org*, 22 Jan 2016)

▸▸ There are reports that Egypt's government is increasingly using the tactic of "enforced disappearance" to crack down on real and imagined opponents. The "disappeared" are not held in the formal legal system – which had already detained thousands of people – but are moved into a network of secretive detention centers, run by the security forces. They are held without charge or access to a lawyer, where their isolation and lack of legal protection enables them to be interrogated harshly – many say they have been tortured – and forced to identify friends and relatives. Detainees are usually released without charge within months, or charged with membership of the outlawed Muslim Brotherhood. But some are missing for much longer, and the dead bodies of others are dumped in morgues. The disappeared include members of the Muslim Brotherhood, but also civil society activists, journalists and members of the public unwittingly caught up in the state's security dragnet. "The goal seems to be to terrorise society, to show that anyone who dares criticise the government will face a similar fate," said Mohamed Elmissiry, a researcher with Amnesty International. Public disquiet has grown over this crackdown, leading to an investigation of the cases of 101 missing people. However, lawyers and human rights groups believe that the investigation will be a whitewash, as the government has already declared that detainees were legally arrested, joined militant groups or had fled Egypt. (*Irish Times*, 27 Jan 2016)

▸▸ In a landmark ruling, the International Criminal Court (ICC) at The Hague has found a warlord guilty for perpetrating rape as an act of war. It also secured a conviction for "command responsibility", which means that a commander can be found guilty of crimes if he or she orders them, even without directly taking part. This verdict was delivered against Jean–Pierre Bemba, head of the Movement for the Liberation of the Congo, who sent his militia into the Central African Republic (CAR) to rampage during a period of turmoil. More than 5200 victims testified that they had been sexually assaulted or their property stolen. Although the court has faced accusations of bias against Africa, it was an African government – the former CAR government – which referred Mr Bemba to the ICC. "The facts have shown that rape was systematic, as was pillage, and was perpetrated in a humiliating way, anywhere, anytime by multiple rapists," said Ms Marie–Edith Douzima–Lawson, the victims' legal representative. Sentencing has yet to be carried out. (*The Economist*, 22 Mar 2016)

▸▸ There is a lack of reliable data on gender violence in Cambodia, but a UN report from 2013 found that 22% of women had experienced violence from a male partner, although only 16% of men admitted violence toward a woman. In addition, 96.2% of men and 98.5% of women believe that a woman should obey her husband, and 67% of women believe they should tolerate violence to maintain the family. This attitude reflects the teachings of the Chbab Srey, a poem on women's role which teaches submission, and which was part of the Cambodian school curriculum until 2007. It is also a legacy of the Khmer Rouge, where an unknown number of women were forced into sex work for survival, into marriage, or were victims of sexual violence. Most cases of domestic violence go unreported, partly due to Cambodia's skeletal judicial system and lack of victim support, although tradition – which emphasizes virginity as a marriage pre–requisite – also causes victims' silence. Rape is also prevalent in Cambodia, with 20% of men admitting to at least one rape – and 38.4% of these men were unpunished. However, the Asia Foundation has funded the development of mobile solutions by women's networks, such as apps to explain the causes and risk factors behind domestic violence, give the names of support organisations, and to file reports anonymously. This is a small step toward making Cambodia safer for women. (*The Diplomat*, 18 Apr 2016)

Science and Technology

At an expert panel moderated by the US Vice President Joe Biden at the World Economic Forum in Davos, the discussion focused on the need to collect, harness and analyze Big Data to finally find a cure for cancer. There has been a huge increase in the volume of oncological data, but researchers' ability to use it has not kept pace. The panelists highlighted the following obstacles to fully realizing the potential of Big Data. First, medical data are not standardised across platforms, and standardising it would increase the volume of data available to scientists working on specific problems. Second, patients' concerns over privacy must be overcome. Third, few cancer patients – only 5% – volunteer for clinical trials (despite such trials offering hope to those who are have otherwise no options left) and most of those who do are not given access to their data. Mr Biden said he is dedicating his last year of office to a "cancer moonshot", and believes that scientists are on the cusp of a breakthrough in cancer treatment. However, he acknowledges that the Big Data challenge is a fundamental challenge, and solving it will require unprecedented co–operation between professionals across many disciplines. (*Forbes*, 26 Jan 2016)

According to a study published in *The Lancet Infectious Diseases*, more than 50% of HIV–positive people who are not responding to treatment have an HIV strain which is resistant to tenofovir, a key antiretroviral drug. 60% of HIV–positive people in Africa have become tenofovir–resistant, compared to 20% in Europe. Second–line drugs are available, but these are generally more expensive and have more side effects. Resistance to tenofovir can be caused by the drug regime being incorrectly followed, or the HIV–positive person becoming infected with another, tenofovir–resistant, strain of HIV. Surveillance, treatment and monitoring of HIV patients must be improved in the wake of this development, and studies are under way to determine how HIV developed tenofovir resistance. (*Tech Times*, 29 Jan 2016)

The Global Trachoma Mapping Project has recently completed a global survey of trachoma – a painful [preventable], neglected tropical disease which can cause blindness – and the scale and quality of the survey means that it could be eliminated by 2020. Ethiopia highlights the advantages of disease mapping – prior to the survey in 2012, only one district had support for tackling trachoma, but now the entire country has been mapped and there is funding and support to deliver the required interventions for the entire country. Following the survey, many areas have been enrolled to receive free antibiotics donated by Pfizer via the International Trachoma Initiative. The software app developed for the survey is also being trialed for other diseases, including schistosomiasis and guinea worm disease. (*Thomson Reuters Foundation*, 10 Feb 2016)

The pharmaceutical company Sanofi has assembled a team of more than 80 experts to start pre–clinical tests of a potential zika vaccine in animals, with the first human trials likely in 2017. Other companies such as Bharat Biotech, Inovio and the US National Institutes of Health are also working on vaccines, but Sanofi has the advantage of being a major vaccine producer, and the first company to develop a vaccine for the related dengue virus. This could speed up development by several years and simplify safety, as the "backbone" of the virus is already in use. Developing a zika vaccine is potentially simpler compared to diseases like HIV, as the virus's genetic code is more than 95% the same across samples. However, designing clinical trials could be complicated as pregnant women are often excluded until the drug's safety is well–established, and ultimately the vaccine may be given to different age groups. (*Fortune*, 4 Mar 2016)

Wei Zexi, a 21–year old student suffering from synovial sarcoma – a rare form of cancer – died after using an experimental cancer therapy he found online. He sought the treatment from a hospital that came top of his Baidu rankings. Baidu is China's largest internet search engine, with 70% market share and more than 660 million people use its mobile search each month. Baidu has already faced criticism for selling listings to bidders without thoroughly checking their claims. Baidu marks its paid–for listings with "promote" in small text, but many argue that this does not adequately identify paid–for listings. Before he died, Wei accused the hospital of misleading him and his family over the treatment's effectiveness, and criticized Baidu for selling medical search listings to the highest bidder. Baidu denies ranking hospitals in paid–for search results based on payment, although investigations have been launched into the hospital. The Chinese government will carry out an official inquiry into Baidu's "search results for sale" business model, and will make the findings public. (*BBC*, 3 May 2016)

Council of Science Editors' award for JoGH's Editor–in–Chief Ana Marušić

Photo: Prof. Ana Marušić

Our Editor–in–Chief Prof. Ana Marušić won the prestigious Meritorious Award from the Council of Science Editors. Prof. Marušić received the Meritorious Award from the Council of Science Editors (CSE) at its 2016 Annual Meeting in Denver, Colorado, USA. The CSE is an international membership organization for editorial professionals publishing in the sciences, serving the scientific, scientific publishing, and information science communities. The Meritorious Award is the CSE's highest award, recognising those who embrace the purpose of the CSE: improving scientific communication through the pursuit of high standards in all activities connected with editing. CSE also aims to foster networking, education, discussion, and exchange, and acts as a resource on current and emerging issues in the communication of scientific information.

Ana Marušić joins such previous recipients as ORCID, COPE, CrossRef, and her fellow EQUATOR Network steering group member Doug Altman. Prof. Marušić has been active in the publishing and editing world throughout her career. She was the Editor–in–Chief of the Croatian Medical Journal for 20 years and is the Founder and Co–editor–in–Chief of the *Journal of Global Health* (JoGH). She is a past president of both the World Association of Medical Editors and the CSE, and is vice president of the European Association of Science Editors.

Ana Marušić is Professor of Anatomy and Chair of the Department of Research in Biomedicine and Health at the University of Split, School of Medicine in Croatia. She is also a Visiting Professor at the Centre for Global Health Research at the University of Edinburgh. She has been involved with the EQUATOR Network since 2010. Whilst her biomedical research focuses on the interactions between the immune and bone systems, her research interests also include peer review and research integrity, and she is committed to research quality, integrity, and transparency. To this end, she founded the Croatian branch of the Cochrane Collaboration and created the first Croatian public registry of clinical trials.

Thomson Reuters lists our Editors–in–Chief Harry Campbell and Igor Rudan among "The World's Most Influential Scientific Minds in 2015"

In their annual publication that tries to identify the scientists amongst the estimated 9 million researchers in the world whose work has earned distinction in the eyes of the scientific community, Thomson Reuters – the large international publishing company behind the most respected scientific citation index, *Web of Science* – identifies a tiny fraction of the authors whose work has consistently resulted in an outsized influence in the form of citations from fellow scientists. Each year the company produces a list of the leading such hundred authors in 21 different areas of science, who are officially designated as "Highly Cited Researchers". Inclusion to this global rank of researchers is based on the number of papers that have been among the 1% most cited in their respective fields in the previous 10 years. This year, this highly prestigious list has included two of JoGH's Editors–in–Chief: Prof. Harry Campbell and Prof. Igor Rudan, who have both been selected based on their research in the field of Molecular Biology and Genetics in the previous decade.

Photo: Prof. Harry Campbell (left) and Prof. Igor Rudan (right)

The Royal Society of Edinburgh elected the JoGH's Editor–in–Chief Igor Rudan as a Fellow

Established by Royal Charter in 1783 by key proponents of the Scottish Enlightenment, the Royal Society of Edinburgh (RSE) serves as Scottish National Academy that admits Fellows from a wide range of disciplines. The work of the RSE includes awarding research funding, leading on major inquiries, informing public policy and delivering events across Scotland to inspire knowledge and learning. This year, RSE admitted the JoGH's co–Editor–in–Chief, Prof. Igor Rudan, as a Fellow of the Royal Society of Edinburgh (FRSE).

Prof. Igor Rudan is particularly credited for his contributions to reduction in global child mortality in the 21st century through generating critical evidence that was required for developing successful health policies, and for developing novel methods for prioritizing investments in global health and development that have been widely used by international organizations. In his efforts to reduce global child mortality, he served as a consultant of the World Health Organization, UNICEF, The Bill and Melinda Gates Foundation, The World Bank, Save the Children and others. He also founded the biobank in isolated populations of Croatian islands, which contributed to the discovery of biomedical role for more than 1000 human genes to date. He joined the University of Edinburgh in 2001. He has published 400 research papers and 7 books focused on global maternal and child health and genetic basis of human disease. He has been awarded 20 national and international research awards and professional recognitions.

Photo: Prof. Igor Rudan's Fellowship of the Royal Society of Edinburgh

Meeting Report: Harmonization of RSV therapeutics – from design to performance

Harish Nair[1], Octavio Ramilo[2], Irmgard Eichler[3], Eric Pelfrene[3], Asuncion Mejias[2], Fernando P Polack[4,5], Koen B. Pouwels[6,7], Joanne M. Langley[8], Marta Nunes[9], Nicoline van der Maas[10], Leyla Kragten–Tabatabaie[11], Eugenio Baraldi[12], Terho Heikkinen[13], Brigitte Fauroux[14], Mike Sharland[15], Cyrus Park[11], Paolo Manzoni[16], Nikolaos G. Papadopoulos[17,18], Federico Martinón–Torres[19], Renato Stein[20], Louis Bont[21], in collaboration with Respiratory Syncytial Virus Network (ReSViNET)

[1] Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh Medical School, Edinburgh, UK
[2] Paediatric Infectious Diseases, Nationwide Children's Hospital, and The Ohio State University, Columbus, Ohio, United States of America
[3] European Medicines Agency, London, UK
[4] Fundacion INFANT, Buenos Aires, Argentina
[5] Department of Pediatrics, Vanderbilt University Medical Center, Nashville, Tennessee, USA
[6] Modelling & Economics Unit, Public Health England, London, UK
[7] Unit of PharmacoEpidemiology & PharmacoEconomics, University of Groningen, the Netherlands
[8] Canadian Center for Vaccinology, Dalhousie University, IWK Health Centre and Nova Scotia Health Authority, Halifax, Canada
[9] Department of Science and Technology/National Research Foundation: Vaccine Preventable Diseases & Medical Research Council: Respiratory and Meningeal Pathogens Research Unit, University of the Witwatersrand, Johannesburg, South Africa
[10] Department Epidemiology and Surveillance of the National Immunisation Programme, CIb–RIVM, the Netherlands
[11] Julius Clinical, Zeist, the Netherlands
[12] Women's and Children's Health Department, Unit of Respiratory Medicine and Allergy, Padova, Italy
[13] Department of Pediatrics, University of Turku and Turku University Hospital, Turku, Finland
[14] Noninvasive ventilation and Sleep Unit, Necker Pediatric University Hospital, Paris Descartes University, Paris, France
[15] Paediatric Infectious Diseases Research Group, St George's University London, UK
[16] Neonatology and Neonatal Intensive Care Unit, S Anna Hospital, Torino, Italy
[17] University of Manchester, Manchester, UK
[18] Allergy Dept 2nd Pediatric Clinic, University of Athens, Athens, Greece
[19] Translational Pediatrics and Infectious Diseases, Pediatrics Department, Hóspital Clínico Universitario de Santiago de Compostela, University of Santiago, La Coruña
[20] Pediatric Pulmonology Unit, Pontifícia Universidade Católica RS, Porto Alegre, Brazil
[21] Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, Netherlands

RSV is a major cause of morbidity and mortality worldwide. Although no treatment or vaccine currently exists, RSV therapeutics and preventative strategies are being evaluated in clinical trials, including phase 3 trials. Despite great prospects, the regulatory pathways of novel RSV therapeutics have been defined insufficiently. Here we report the results from the ReSViNET 2nd High–level expert meeting 2016 on RSV therapeutics, which was held in Zeist, the Netherlands on March 2nd and 3rd. This meeting was organized to advance discussion on regulatory pathways, clinical development, clinical trials, and health technology models in the RSV therapeutics field. During this meeting regulators, public health specialists, academia, non–governmental organizations and pharmaceutical companies openly discussed and addressed the needs for the successful development of RSV therapeutics and prophylaxis.

THE NEED AND THE OPPORTUNITY

RSV infection is one of the leading causes of acute lower respiratory infection (ALRI) related hospitalization and mortality during early childhood [1]. The current burden estimates are based on limited data. To overcome this data gap and to inform policy for introduction of RSV vaccine (which appears likely in the next 5–7 years), the Bill and Melinda Gates Foundation (BMGF) funded the RSV Global Epidemiology Network (RSV GEN), a platform to bring together RSV researchers from low and middle income countries to share unpublished data from ongoing / recently completed studies. This network has contributed data from more than 75 sites for the revised RSV burden estimates for 2015 which were presented at the meeting by *Dr Nair*. It is anticipated that these results which are of huge interest to clinicians, donor agencies and policy makers will be published soon.

The 2016 ReSViNET meeting aimed to discuss the development of different RSV therapeutics, the existing hurdles and strategies to overcome these barriers. *Dr Ramilo* provided an overview of the different target populations for therapy against RSV–infection. Current therapies are limited to specific populations. Prophylaxis with anti–RSV monoclonal antibodies is directed to high–risk populations only including preterm infants, children with bronchopulmonary dysplasia and those with congenital heart disease (CHD), as well as selected children with cystic fibrosis (CF) and immunocompromised conditions. However, the great majority of patients with RSV infection including children hospitalized with severe lower tract infection, and those with mild disease managed as outpatients are treated symptomatically. Not much is known yet about the burden of RSV disease among the elderly and in individuals with Chronic Obstructive Pulmonary Disease (COPD) and there are no RSV–specific therapies for these individuals. Ramilo gave a review of vaccines in development, mentioning the pros and cons of the different vaccines strategies and he also introduced antivirals and monoclonal antibodies (Mabs) in clinical development [2]. In the presentation, the need to optimize the design of clinical studies aimed at developing both treatment and preventive strategies against RSV was discussed. He reviewed how to best select the patient populations and the clinical context to evaluate the different interventions, to define clinical endpoints, and how to put into practice the lessons that can be learned from the development of antiviral therapy for HIV or from the vaccines against pneumococcus. For defining and adapting clinical endpoints, Ramilo outlined challenges as well as opportunities and emphasized the importance of obtaining robust and comprehensive clinical data, because society wants the interventions to be cost–effective. Finally, Ramilo showed the importance of selecting laboratory markers, virologic as well as immune markers and encouraged the participants to incorporate collection of clinical samples in their clinical trials to facilitate laboratory analyses that permit a better understanding of how interventions work (or not) against RSV infection.

REGULATORY CONSIDERATIONS IN INITIATING PAEDIATRIC CLINICAL TRIALS FOR RSV THERAPEUTICS

Overview of pediatric development plans evaluated by Paediatric Committee PDCO: regulatory considerations for initiating pediatric trials

Dr Eichler gave an overview of Paediatric Investigation Plans (PIPs) for RSV–antivirals and monoclonal antibodies for which the information is available in the public domain. Main challenges for the PDCO of the European Medicines Agency (EMA), responsible for assessing the content of Paediatric investigation plans, are the current lack of generally agreed recommendations by the scientific community regarding when to best initiate antiviral treatment, and the lack of validated and agreed clinically meaningful outcome measures for evaluating the effect of RSV antivirals.

Upper respiratory tract infections caused by RSV are not considered a major clinical problem, necessitating antiviral treatment. However, in children at high risk for severe lower respiratory disease (LRD) caused by RSV, the initiation of antiviral treatment early in course of RSV infection, ie, before major tissue injury in the lower airways has occurred, could benefit them to prevent severe LRD. At present, clear definitions for bronchiolitis and/or severe LRD are lacking.

The highest disease burden is considered in the first 2 years of life; consequently, this age cohort should be included in clinical trials [3]. It would be desirable to also evaluate the efficacy and safety of antivirals in older children at high risk to develop serious RSV LRD, ie, children with chronic underlying conditions, such as immunodeficiency or neuromuscular disease. As these populations are very heterogeneous and small, the conduct of dedicated clinical studies in these populations is most likely not feasible. At present, it is unknown to what extent extrapolation of efficacy is possible from young children in whom severe RSV infection manifests clinically as bronchiolitis to older children with underlying chronic diseases, in whom RSV infections manifest as RSV pneumonia. Therefore, the generation of limited clinical data in older children with chronic underlying conditions needs to be discussed.

In the absence of validated endpoints, a consensus definition of a set of core outcome measures which should be measured and reported in all clinical trials is highly warranted to allow comparability of trial results and to validate candidate endpoints.

Regulatory aspects related to development of vaccines for RSV

Dr Pelfrene outlined the main regulatory considerations for maternal immunization and subsequent determination of protective efficacy in the offspring, the primary vaccination in infants as well as the desired safety database and duration of safety follow–up for both strategies. Foremost was stressed that the trial sponsor should define the intended aim of vaccination, since ultimately this will inform the labeling claim, ie, prevention of a specific clinical presentation. The case definition for RSV disease will require subjects to meet both clinical and laboratory criteria [4]. In this regard, it was emphasized that specific and sensitive assay methods for detection of RSV breakthrough cases should be employed in a standardized manner across participating study centers. The need and feasibility of a central laboratory confirmation should also be determined. Study protocols will need to define and justify the method of specimen collection (eg, nasopharyngeal aspirate or nasal swab) and provide details on sample storage and shipping conditions.

With respect to vaccination during pregnancy, it was stated that background rates of fetal demise, prematurity and congenital aberrations need to be available. Prior to conducting clinical trials in pregnant women, demonstration of safety and immunogenicity data in healthy adults (including non–pregnant women) will be necessary, as well as favorable preclinical data to be obtained on immunology and toxicology, including experiments performed in late stage pregnant animals. Exploratory trials may provide sufficient data on trans–placental transfer and persistence of maternal antibodies in the infant. In this sense, it is recognized that duration of protection may be trial setting dependent and a function of antibody titer at birth and rate of decline in the infant. Efficacy trials are expected to be broadly inclusive but there will likely be a need to stratify or to exclude sub–groups with recognized poor trans–placental transfer, such as HIV infected subjects. For the confirmatory trials, due consideration should be given to seasonality: immunization in 3rd trimester pregnancy will need to be scheduled within an appropriate time–window so that ensuing delivery coincides with the early part of the RSV season. With regard to infant immunization, it was stressed that the aim would be to elicit a strong neutralizing antibody response with a non–T helper type 2 (Th2) biased cellular immune response. Vaccine development strategy will most probably be featuring an age–de–escalation approach, and thus first be administered to RSV seropositive adults and children before progressing to RSV–naïve infants. The question of optimal timing of immunization was raised, with primary series to start as early as feasible in infancy, taking into account the inhibitory effect of maternal antibodies. In the case of vaccinating infants born to vaccinated mothers, the estimate of duration of passive protection should be known before deciding when to start active immunization. Further on, it was conveyed that the current EU general expectation regarding a pre–licensure safety database for a novel vaccine is a minimum of 3000 exposed persons to the final dose regimen of the vaccine. Though, for the vaccination of infants, it needs to be discussed what the breadth of evidence should be, to support negligible risk of disease enhancement in the RSV–naïve population.

During the discussion, it was acknowledged that the major RSV burden and mortality occurs in low–income countries. As such, it was asserted that ideally, vaccine– and therapeutic development programs need to consider a global perspective. Hence, it would be desirable that case definitions include clinical features which are easily standardized and generalizable across the different settings. EMA urged the use of the same scoring systems or scales in the trials and meeting participants were encouraged to consider scientific qualification advice for potential candidate biomarkers/outcome measures.

MEASURING SAFETY

Assessing the impact of anti–RSV interventions: clinical endpoints and biomarkers

It is possible to combine clinical endpoints to evaluate anti–RSV interventions with viral factors, host immune profiles and antibody responses for the diagnosis, pathogenesis and assessment of RSV disease severity [5]. *Dr Mejias* discussed that as age goes up, a decrease of antibodies against RSV are seen in the infant, acutely infected with RSV, reflecting maternal antibody transfer. Having measured neutralizing activity, they didn't find a perfect correlation between concentration and neutralization, which needs to be understood. In fact, standardizing antibody assays and identifying a consistent antibody threshold indicative of protection still needs to be defined. However, Mejias made clear that other biomarkers such as genomics markers have a great value as predictive tools and to objectively assess disease severity. The team formed by Mejias and Ramilo found in infants hospitalized with RSV bronchiolitis significant correlations between a molecular genomic score and (1) the clinical disease severity score, (2) duration of hospitalization and (3) duration of supplemental oxygen. They are

now using these tools to understand responses to therapeutic interventions, to monitor disease progression as well as to study the normal maturation of the immune system in infants. Furthermore, in collaboration with Dr Bogaert, Mejias studied the role of bacterial colonization on RSV disease severity, and noticed that nasopharyngeal bacterial colonization with specific pathogens did not appear to be a passive phenomenon. Specifically, infants with RSV infection and colonized with *S. pneumoniae* or *non–typable H. influenzae* displayed a more severe clinical phenotype and different host transcriptional profiles. Understanding the RSV–bacterial interactions in these children is of key importance when evaluating the benefit of RSV therapeutics. The challenge consists of developing composite endpoints that include clinical, virologic and laboratory parameters to monitor responses to clinical interventions.

Enhanced RSV disease and vaccines

There are a number of new strategies for RSV vaccines, mentioning that each formulation may present individual characteristics that theoretically decrease or increase the risk for enhanced RSV disease (ERD) [6]. *Dr Polack* discussed ERD, which has been seen only in sero–negative infants and young children who were previously immunized with a formalin inactivated RSV vaccine. For many years, the consensus was that nothing but live attenuated RSV vaccines would ever be used to immunize infants. Therefore, the characterization of ERD phenotypes was of academic interest but had limited regulatory implications. Based on the outcome in mice studies and limited human data, it may be assumed that it is worthwhile to look for Th2 endpoints. However, there is a need for a consensus definition of Th2 bias and/or a consensus set of control groups for studies. ERD does not impact sero–positive children, because better antibodies precede immunization. Polack told the audience not to rely on older subjects in phase 1 studies, since there is no ERD in patients that had RSV infection before, so that would prove to be futile.

Polack concluded that the monoclonal antibodies and maternal immunization strategies are safe. For other vaccines, awareness of steric hindrance is needed and he suggested waiting between immunization and challenge, because of the danger of misunderstanding the read out. In summary, Polack said that RSV vaccines should elicit a long–lived protective antibody of high avidity for RSV protective antigens and specific cytotoxic T lymphocytes, and, neither elicit lung eosinophils, nor bias the response to Th2.

During the discussion, the participants discussed how to measure low affinity antibodies in the clinical setting, and how to assess the risk for developing ERD. How to use the information for vaccine studies was food for thought, since a perfect model of ERD is lacking. Dr Polack stressed that if ERD occurs, it probably will manifest in many individuals, as early trials had disease rates above 50%.

RSV COST EFFECTIVENESS AND EPIDEMIOLOGY

RSV vaccine in development: assessing the potential cost–effectiveness in high risk adult populations

The potential cost–effectiveness of RSV vaccination in high risk adults was discussed, with a main focus on the elderly population. *Dr Pouwels* noted the importance of thinking about the target population, when evaluating vaccines. The most obvious choice would be infants, given the well–established burden among infants. However, there is accumulating evidence that RSV causes a substantial burden in high–risk adults and elderly [7]. To date, two health economic studies that evaluated vaccination of the elderly against RSV have been published [8,9]. Both studies indicated that vaccination of the elderly has the potential to be cost–effective, especially among high–risk elderly. However, both studies were performed with limited data about the burden of RSV among the elderly. Moreover, one of the main drivers of the successes of several vaccination campaigns–indirect protection by reducing transmission–was not taken into account. Hence, there is a need for an updated transmission dynamic cost–effectiveness model to evaluate which vaccination strategy is most cost–effective. A recent transmission dynamic model from Kenya, which did not focus on the elderly population, concluded that vulnerable infants could be indirectly protected by annual vaccination of all school–age children [10]. Another study from the same region, estimated that it may be sufficient to vaccinate children aged 5–10 months [11]. It is clear that more models, also incorporating direct and indirect protection of the elderly, are needed. Pouwels concluded that to assess the impact and cost–effectiveness of the different strategies using transmission dynamic models, better age–group specific virological surveillance and more data on the age– and risk–group specific burden are needed.

Collaboration to conduct research on vaccine preventable diseases in Canada–what we are learning

The Canadian Immunization Research Network (CIRN; http://cirnetwork.ca/) includes a hospital–based surveillance network (Serious Outcomes Network) which evaluates morbidity associated with vaccine–preventable infections (eg, influenza) and vaccine efficacy in adults. CIRN collaborates with the Immunization Monitoring Program

ACTive (IMPACT) which conducts similar research as well monitoring adverse events following immunization, in pediatric health centers [12]. Dr Langley noted that these networks could be of interest to the RSV field for determining the burden of disease. The supportive networks within CIRN, Social Sciences and Humanities Network, Reference Laboratory Network, and Modeling and Economic Research Network actively take part in study design the studies [13]. The focus of CIRN's work is vaccine safety, immunogenicity and effectiveness, vaccine coverage, vaccine hesitancy, and program implementation and evaluation. IMPACT started a working group on RSV in 2015, looking at what is available on Canadian epidemiology and planning for a surveillance project on severe outcomes in hospitalized children. These networks have found that transparent, open processes, standard operating procedures for study processes, project review and funding of have been essential in conducting research in multiple provinces across a large country.

The discussion was based on the influenza surveillance data generated by these networks, including the studies controlling for frailty in older persons. The need for significantly more data to make the right cost–effective models impressed the audience, but the overall expectation was that in a year, based on CIRN and other studies of RSV illness in the community, there will be more sophisticated models available to accurately assess this burden.

During the discussion it was emphasized what we can learn from a recent transmission dynamic cost–effectiveness model that led to the decision to extend influenza vaccination to children in the UK. A lot more data are needed to build a similar model for RSV vaccination, both in terms of better age–specific RSV virological surveillance, short– and long–term consequences of RSV infection among different age–groups, and the effect of vaccination on transmissibility of RSV. Meanwhile, cost–effectiveness models should be updated when clinical trial data and improved burden of disease estimates become available in the near future.

DEVELOPMENT OF RSV VACCINES FOR USE IN PREGNANCY

Clinical endpoints in trials of RSV vaccines in pregnant women: study design issues, assessment of safety and effectiveness

Maternal influenza studies were discussed as a comparator for vaccination of pregnant women against RSV. *Dr Nunes* discussed the challenges for maternal immunization especially in low– and middle–income countries (LMIC), and identified assessing the accurate gestational age as a main problem. The best method would be early ultra sound, which is not always available. Nunes emphasized the importance to know the study population for designing such a trial and gave examples of risk–factors which are more prevalent in LMIC such as co–morbidities like anemia and concomitant illnesses (HIV, malaria) that may affect the placental function. For experimental vaccines a randomized, placebo–controlled trial would be the desired study design, with a primary objective of evaluating the efficacy of RSV vaccination of pregnant women against laboratory–confirmed RSV LRTI in their infants up to 3 months of age. A major concern in maternal immunization trials is safety endpoints; in this regard the WHO requested the Brighton Collaboration (BC) to develop a guidance document harmonizing safety assessments during maternal and neonatal vaccine trials in all resource settings ie, in LIC and high IC. Although promising, maternal vaccination might be limited by transplacental antibody transfer, antibody decay rates in the infants and safety in pregnant women. Information on RSV–associated disease burden in pregnant women is lacking and will be obtained from virological analyses from the recent large maternal influenza vaccine trials.

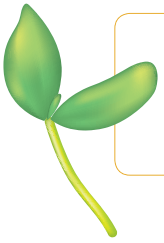
Lessons learned from non–RSV maternal immunization – safety, immunogenicity and effectiveness

Safety of maternal vaccination was discussed by *Dr Van der Maas*. She mentioned that based on the monitoring in a Norwegian influenza immunization study after the influenza A (H1N1) pandemic in 2009, no increased risk was found for fetal death in vaccinated women compared to non–vaccinated women, but women who contracted influenza during pregnancy did have a significantly greater risk of fetal death compared with pregnant women who did not suffer from influenza [14]. It was concluded that vaccination works and is safe, backed by other papers and a Dutch study. Follow up of the infants up to 1 year showed no difference for growth, development and GP infection–related contact rates, in infants of vaccinated and non–vaccinated mothers [15–17]. Furthermore, Van der Maas referred to the re–emergence of pertussis in the world, and emphasized the importance of monitoring the maternal pertussis vaccination for effectiveness, immunogenicity and safety. Regarding immunogenicity and the transplacental IgG transport, the timing of the vaccination is crucial. One of the lessons learned from non–RSV maternal vaccination is the essential monitoring of safety, effectiveness and immunogenicity, in order to maintain the public trust, the occurrence of disease in infants ("vaccine failure" vs "failure to vaccinate"), and to optimize the infant and maternal vaccination schedule.

It was discussed that the most common cause of non–obstetrical fatal illness in pregnant women is in fact respiratory disease. Another point was what would be the best settings for conducting trials of such a vaccine, since the burden is highest in populations that cannot afford private health care and high cost interventions. The participants agreed that the trials should be conducted both in developed countries and in LMIC. Besides, the BMGF has partnered with industry with the aim of making these vaccines available at a lower cost in developing countries.

CONCLUDING REMARKS

The 2016 ReSViNET meeting was organized with participation of pharmaceutical companies, public health advocates, academia, WHO, FDA, EMA and the BMGF. There was a focus on regulatory requirements for upcoming RSV therapeutics. The meeting integrated information from many of the stakeholders, including views of the regulators and public health. Alignment of the regulatory requirements with the developments in the pharmaceutical field was identified as a major challenge. Integrating the views of all stakeholders, including the patient's perspective, will optimize the development of novel therapeutics against a respiratory virus which continues to cause so much disease to so many people worldwide.



Disclaimer: The views expressed in this paper are the personal views of the authors and must not be understood or quoted as being made on behalf of or representing the position of the EMA or one of its committees or working parties.

REFERENCES

1 Nair H, Nokes DJ, Gessner BD, Dherani M, Madhi SA, Singleton RJ, et al. Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta–analysis. Lancet. 2010;375:1545-55. Medline:20399493 doi:10.1016/S0140-6736(10)60206-1

2 Mazur NI, Martinon–Torres F, Baraldi E, Fauroux B, Greenough A, Heikkinen T, et al. Lower respiratory tract infection caused by Respiratory Syncytial Virus: current management and new therapeutics. Lancet Respir Med. 2015;3:888-900. Medline:26411809 doi:10.1016/S2213-2600(15)00255-6

3 Smyth RL, Openshaw PJ. Bronchiolitis. Lancet. 2006;368:312-22. Medline:16860701 doi:10.1016/S0140-6736(06)69077-6

4 Modjarrad K, Giersing B, Kaslow DC, Smith PG, Moorthy VS. WHO consultation on Respiratory Syncytial Virus Vaccine Development Report from a World Health Organization Meeting held on 23–24 March 2015. Vaccine. 2016;34:190-7. Medline:26100926 doi:10.1016/j.vaccine.2015.05.093

5 Mejias A, Dimo B, Suarez NM, Garcia C, Suarez–Arrabal MC, Jartti T, et al. Whole blood gene expression profiles to assess pathogenesis and disease severity in infants with respiratory syncytial virus infection. PLoS Med. 2013;10:e1001549. Medline:24265599 doi:10.1371/journal.pmed.1001549

6 Acosta PL, Caballero MT, Polack FP. Brief history and characterization of enhanced Respiratory Syncytial Virus disease. Clin Vaccine Immunol. 2015;23:189-95. Medline:26677198 doi:10.1128/CVI.00609-15

7 Falsey AR, Hennessey PA, Formica MA, Cox C, Walsh EE. Respiratory Syncytial Virus infection in elderly and high–risk adults. N Engl J Med. 2005;352:1749-59. Medline:15858184 doi:10.1056/NEJMoa043951

8 Meijboom MJ, Pouwels KB, Luytjes W, Postma MJ, Hak E. RSV vaccine in development: assessing the potential cost–effectiveness in the Dutch elderly population. Vaccine. 2013;31:6254-60. Medline:24148573 doi:10.1016/j.vaccine.2013.10.023

9 Gessner BD. The cost–effectiveness of a hypothetical Respiratory Syncytial Virus vaccine in the elderly. Vaccine. 2000;18:1485-94. Medline:10618546 doi:10.1016/S0264-410X(99)00425-9

10 Poletti P, Merler S, Ajelli M, Manfredi P, Munywoki PK, Nokes D, et al. Evaluating vaccination strategies for reducing infant Respiratory Syncytial Virus infection in low–income settings. BMC Med. 2015;13:49. Medline:25857701 doi:10.1186/s12916-015-0283-x

11 Kinyanjui TM, House TA, Kiti MC, Cane PA, Nokes DJ, Medley GF. Vaccine induced herd immunity for control of Respiratory Syncytial Virus disease in a low–income country setting. PLoS ONE. 2015;10:e0138018. Medline:26390032 doi:10.1371/journal.pone.0138018

REFERENCES

12 Scheifele DW, Halperin SA. Immunization Monitoring Program, Active: a model of active surveillance of vaccine safety. Semin Pediatr Infect Dis. 2003;14:213-9. Medline:12913834 doi:10.1016/S1045-1870(03)00036-0

13 Edwin JJ, Reyes Domingo L, Booth TF, Mersereau T, Skowronski DM, Chambers C, et al. Surveillance summary of hospitalized pediatric enterovirus D68 cases in Canada September 2014. Canada Communicable Disease Report 2015. Available: http://www.phac-aspc.gc.ca/publicat/ccdr-rmtc/15vol41/dr-rm41s-1/surveillance-1-eng.php. Accessed: 3 March 2016.

14 van Lier A, Steens A, Ferreira JA, van der Maas NA, de Melker HE. Acceptance of vaccination during pregnancy: experience with 2009 influenza A (H1N1) in the Netherlands. Vaccine. 2012;30:2892-9. Medline:22374374 doi:10.1016/j.vaccine.2012.02.030

15 van der Maas N, Dijs–Elsinga J, Kemmeren J. van LA, Knol M, de Melker H. Safety of vaccination against influenza A (H1N1) during pregnancy in the Netherlands: results on pregnancy outcomes and infant's health: cross–sectional linkage study. BJOG. 2016;123:709-17. Medline:25754741 doi:10.1111/1471-0528.13329

16 Pasternak B, Svanström H, Mølgaard-Nielsen D, Krause TG, Emborg HD, Melbye M. Risk of adverse fetal outcomes following administration of a pandemic influenza A(H1N1) vaccine during pregnancy. JAMA. 2012;308:165-74. Medline:22782418 doi:10.1001/jama.2012.6131

17 Pasternak B, Svanström H, Mølgaard-Nielsen D, Krause TG, Emborg HD, Melbye M, et al. Vaccination against pandemic A/H1N1 2009 influenza in pregnancy and risk of fetal death: cohort study in Denmark. BMJ. 2012;344:e2794. Medline:22551713 doi:10.1136/bmj.e2794

Correspondence to:
Prof Louis Bont
Wilhelmina Children's Hospital
Pediatric Infectious Disease and Immunology
University Medical Center
Utrecht
3584 EA
The Netherlands
l.bont@umcutrecht.nl

Setting health research priorities using the CHNRI method: I. Involving funders

Igor Rudan[1], Sachiyo Yoshida[2], Kit Yee Chan[1], Simon Cousens[3], Devi Sridhar[1], Rajiv Bahl[2], Jose Martines[2]

[1] Centre for Global Health Research, The Usher Institute for Population Health Sciences and Informatics, The University of Edinburgh, Scotland, UK

[2] Department of Maternal, Newborn, Child and Adolescent Health, World Health Organization, Geneva, Switzerland

[3] London School of Hygiene and Tropical Medicine, London, UK

In 2007 and 2008, the World Health Organization's Department for Child and Adolescent Health and Development (later renamed as WHO MNCAH – Maternal, Newborn, Child and Adolescent Health) commissioned five large exercises to define research priorities related to the five major causes of child deaths for the period up to the year 2015. The exercises were based on the CHNRI (Child Health and Nutrition Research Initiative) method, which was just being introduced at the time [1,2]. The selected causes were childhood pneumonia, diarrhoea, birth asphyxia, neonatal infections and preterm birth/low birth weight [3–7]. The context for those exercises was clearly defined: to identify research that could help reduce mortality in children under 5 years of age in low and middle income countries by the year 2015. The criteria used in all five exercises were the "standard" CHNRI criteria: (i) answerability of the research question; (ii) likelihood of the effectiveness of the resulting intervention; (iii) deliverability (with affordability and sustainability); (iv) potential to reduce disease burden; and (v) effect on equity [3–7].

The five criteria used by the scorers were intuitive as they followed the path from generating new knowledge to having an impact on the cause of death. They were chosen with a view to identifying research questions that were most likely to contribute to finding effective solutions to the problems. However, after the five exercises – all of which were published in respected international journals [3–7] – the WHO officers were left with an additional question: how "fundable" were the identified priorities, ie, how attractive were they to research funders? More specifically, should another criterion be added to the CHNRI exercises, which would evaluate the likelihood of obtaining funding support for specific research questions?

To answer these questions, coordinators of the CHNRI exercises at the WHO agreed that it would be useful to invite a number of representatives from large funding organizations interested in child health research to take part in a consultation process at the WHO. The process aimed to explore funders' perspective in prioritization of health research. The funders would be presented with the leading research priorities identified through the CHNRI exercises and asked to discuss any potential variation in their likelihood of being funding. If all the leading priorities were equally attractive to funders and likely to attract funding support, this would indicate that the "standard" CHNRI criteria were sufficient for the process of prioritization. However, if there were large differences in attractiveness of the identified research priorities to funders, then adding another criterion to the exercise – "likelihood of obtaining funding support", or simply "fundability" – would be a useful addition to the standard CHNRI framework.

In 2007 and 2008, the World Health Organization's Department for Child and Adolescent Health and Development commissioned five large research priority setting exercises using the CHNRI (Child Health and Nutrition Research Initiative) method. The aim was to define research priorities related to the five major causes of child deaths for the period up to the year 2015. The selected causes were childhood pneumonia, diarrhoea, birth asphyxia, neonatal infections and preterm birth/low birth weight. The criteria used for prioritization in all five exercises were the "standard" CHNRI criteria: answerability, effectiveness, deliverability, potential for mortality burden reduction and the effect on equity. Having completed the exercises, the WHO officers were left with another question: how "fundable" were the identified priorities, i.e. how attractive were they to research funders?

THE MEETING WITH THE FUNDERS (GENEVA, 27–29 MARCH 2009)

In March 2009, MNCAH invited 40 representatives from funding organizations, including the Bill and Melinda Gates Foundation, the Wellcome Trust, National Institutes of Health USA, Department for International Development UK, Save the Children, INCLEN, EPICENTRE, UNICEF, USAID, PATH, Ministry of Science and Technology of India, Ministries of Health of Zambia, Pakistan and Brazil, Global Forum for Health Research, Trinity Global Support Foundation, Children's Investment Fund Foundation, Osaka Research Institute for Maternal and Child Health. Eventually, 16 representatives of funding agencies agreed to take part in the exercise under the condition of anonymity. Moreover, it was understood that their input would not necessarily be the official position of their respective funding agencies, nor would it create any form of funding obligation.

Having explained the aims of the consultation meeting to the representatives of funding agencies, the 16 participants were presented with a list of the top 10 research priorities for each of the five major causes of child deaths: pneumonia, diarrhea, birth asphyxia, neonatal infections and preterm birth/low birth weight [3–7]. This set of 50 research priorities represented roughly the top 5% of all the research ideas submitted for scoring during the CHNRI exercises.

The WHO coordinators (RB and JM) explained each of the 50 leading research priorities to the 16 donor representatives. Then, the 16 donor representatives were provided with the list of research priorities and asked to individually identify those that were most likely to receive funding support from their respective organizations.

Funding attractiveness was measured in two ways. First, funder representatives were asked to rank the identified research priorities according to their likelihood to receive funding support under an organization's current investment policies and practices. Second, funding attractiveness was measured by asking funder representatives to distribute a theoretical US$ 100 among the research priorities that seem most fundable. Results were used to facilitate discussion on what makes a research question attractive (or unattractive) for funding support. The scoring sheet that was given to meeting participants is shown in **Figure 1**. While they did not need to provide their name or organization, they were asked to assign ranks 1–10 to the ten research priorities identified for each of the five causes of death (column 1), and also to distribute a hypothetical US$ 100 to different research priorities in concordance to the likely funding support that they may obtain.

Sixteen participants scored the identified research priorities according to the instructions (**Figure 1**). The average ranks across the 16 participants (1 = most likely to be funded; 10 = least likely to be funded) assigned to the 50 research priorities ranged from 3.7 to 7.2. The average US$ amount assigned to research priorities ranged from US$ 20.1 to US$ 2.5. There was general consistency between ranks and the US$ assigned to research priorities.

CAUSE OF DEATH (1 / 5) - e.g. PNEUMONIA

NAME:
ORGANIZATION:

	RANK?	US$?
RESEARCH QUESTION 1		
RESEARCH QUESTION 2		
RESEARCH QUESTION 3		
RESEARCH QUESTION 4		
RESEARCH QUESTION 5		
RESEARCH QUESTION 6		
RESEARCH QUESTION 7		
RESEARCH QUESTION 8		
RESEARCH QUESTION 9		
RESEARCH QUESTION 10		

Figure 1. A questionnaire that was given to 16 funder representatives at the meeting to obtain information useful to understanding funding attractiveness of different research priorities.

Importantly, the analysis of the collective input based on the 2nd column (ie, assigned US$), presented in **Figure 2**, clearly shows that there was a rather substantial departure of the assigned funds from that expected at random: if all research priorities were equally likely to obtain support from the funders, then all the bars would be extending only to the line that represents an investment of US$ 10.0. Furthermore, 4 research priorities (8%) clearly stood out from the rest [8]. It was agreed that they might provide a starting point from which MNCAH Department could concentrate its efforts. These 4 research priorities are shown in **Table 1**.

Table 1. The 4 research priorities (8%) that were identified as positive outliers in terms of their likelihood to obtain funding support

Evaluate the quality of community workers to adequately assess, recognize danger signs, refer and treat acute respiratory infections (ARI) in different contexts and settings.
What are the barriers against appropriate use of oral rehydration therapy?
What are the feasibility, effectiveness and cost of different approaches to promote the following home care practices (breastfeeding, cord/skin, care seeking, handwashing)?
What are the feasibility, effectiveness and cost of a scheme of routine home visits for initiation of supportive practices, detection of illness and newborn survival?

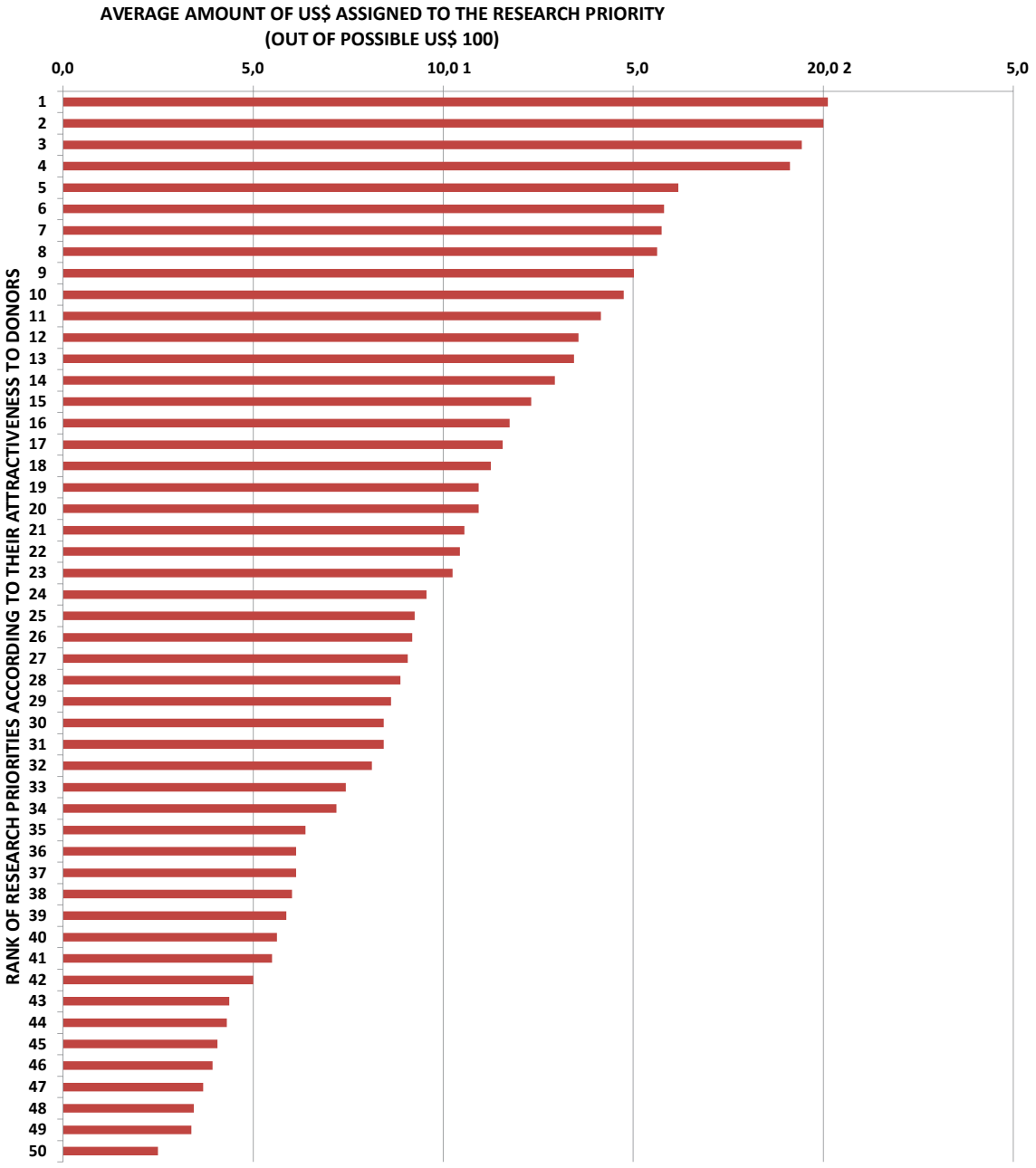


Figure 2. The results of the collective input from 16 funder representatives, showing large differences in funding attractiveness between 50 research priorities. No substantial differences in funding attractiveness would be indicated by equality of the scores on the horizontal axis at the US$ 10.0 line.

AN ANALYSIS OF THE EXERCISE WITH FUNDER REPRESENTATIVES

The results were analysed after the first day of the meeting and presented to donor representatives at the beginning of the second day of the meeting. An open discussion was held with participants to understand and interpret the results of their collective input. Participants agreed that the most important criteria for research prioritisation differed between researchers and funders. Researchers tended to value answerability, effectiveness, deliverability, impact on the burden and equity. Funders were also interested in the clarity and specificity of research ideas, value for money, novelty, international competitiveness of the groups proposing the research, linkages to broader societal issues, and complementarity with other long–term strategic investments that were already made. An important point in the discussion was that researchers and research funders, especially those in the private sector, often speak quite different languages. Researchers need to be clear on what their goals are and communicate these in more readily understood terms. This point is particularly important because it implies that the CHNRI exercises' research priorities that were identified as most likely to generate useful new knowledge may not be considered equally relevant by the funders.

In March 2009, WHO officers invited 40 representatives from organizations that provide substantial funding support for global child health research to take part in a consultation process at the WHO. The process aimed to explore funder's perspective in prioritization of health research. Eventually, 16 funders' representatives agreed to take part in the exercise under the condition of anonymity. Participants agreed that the most relevant criteria for prioritisation differed between researchers and funders. Funders are interested in clarity and specificity of research ideas, value for money, novelty, international competitiveness of the groups proposing the research, links to broader societal issues, and complementarity with other long–term strategic investments that they have already made. Some may be particularly interested in the potential for forming partnerships between researchers and industry to improve the translation of findings and their application

This should certainly be taken into account when presenting and discussing the results of the CHNRI exercises.

Moreover, there seem to be important differences between the categories of funders in the criteria that they use to decide on research priorities. Generally, all investors in health research are concerned with answerability of the proposed research ideas in an ethical way, feasibility and value for money. However, some may be particularly interested in potential for forming partnerships between researchers and industry to increase the translation of findings and their application. Ministries and international organizations appeared more interested in deliverability, affordability and sustainability of the resulting interventions, local and national research capacities to carry out the proposed research ideas, and whether a research question is linked to an ongoing public debate or an important societal issue. Industrial donors may be primarily motivated to generate patents and translate research results into commercial products. Finally, society as a whole may be more concerned with issues of safety and equity issues and ask whether implementation of research results would widen the existing socio–economic gaps..

Transparency of research priority setting processes must, therefore, begin with those who invest. Perceived returns on investments in health research should be clearly stated at the beginning of the process. They may be defined as reduction in disease burden wherever public money is being invested. Investors from industries may see patentable products as their preferred returns. Non–profit organizations may be primarily interested in increased media attention for their agenda. The context in which investment prioritization takes place is thus primarily defined by expected returns of the funders. Moreover, their investment styles may be balanced and responsible (suggested for those investing public funds), risk–averting (which may be preferred among some industrial partners) or risk–seeking and biased towards high risk – high profit avenues of health research (which may be typical for some industry and not–for–profit organizations).

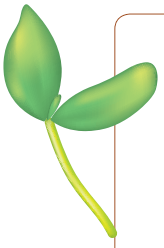
Apart from funders' perceived returns and their investment styles, the population, geographic area and disease burden of interest, the time frame in which returns are expected is an important defining component of the overall context. Priorities can differ substantially if the overall context is one of great urgency to tackle a problem, or whether decisions are made on very long–term, strategic investments.

CONCLUSIONS

The meeting with research funders organized by the WHO MNCAH department in March 2009 was exceptionally useful in understanding that funders certainly have their own views on what represents an attractive funding option.

Those views are not generalizable and may differ between categories of funders. Moreover, funders' perspectives are often quite different from those of researchers, or wider stakeholder groups. It is important to involve funders early in the process of setting research priorities, such as the CHNRI process, to encourage their ownership of the results. Funder–supported criteria must be taken into account, in addition to those preferred by the researchers and wider stakeholders. Otherwise, the outcomes of research prioritization exercises may have very limited impact on funders' decision making.

The key value of the CHNRI method to funders lies in its ability to transparently lay out the potential risks and benefits associated with investing in many competing research ideas, drawing on collective knowledge of the broad research community. Results of the CHNRI process represent an attempt on the part of researchers to communicate their views and opinions to funders in a way that is easily understood, transparent, replicable and intuitive. It provides useful additional information that funders may, or may not take into account when deciding on their own research agenda. From a methodological perspective, finding appropriate and effective ways of involving funders in future CHNRI exercises, communicating the outcomes clearly, and securing their commitment to acknowledge the results of the CHNRI process remain considerable challenges. An even greater challenge in future years will be to develop tools that can detect and evaluate the impact of CHNRI exercises on funder decision making and any change in funding priorities as a direct result of the CHNRI process. This should be particularly relevant to those who make decisions about investing public funds, whose primary agenda should be improving public health in the most cost–effective way – a target that CHNRI exercises should serve quite well.



Acknowledgements: We sincerely thank the representatives of the funding agencies for taking part in the consultation process and providing useful information under the condition of anonymity. SY, RB and JM are the employees of the World Health Organization. Their views expressed here do not necessarily represent the views of the Organization.

Funding: None.

Authorship declaration: JM and RB organized the consultation at the World Health Organization and IR moderated the consultation process. IR, SY and KYC wrote the article. SC and DS revised the article and made useful comments.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare no conflict of interest.

REFERENCES

1 Rudan I, Chopra M, Kapiriri L, Gibson J, Lansang MA, Carneiro I, et al. Setting priorities in global child health research investments: universal challenges and conceptual framework. Croat Med J. 2008;49:307-17. Medline:18581609 doi:10.3325/cmj.2008.3.307

2 Rudan I, Gibson JL, Ameratunga S, El Arifeen S, Bhutta ZA, Black M, et al. Setting priorities in global child health research investments: guidelines for implementation of the CHNRI Method. Croat Med J. 2008;49:720-33. Medline:19090596 doi:10.3325/cmj.2008.49.720

3 Rudan I, El Arifeen S, Bhutta ZA, Black RE, Brooks A, Chan KY, et al. Setting research priorities to reduce global mortality from childhood pneumonia by 2015. PLoS Med. 2011;8:e1001099. Medline:21980266 doi:10.1371/journal.pmed.1001099

4 Fontaine O, Kosek M, Bhatnagar S, Boschi–Pinto C, Chan KY, Duggan C, et al. Setting research priorities to reduce global mortality from childhood diarrhoea by 2015. PLoS Med. 2009;6:e41. Medline:19278292 doi:10.1371/journal.pmed.1000041

5 Bahl R, Martines J, Ali N, Bhan MK, Carlo W, Chan KY, et al. Research priorities to reduce global mortality from newborn infections by 2015. Pediatr Infect Dis J. 2009;28 Suppl:S43-8. Medline:19106763 doi:10.1097/INF.0b013e31819588d7

REFERENCES

6 Lawn JE, Bahl R, Bergstrom S, Bhutta ZA, Darmstadt GL, Ellis M, et al. Setting research priorities to reduce almost one million deaths from birth asphyxia by 2015. PLoS Med. 2011;8:e1000389. Medline:21305038 doi:10.1371/journal.pmed.1000389

7 Bahl R, Bhandari N, Bhutta ZA, Biloglav Z, Edmond K, Iyengar S, et al. Setting research priorities to reduce global mortality from low birth weight by 2015. J Glob Health. 2012;2:010403. Medline:23198132 doi:10.7189/jogh.01.010403

8 World Health Organization. Consultation Proceedings: Identifying priorities for child health research to achieve Millennium Development Goal 4. DRAFT April 30, 2009. Geneva: WHO, 2009, pp. 1–29.

Correspondence to:
Professor Igor Rudan
Centre for Global Health Research
The Usher Institute for Population Health Sciences and Informatics
The University of Edinburgh Medical School
Teviot Place
Edinburgh EH8 9AG
Scotland, UK
igor.rudan@ed.ac.uk

Setting health research priorities using the CHNRI method: II. Involving researchers

Sachiyo Yoshida[1], Simon Cousens[2], Kerri Wazny[3], Kit Yee Chan[3,4]

[1] Department for Maternal, Newborn, Child and Adolescent Health, World Health Organization, Geneva, Switzerland
[2] Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK
[3] Centre for Global Health Research, the Usher Institute for Population Health Sciences and Informatics, the University of Edinburgh, Edinburgh, Scotland, UK
[4] Nossal Institute for Global Health, University of Melbourne, Melbourne, Australia

Large groups of researchers who agree to offer their research ideas and then score them against pre–defined criteria are at the heart of each CHNRI priority–setting exercise. Although the roles of funders and other stakeholders are also very important, much of the exercise is focused on selecting and engaging a large group of researchers, obtaining their input and analysing it to derive the initial results of the process. In a sense, a CHNRI exercise serves to "visualise" the collective knowledge and opinions of many leading researchers on the status of their own research field. Through a simple "crowdsourcing" process conducted within the relevant research community, the CHNRI approach is able to collate a wide spectrum of research ideas and options, and come to a judgement on their strengths and weaknesses, based on the collective knowledge and opinions of many members of the research community. In doing so, it provides valuable information to funders, stakeholders and researchers themselves, which is obtained at low cost and with little time necessary to conduct the exercise.

Success in involving researchers within each research community, and ensuring their voluntary participation and engagement, is therefore essential to the successful completion of a CHNRI exercise. Over the past few years, we have been involved in assembling groups of researchers to participate in several CHNRI research priority–setting exercises. In this paper, we share our experience of what works well and what works less well and try to answer the most frequently asked questions when it comes to engaging researchers in the CHNRI exercises.

Figure 1 shows where within the CHNRI process researchers should be involved –which is after the funders have provided their input, and before other stakeholders are approached and asked to contribute.

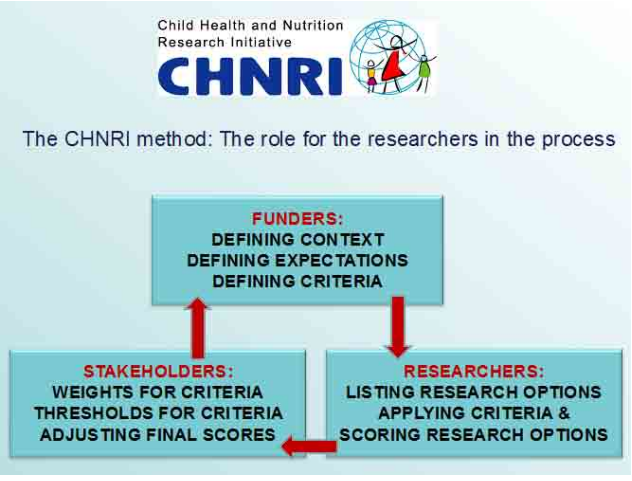


Figure 1. The role of researchers shown within the broader CHNRI process.

WHY DO RESEARCHERS NEED TO BE INVOLVED IN THE CHNRI EXERCISE?

Following input from funders, as described in a previous paper of this series [1], the managers of the CHNRI process then need to involve a sufficiently large sample of researchers. We discuss the considerations relevant to the optimal size of this sample of researchers in another paper of this series [2]. Researchers have two important roles in the CHNRI process: (i) providing the managers with a broad spectrum of research ideas, which usually span the spectrum of "description", "delivery", "development" and "discovery" research; and (ii) providing their own judgement on the likelihood that each submitted research idea will meet a set of pre–defined criteria. These judgements allow the ranking of a large number of submitted research ideas.

At this point, we should explain why CHNRI uses only researchers to provide research ideas, and not other groups of people–eg, funders, programme leaders and managers, other stakeholders, or simply members of the public. This is typically justified on the grounds that researchers are expected to possess far more knowledge and understanding of the state of their research field and the questions that have real potential to generate new knowledge. Importantly, their judgement of each research idea against the priority–setting criteria will also be based on an understanding of the realities of the research process and the success rate in their field. Including participants without this prior knowledge would likely introduce "random noise" into the exercise, resulting in most or all of the ideas receiving similar scores. Thus, restricting participation in these steps to researchers is expected to improve discrimination between the competing research ideas by using the collective knowledge and opinion of a small group of very knowledgeable people.

There is also a practical reason for this: by selecting the most productive, or highly cited researchers over the several preceding years, we are targeting the very group of people who will be most competitive for the research grant calls and likely be awarded the majority of the grants in the immediate future. We should also stress that this is, potentially, a "double edged sword", because researchers may not be entirely objective in their scoring and may tend to score highly their own preferred areas. This is why the chosen group always needs to be large enough, to prevent anyone's individual input having a substantial effect on the overall scores. Therefore, the leading researchers are given power through this method to influence the priorities and shape the topics for the future grants, ie, influence the subjects of the calls that are advertised by the funders, rather than simply responding to them. This could also be helpful to the funders, who do not have an easy access to a collective opinion of their research field.

It is worth bearing in mind that an important characteristic of the CHNRI method is its flexibility. Suggestions provided in the guidelines are not prescriptive, and each exercise can be tailored to meet the specific needs of the exercise. For example, some exercises may be mainly focused on implementation ("delivery") or fundamental ("discovery") research, particularly if the exercise is related to a specific intervention or geographic context. There have been several examples of such exercises, eg, the implementation of zinc interventions [3], implementation research for maternal and newborn health [4], emerging (discovery–based) interventions for childhood pneumonia and diarrhoea [5,6] and others. In such cases, there is scope for involving further groups of people whose knowledge and experience can provide informative input, particularly if this input is limited to the priority–setting criteria where the researchers would be unlikely to possess any first–hand knowledge. For example, many programme managers contributed to the scoring of questions on the newborn research agenda in relation to its deliverability, affordability and sustainability [7]. Our analyses of previous exercises have shown that the researchers tend to be less optimistic than programme managers on the criterion of answerability, while they tend to be more optimistic on the criterion of deliverability, affordability, sustainability and maximum potential for burden of disease reduction; similarly, programme managers tend to prioritise implementation research questions, whereas researchers prioritised technology–driven research [2,8]. Clearly, a good understanding of the complexities and challenges involved tends to make the experts–whoever they are–more cautious about the prospects of the suggested research ideas.

HOW TO INVOLVE RESEARCHERS IN THE CHNRI EXERCISE?

In planning the involvement of the group of researchers, the minimum target sample size needs to be decided early in the process. The optimal number will be derived based on the analyses conducted by Yoshida et al. [2], as mentioned previously. Yoshida's analyses suggest that the ranking of proposed research ideas, relative to each other, stabilises at surprisingly small sample sizes–ie, once that 30–50 people with private knowledge on the topic are involved, it is unlikely that the ranking of proposed research ideas will change markedly with the addition of further researchers and their opinions. Given this finding, targeting sample sizes of 50 or greater should result in a replicable CHNRI priority–setting exercise [2].

However, in planning the number of scorers needed, an important issue needs to be considered, which can reduce not only the actually achieved sample size quite substantially, but also introduce potential bias that can invalidate

the entire exercise. This is the issue of *(self–)selection bias*. The nature of CHNRI process means that researchers are usually invited (using e–mail or other means) by the management team to take part in the exercise. Their participation is needed in two consecutive steps of the process: (i) providing research ideas that they think would stand a good chance against all other ideas, given the pre–defined priority–setting criteria; and (ii) scoring a long list of research ideas against the pre–defined criteria. While the first step, providing research ideas, is not very time–consuming for researchers, the second step is a lot more time consuming and it may require several hours of input.

In an analysis of the first 50 CHNRI exercises, in which more than 5000 scorers were approached, Rudan et al. reported that the initial response rate (ie, submitting research ideas) was about 60%, with each expert submitting an average of about 3 research ideas. However, when all the initially invited experts were approached again to score the "consolidated" list of research ideas, the response rate dropped to only about 35%. Thus 40% of potential scorers are lost at the first stage, and further 25% of the total number are lost at the second stage (Rudan I, personal communication). The reason for re–contacting everyone who was initially invited to participate, even if they didn't offer any research ideas, is that there may be experts who are not keen giving away their ideas, but would be prepared to score ideas generated by others. This may help to preserve the initial sample that was contacted to the maximum extent possible.

Non–response has two important implications for an exercise. First, it reduces the actual sample size. This can be accounted for–eg, if the desired sample is 100 scorers, then about 300 probably need to be invited to participate in the exercise. Second, and more worrying, is the potential for bias in the results if responders and non–responders differ in their opinions. Results based on inputs from only about one third of the initial pool of researchers contacted may suffer from self–selection bias. For example, if individuals are more likely to respond to an invitation from the management group if they know the members of that group well, they may also be more likely to share similar views with the management group members. Others, who may disagree with those views and may, in fact, be in a majority in that particular research community, would not have their opinions recorded, or would be underrepresented. The high proportion of non–responders in many CHNRI exercise is therefore an important issue and we plan to conduct further work to explore non–response in previous exercises by comparing the characteristics of responders vs non–responders. The important thing to realise in relation to this self–selection bias is that it cannot be attenuated or controlled by further increasing sample size with new invitees because, no matter how large the sample size, they may still be based on the opinions of an unrepresentative subset of research community. In summary, increasing the achieved sample size can be done by inviting more people to participate, or by improving the response rate. The former approach will not attenuate possible self–selection bias, while the latter would tend to reduce the scope for bias and should be preferred. Several reminders are, therefore, usually sent to all invited participants to maximise the response rate.

SELECTING AND APPROACHING THE RESEARCHERS

The approach to identifying whom to invite to participate in the exercise can be very flexible, but must be credible to both the reviewers of the resulting publication, and also to any researchers who are left out of the exercise (ie, don't get an invitation). We present three examples of previous CHNRI exercises to examine how different strategies may work in different specific situations.

EXAMPLE OF THE CHNRI EXERCISE ON RESEARCH PRIORITIES FOR CHILDHOOD PNEUMONIA MORTALITY REDUCTION

This exercise [9], published in 2011, involved a small community of researchers working on childhood pneumonia in the low– and middle–income (LMIC) settings. A search for publications on childhood pneumonia in low–resource settings over the previous 5 years listed by the Web of Science identified only a few hundred publications in total. Ranking the authors of these publications ranked by the number of those papers that they had co–authored, revealed that the 100 most productive names were associated with a large majority of papers, and that those authors who were not among the most productive 100 had each contributed 3 papers or fewer over the previous 5 years. The decision was therefore taken to invite the most productive 200 researchers on the basis that this would cover almost the entire research community on this topic, regardless of the nature or importance of their discoveries.

It was agreed that an official approach through the World Health Organization (WHO), that agreed to serve as the hosting hub for the management group, would be most likely to persuade invited researchers to participate in the exercise. Moreover, mentioning that they were selected based on their placement among the 200 most productive researchers in this field would help to make them feel appreciated and that their work is valued. Nevertheless, even with these measures taken, the final response rate in terms of scoring in this small research community was 45/200 (22.5%).

Initially, the researchers were contacted through individual e–mails sent from the WHO, which explained the aim

of the exercise, acknowledged the contribution of each researcher to the field, and explained the type of the research idea that was sought – ie, neither too broad, nor too specific (this was further explained in the guidelines for implementation of the CHNRI method) [10]. They were also asked to consider different instruments of health research, ie, "description", "delivery", "development" and "discovery" and they were given an example of a "valid" research idea from each of those four types of research. They were initially given up to one month to submit as many research ideas as they wished, and two further reminders were sent at two weekly intervals following the initial deadline before the total number of submitted ideas reached 500. At that point, reminders were stopped and the management group studied the potential bias introduced because some researchers submitted many more ideas than others. At that point, a "consolidation" of the list of research ideas was conducted to ensure that the retained questions were evenly distributed across different research instruments and main research avenues and cover them all reasonably well. In this phase, all duplicate ideas were removed, while similar ideas were compressed into a single research question. This resulted in the reduction of the number of research ideas considered for scoring from 500 to 158, thus also making the scoring process more manageable.

Depending on the number of research ideas and the anticipated time required for scoring, one option is to offer the scorers the option of only scoring the criteria that they feel most comfortable with scoring – another flexibility in the CHNRI method. It is important that each scorer scores all research ideas on the same criterion, rather than scoring some but not all ideas for all criteria. This ensures that each research idea is scored by the same set of scorers, avoiding any personal preferences towards some ideas and keeping the process transparent and fair.

Given that scoring is time consuming, it was considered reasonable to allow the scorers about a month to reply, with two further reminders sent at monthly intervals after the deadline. After 3 months, the scoring process should typically be considered completed, the drop–out rate recorded, and the analyses can begin. The process of analysis of the scores is described in great detail in another paper [10].

EXAMPLE FROM THE CHNRI EXERCISE ON RESEARCH PRIORITIES FOR NEWBORN HEALTH

This study has been published in its extended form in this theme issue [7]. Although the field of newborn health in low–income settings is very recent and the research community is still quite small, and although the process of involving researchers followed many steps that were in common to the exercise on pneumonia 5 years earlier, several important innovations were introduced.

Similarly to the pneumonia exercise, the management group selected the 200 most productive researchers, based on the number of co–authored publications in peer–reviewed journals in the previous 5 years. However, the composition of those 200 researchers was more targeted in this case: in addition to inviting the 100 most productive researchers on newborn health globally, the 50 most productive researchers affiliated to institutions in low and middle–income countries (LMIC) were also invited. The final 50 invitations were reserved for the most productive researchers in the area of stillbirth research globally. The purpose of this approach to sampling was to avoid under–representation of researchers from LMIC and the small number of researchers who worked on the increasingly important issue of stillbirths. This was a carefully thought–through approach and is another example of the flexibility allowed in the CHNRI process. It is important to "design" the sampling process in a way that captures researchers who could be most informative for the specific exercise, which is likely to be more important for exercises that are very broad in scope and less important for those which are very narrow.

Photo: Researchers in Bangladesh working in their laboratory (Courtesy of Dr Ozren Polašek, personal collection)

Another innovation in this newborn health exercise was the inclusion of

programme managers, identified through the Healthy Newborn Network database. This was a suggestion made by several members of the management board in light of broad agreement that "description" research was no longer a priority and that the new focus should be on implementation. Therefore, the group recognised the need to include experts with first–hand understanding of the challenges with delivery, cost and sustainability of newborn health and stillbirth prevention programmes in LMIC settings. This resulted in about 600 potential scorers being invited to participate in the exercise, of which the majority (400) were program managers familiar with the challenges in low–resource settings. Eventually, 132 persons participated in the generation of ideas and 91 in scoring, bringing the final response rate to about 15%.

Another innovation in this exercise was the use of "Survey Monkey", which allowed the management group to keep track of the age, gender, geographic area, background and affiliation of each participating researcher/programme manager in real time. This innovation was seen as very useful, because it allowed more intense reminders that were being sent to specific groups of invitees who were falling behind and becoming under–represented.

To improve the response rate, the management team sent four and five reminders to the invitees for both submitting the ideas and the scores. The team met in Geneva for a week to consolidate the initial list of research ideas they had received from about 400 down to about 200 that were eventually scored. In summary, this exercise stands out in three ways: (i) the targeted sampling of researchers; (ii) the inclusion of programme managers as the majority of invited scorers, to better reflect the community with useful knowledge on the criteria, which is not necessarily reflected in academic articles; and (iii) the tracking of score responses in real time using survey monkey [7].

EXAMPLE FROM THE CHNRI EXERCISE ON RESEARCH PRIORITIES FOR DEMENTIA

The examples on childhood pneumonia and newborn health are both relevant to research fields with relatively small research communities. In both exercises, the CHNRI method was used primarily as a way to galvanise the community and define the strategy for the development of the field. The small number of productive researchers in both fields meant that nearly everyone who had contributed to the research field over the previous 5 years was invited to participate in the exercise. However, how should we select researchers when the research field is very large and has tens of thousands of actively participating researchers? One such recent example is the CHNRI exercise on dementia and Alzheimer disease, a field in which tens of thousands of researchers are active. This exercise represents a good example of the strategies that can be used to solicit input from researchers in such circumstances.

The management group numbered 15–20 members at various stages of the process and included representatives of the World Health Organization, several international societies and funders interested in this topic (eg, Alzheimer Disease International, USA–based Alzheimer Association, UK's National Institute for Health Research, Canadian Institute for Health Research and USA–based National Institute of Aging), together with leading researchers and opinion–leaders in the field who were based in academic institutions (Rudan I, personal communication). This diverse group needed to devise a plan for recruiting a large number of researchers to provide research ideas and scores for the vast multi–disciplinary field of dementia and Alzheimer disease research. They held several meetings and teleconferences during which they discussed the best strategy to address this difficult task.

Their discussions soon focused on finding the proper justification for inviting some researchers, while leaving many thousands of others outside of the exercise. The group started to look for an appropriate response to a likely *post–hoc* question "*Why wasn't I invited to participate, and other colleagues were?*" that would eventually be acceptable to all those who might ask this question. The group eventually agreed that a justification that was likely to be accepted by researchers in this area should have the following format: "*You were not invited because: (i) you were not among the most productive 500 researchers (in terms of the number of publications) in this field in the past 5 years; (ii) you were neither the lead, nor the senior author on any of the 50 most cited papers in each of the past 5 years; and (iii) you don't belong to any of the groups of researchers specifically targeted for inclusion (even if they do not fall into the first two categories); this mainly relates to the few researchers from low– and middle–income countries (LMICs)*".

Given that the line of whom to invite needs to be drawn somewhere, the CHNRI management group agreed that the justification provided above would have a good chance for being accepted by the entire research community. Indeed, if a researcher isn't among the 500 most productive in the field in the previous 5 years, they cannot easily take an issue over those 500 more productive researchers being invited. Moreover, if a researcher hasn't led the research on a paper that was later ranked among the 50 most cited papers on the topic in each of the 5 previous years, then they cannot easily take an issue over the invitation of those 500 further authors who were in this position (5 years × 50 papers × (1 lead +1 corresponding author) = 500 authors). This rule implied that up to 1000 researchers would be invited to participate – some based mainly on their productivity in this field, and others mainly on high impact of their work, with some overlap expected between the two groups.

Finally, given that the exercise was global in terms of geographical scope, and that the vast majority of the most productive and/or cited authors were based in wealthy countries, the group concluded that every effort should be invested to identify the third group to invite – composed of an unrestricted, but likely quite small number of prominent published researchers based in low– and middle–income countries, which would be sought for through a separate effort.

The productive authors for the first group were identified through a search of Web of Sciences' "Core Collection", which ranked all researchers in the world in the field of dementia or Alzheimer disease by the number of publications, limited to the output in the preceding 5 years (2009–2013). This allowed the CHNRI management group to identify the 500 most productive researchers. The group also needed to check and merge results for the same author who published with different initials (ie, interchangeably using only one or both initials in their papers). The contact details were then successfully obtained from their publications for a sizeable subset, although not for all. This potentially introduced a bias related to dropping those who couldn't be contacted from further stages of the process.

The group then used Web of Science's "Core Collection" to rank the papers published in each of the years 2009–2013 by the number of citations that each paper received by the end of 2014. For the 50 most cited papers in each year, the group identified the lead and the corresponding author (ie, the first and last listed). After removing duplicate entries – because some authors would be found on several such papers, and then also on the previous list of the most productive authors – the identified authors would be invited to participate in the exercise wherever their contact details could be found. All duplicates were removed, but the "new free places" were not filled with further scientists, because the justifications for inclusions were pre–set and it was not clear whether to keep filling the places based on productivity, citations, or some other criterion. This meant that the final number of invited researchers would decrease from 1000 to a smaller number. Due to the overlap, the described process yielded 672 researchers to be contacted.

In addition, Chinese databases were systematically searched. The papers published in those databases didn't have many citations (as checked through Google Scholar), so the ranking of papers by citations received could not have been used as a selection criterion in a truly meaningful way. The group therefore invited the most productive 50 authors from the Chinese literature over the preceding 5–year period (2009–2013). To identify the few researchers from other low– and middle–income countries, the Alzheimer Association, Alzheimer Disease International (ADI, which is the global umbrella organization of all national Alzheimer associations) and 10/66 dementia research group (broad network of researchers from low and middle income countries) were actively involved in identifying and contacting the experts in LMIC. In the end, about 800 researchers were identified for contact, and the contact details were successfully obtained for 69% of them, each of whom was asked to submit 3–5 research ideas. Then, a total of 201 experts responded and submitted 863 research ideas. Those ideas pertained to prevention, diagnosis, treatment or care for dementia and represented "basic", "clinical–translational" or "implementation" research, as categorized by the management group. The management group then decided that this number was too large to score, so they convened a meeting to review all received research ideas. They consolidated the list to 59 representative "research avenues/themes", which were broader than specific research ideas/questions. These broader avenues/themes were then scored using a slightly modified set of the 5 standard CHNRI criteria. Thus, this exercise developed not only an approach to the sampling of experts when a very large number of experts exists in the world, but also developed an approach to deal with an unmanageable number of specific research ideas/questions received from such a large expert group. It is possible that, in the final version of the published paper (which is now still under review), some minor practical modifications from this protocol will be observed (Rudan I, personal communication).

ETHICAL AND OTHER CONSIDERATIONS

Given that the CHNRI method essentially relies on input from human subjects (who are researchers in this case), we consider here the ethical aspects of conducting CHNRI exercises. The CHNRI exercises are a form of research that uses various measures of collective opinion as an output – eg, the level of collective support for a particular research idea, the extent of agreement within the collective, the variance in all expressed opinions, the average level of support across several criteria, and possibly others. Nevertheless, the input is based on individual opinions received from individual participants.

The method itself, as initially proposed [10], underwent ethical scrutiny at the institution where it was conceived – at the Croatian Centre for Global Health at the Faculty of Medicine of the University of Split, Croatia. The following recommendations were made:

(i) It is important to let all participants know, at the stage of inviting them to participate in the CHNRI exercise, that by responding to the invitation through submitting their ideas, and then their numerical scores, they acknowledge their voluntary participation in the exercise; this will deal

with the ethical concern over whether their participation is voluntary, and they would not need to sign a special informed consent;

(ii) Although the input received from the participants is encoded as a sequence of numbers (the scores), if it is presented in the supplementary material of the resulting papers under the scorers' personal names or surnames, and aligned against the research ideas that were scored, this can still be used to reconstruct their personal opinions on a wide range of research topics; this may make the participants (ie, scorers) uncomfortable. Therefore, unless specific approval is obtained at the individual or a group level to disclose all individual scores in the interest of transparency of the CHNRI exercise (which is a motivation that can be seen as being in conflict with ethics concerns in this case), we recommend that all scores disclosed in the public domain through publications should be anonymized. If the scores received from the scorers are anonymized in a proper way, and only the opinion of the entire collective is studied and interpreted, there should not be any ethical concerns related to the CHNRI exercise.

(iii) We see another theoretical ethical concern that should potentially be carefully managed; namely, if all participants and their scores are disclosed in the public domain, and the participants haven't been anonymised at their own request (ie, in the interest of transparency and legitimacy of the CHNRI exercise), then the participants should still be warned that further statistical analyses could potentially be performed on the data set that involves their names. Those analyses could focus on participants themselves as subjects, and "ranking" and comparisons among the participants, rather than research ideas. Therefore, everyone's input could be statistically compared to that of one or more other participants. Although this is never the intention or a focus of the CHNRI exercise, it is a theoretical possibility and it could identify some scorers as "outliers" in terms of scoring with respect to their colleagues, which may cause them an unforeseen concern.

If these theoretical concerns are appropriately addressed and managed, which can most easily be achieved through informing the participants of the scope of the exercise, explaining that by self–selecting themselves for the exercise they are acknowledging their voluntary participation, and anonymising their scores once they are received, the CHNRI method should be considered free from ethics concerns.

The managers of CHNRI exercises often ask whether the results of the exercise should be returned to all participants. We endorse this practice, because we can see no reason why this should not happen. It is in everyone's interest to inform them of the collective optimism/pessimism towards various research ideas within each research community, especially when the participants have freely offered their ideas and time for scoring.

This brings us to another frequent question, which is how to thank the participants for their contributions in terms of suggesting research ideas and dedicating their time to scoring? In the vast majority of the previously conducted CHNRI exercises, this was done through involving the participants in the resulting publication. This involvement could either take the form of equal co–authorship, or listing under the group co–authorship, or simply acknowledging their contribution in the acknowledgement at the end of the paper. The decision as to which of these three options to employ typically depends on the number of participants, the realistic prospects in involving them in other stages of writing of a resulting CHNRI publication (beyond purely providing the scores), and the preferences, restrictions, or authorship criteria of the journals to which the papers have been submitted. It is also possible to motivate the participants to participate in the CHNRI exercise by organising a meeting in a convenient location and supporting participants' travel and accommodation expenses, and then conduct the entire exercise over a few days in a location of preference or convenience. In some cases, this has been done to expedite the scoring process when speed is important as exercises can take quite a long time when conducted via e–mail [4–8].

CONCLUSIONS

To date, we have gained considerable experience with involving researchers as participants who provide research ideas and scores for the CHNRI exercises. We have tried to summarise some informative examples in this paper, irrespective of whether the chosen examples were necessarily the most successfully conducted CHNRI exercises. Indeed, it is difficult to judge whether the CHNRI exercise has been "successful", and what criteria should be used to do so. Clearly, a high participation rate should limit the scope for response bias (through self–selection), which is a major concern with CHNRI exercises. Then, a large and broadly inclusive spectrum of research ideas provided by participants and made available for scoring would certainly signal a success in conducting the exercise, although it is difficult to quantify this inclusiveness. Moreover, it would reflect researchers' willingness to share their ideas freely and take part in the process. Large differences in the final research priority scores (RPSs) received by various research ideas indicate that the criteria used are able to discriminate between ideas. If an exercise results in only small differences in RPSs then any ranking of research ideas based on the scores is unlikely to be very robust, and the exercise will have largely failed to meet its own objectives.

Finally, if the exercise is conducted reasonably quickly (typical time is about 3–6 months) and at low cost (typical direct financial costs are up to US$ 15000, unless the costs of organizing one or more meetings are envisaged), and all participants accept the results and co–author a resulting publication, then the exercise has served its purpose. This will be even more so whenever there is a vision of a follow–up to the exercise, in which a workshop is organised to arrange research proposal writing, or a special meeting with the funders is agreed to ensure that the priorities have been properly communicated. Dissemination of the results and an appropriate follow–up at national, regional and global levels are important parts of the CHNRI process, to increase the likelihood that the research on identified priorities is conducted in the near future. Evaluating whether CHNRI exercises have had an impact on those who invest in health research and influenced investment decisions is challenging and is will be addressed in future papers on the CHNRI method.

Acknowledgements: We sincerely thank the representatives of the researchers for taking part in the previous CHNRI exercises and providing useful information. SY is the employee of the World Health Organization. Her views expressed here do not necessarily represent the views of the Organization.

Funding: None.

Authorship declaration: SY wrote the article under supervision by her PhD mentor, Professor Igor Rudan. SC, KW, KYC revised the article, made many useful comments and provided important intellectual content.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare no conflict of interest.

REFERENCES

1 Rudan I, Yoshida S, Chan KY, Cousens S, Sridhar D, Bahl R, et al. Setting health research priorities using the CHNRI method: I. Involving funders. J Glob Health. 2016;6:010301. Medline:26401269 doi:10.7189/jogh.06.010301

2 Yoshida S, Rudan I, Cousens S. Setting health research priorities using the CHNRI method: VI. Quantitative properties of human collective opinion. J Glob Health. 2016;6:010503. doi:10.7189/jogh.06.010503

3 Brown KH, Hess SY, Boy E, Gibson RS, Horton S, Osendarp SJ, et al. Setting priorities for zinc–related health research to reduce children's disease burden worldwide: an application of the Child Health and Nutrition Research Initiative's research priority–setting method. Public Health Nutr. 2009;12:389-96. Medline:18426636

4 GAPPS Expert Group on Community Based Strategies and Constraints, George A, Young M, Bang A, Chan KY, Rudan I, et al. Setting implementation research priorities to reduce preterm births and stillbirths at the community level. PLoS Med. 2011;8:e1000380. Medline:21245907 doi:10.1371/journal.pmed.1000380

5 Rudan I, Theodoratou E, Zgaga L, Nair H, Chan KY, Tomlinson M, et al. Setting priorities for development of emerging interventions against childhood pneumonia, meningitis and influenza. J Glob Health. 2012;2:010304. Medline:23198129 doi:10.7189/jogh.01.010304

6 Bhutta ZA, Zipursky A, Wazny K, Levine MM, Black RE, Bassani DG, et al. Setting priorities for development of emerging interventions against childhood diarrhoea. J Glob Health. 2013;3:010302. Medline:23826502 doi:10.7189/jogh.03.010302

7 Yoshida S, Martines J, Lawn JE, Wall S, Souza JP, Rudan I. Setting research priorities to improve global newborn health and prevent stillbirths by 2025. J Glob Health. 2016;6:010508. Medline:26401272 doi:10.7189/jogh.06.010508

8 Wazny K, Sadruddin S, Zipursky A, Hamer DH, Jacobs T, Källander K, et al. Setting global research priorities for integrated community case management (iCCM): Results from a CHNRI (Child Health and Nutrition Research Initiative) exercise. J Glob Health. 2014;4:020413. Medline:25520803 doi:10.7189/jogh.04.020413

9 Rudan I, El Arifeen S, Bhutta ZA, Black RE, Brooks A, Chan KY, et al. Setting research priorities to reduce global mortality from childhood pneumonia by 2015. PLoS Med. 2011;8:e1001099. Medline:21980266 doi:10.1371/journal.pmed.1001099

10 Rudan I, Gibson JL, Ameratunga S, El Arifeen S, Bhutta ZA, Black M, et al. Setting priorities in global child health research investments: guidelines for implementation of CHNRI method. Croat Med J. 2008;49:720-33. Medline:19090596

Correspondence to:
Dr Sachiyo Yoshida
Department for Maternal, Newborn, Child and Adolescent Health
World Health Organization
Avenue Appia 20, CH–1211
Geneva 27
Switzerland
yoshidas@who.int

Setting health research priorities using the CHNRI method: III. Involving stakeholders

Sachiyo Yoshida[1*], Kerri Wazny[2*], Simon Cousens[3], Kit Yee Chan[2,4]

[1] Department for Maternal, Newborn, Child and Adolescent Health, World Health Organization, Geneva, Switzerland
[2] Centre for Global Health Research, The Usher Institute for Population Health Sciences and Informatics, The University of Edinburgh, Edinburgh, Scotland, UK
[3] Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK
[4] Nossal Institute for Global Health, University of Melbourne, Melbourne, Australia
[*] These authors contributed equally to the work.

Setting health research priorities is a complex and value–driven process. The introduction of the Child Health and Nutrition Research Initiative (CHNRI) method has made the process of setting research priorities more transparent and inclusive, but much of the process remains in the hands of funders and researchers, as described in the previous two papers in this series [1,2]. However, the value systems of numerous other important stakeholders, particularly those on the receiving end of health research products, are very rarely addressed in any process of priority setting. Inclusion of a larger and more diverse group of stakeholders in the process would result in a better reflection of the system of values of the broader community, resulting in recommendations that are more legitimate and acceptable.

The CHNRI method, as originally proposed, took into account the importance of stakeholders and made provisions for their participation in the process. Although the involvement of a large and diverse group of stakeholders is desirable, they were not expected to propose research ideas, or score them against the set of pre–defined criteria. Because of this, the original CHNRI method proposed that stakeholders should be allowed to "weigh" pre–defined criteria and set "thresholds" for a minimum acceptable score against each criterion that would be required for a research idea to be considered a "research priority". In choosing the stakeholders, the context of each exercise will be very important and the goals of the specific exercise should be defined before choosing an appropriate "stakeholder group". Among stakeholders, we would expect to see those affected by the disease of interest and their family members, their carers and health workers, members of general public, media representatives interested in the topic, community leaders, representatives of the consumer groups and industry, but also potentially researchers and funders themselves. Although the latter two groups – researchers and funders – already have a different role assigned in the CHNRI process, this does not exclude them from also being stakeholders in the process [1,2]. In this paper, we aim to review and analyse the experiences in stakeholder involvement across the 50 CHNRI exercises published in the 10–year period between 2007 and 2016, the proposed approaches to involving stakeholders and their effects on the outcome of the prioritization process.

One paper in the original CHNRI method series focused on involving stakeholders [3]. That paper presented practical experiences from three separate attempts to involve stakeholders that took place in 2006. The three groups approached were: (i) members of the global research priority setting network; (ii) a diverse group of national–level stakeholders from South Africa; and (iii) participants at a con-

Setting health research priorities is a complex and value–driven process. The introduction of the CHNRI method has made the process of setting research priorities more transparent and inclusive, but much of the process still remains in the hands of funders and researchers. However, the value systems of numerous other important stakeholders, particularly those on the receiving end of health research products, are very rarely addressed in any process of priority setting. Including a larger group of stakeholders in the process would result in a better reflection of the system of values of the broader community, resulting in recommendations that are more legitimate and acceptable.

ference related to international child health held in Washington, DC, USA. Each group was asked to complete a short questionnaire to assess the relative importance of the five original CHNRI criteria. Different versions of the questionnaire were used with each group [3]. The results of this exercise indicated that groups of stakeholders vary in the weights they assigned to the 5 criteria, reflecting divergence in the "value" placed on each criterion by each stakeholder group.

The diverse group of respondents within the priority–setting network placed the greatest weight on the criterion of "maximum potential for disease burden reduction" and the most stringent threshold on "answerability in an ethical way". Among the attendees at the international conference on child health, the criterion of "deliverability, answerability and sustainability" was identified as the most important. Finally, in South Africa, where inequity has been a national problem historically, the greatest weight was placed on the "predicted impact on equity" criterion.

This comparative analysis by Kapiriri et al. [3] effectively demonstrated that involving a wide range of stakeholders is an important goal for any research priority setting exercise. The criteria that may be of importance to funders, scientists and other technical experts involved in the process of planning and conducting the exercise may not be well aligned with the values of those who should eventually benefit from health research, or with the sentiments of wider society as a whole [3]. This is an important observation, because if the CHNRI process is conducted without regard for the broader social value or research then it is unrealistic to expect it to fulfil its purpose of being accepted as a fair, transparent and legitimate process for setting investment priorities for health research.

THE CONCEPTS OF THRESHOLDS AND WEIGHTS IN THE CHNRI METHOD

These concepts were introduced as a part of the initial CHNRI method description [4,5]. The multi–disciplinary working group that developed the CHNRI method recognised the need to find a practical way to involve a much larger group of stakeholders in the priority–setting process. An agreement was reached that, at least in principle, most members of the public would not be expected to generate research ideas or score them, because they do not possess the knowledge that would enable them to discriminate among the proposed research ideas. Instead, it was agreed that their contribution to the process and the final results of the exercise would be in the assignment of "weights" to the criteria that reflect their collective preferences and beliefs. Over the years of CHNRI implementation, it has been shown that stakeholders originating from funding institutions or political organizations prefer the criterion of maximum potential for disease burden reduction, because their targets are usually set around this criterion; programme managers are typically more focused on the deliverability and sustainability criterion; stakeholders from the industry tend to prefer knowing the likelihood of effectiveness of resulting interventions; while members of the general public often emphasize equity and ethics as their preferred criteria [6].

In addition to placing more "weight" on some criteria than others, which could affect the final rankings of *all* research ideas as a result of stakeholders' input into the CHNRI process, the stakeholders can also disqualify *some* research ideas using the system of "thresholds". This means they may agree a priori that a research idea will not be considered a priority unless it reaches a certain minimum score against a particular priority–setting criterion. This can be important in a specific context; eg, in the aforementioned example of South Africa, where equity was a very important concern for all stakeholders, they could have insisted that a research idea must have a minimum score of 80% on the "equity" criterion to qualify as a priority. In practice, this means that a research idea with scores 50–70% on all other criteria, but 90% on "equity", could be considered a research priority. However, another idea with scores of 80–90% on all other criteria, but 60% on "equity" would be disqualified from the exercise – or at least delayed, until it addresses the recognized issues with equity. Common examples of the latter are the new, high technology–based interventions that would likely first be utilised by the wealthy. In this way, research ideas with lower overall

The original CHNRI method proposed that large and diverse groups of stakeholders should "weigh" different criteria according to their perceived value and importance for society as a whole. They were asked to set "thresholds" for minimum acceptable scores for each of the pre–defined criteria. In this paper, we aim to review and analyse the experiences with stakeholder involvement across the 50 CHNRI exercises published in the 10–year period between 2007 and 2016, the proposed approaches to involving stakeholders and their effects on the outcome of the prioritization process.

scores could be seen as greater priorities if they pass all the pre–defined "thresholds" [3,4].

Although the interdisciplinary group that developed the CHNRI method considered this approach as practical and inclusive, the question remained of how best to select the stakeholders and ensure their representativeness to the entire community of interest. Possibly the best solution to this problem to date has been achieved by Kapiriri et al. [3] who aimed to develop a "global" group of stakeholders by conducting an internet–based survey of the affiliates to the "Global research priority setting network", which had been assembled in the years prior to the development of the CHNRI method by the staff from the University of Toronto, Canada. Between March and May 2006 a large number of affiliates to the "Global research priority setting network" agreed to participate in a pilot on the condition of anonymity. They agreed to provide stakeholder input to five forthcoming exercises that aimed to set research priorities to address the five major causes of global child mortality. Respondents included a very diverse mix of researchers, policymakers and health practitioners with an interest in priority setting in health care from high–, middle– and low–income countries. Participants were given a simple version of the questionnaire, and were asked to rank the five "standard" CHNRI criteria from 1st to 5th in the order of their perceived importance of the criteria. They were also asked to set a threshold for each of the five criteria. The respondents placed the greatest weight (1.75) on potential for disease burden reduction, while the weights for the remaining four criteria were similar to each other, and ranged between 0.86 to 0.96. The highest threshold was placed on the criterion of answerability in an ethical way (0.54), while the lowest was placed on potential for disease burden reduction (0.39).

CASE STUDIES OF STAKEHOLDER INVOLVEMENT IN CHNRI EXERCISES

We identified 50 research prioritization exercises using the CHNRI method that were published between 2007 and 2016. Of the 50 exercises, 38 (76%) did not seek inputs from stakeholders and 12 (24%) involved stakeholders as their larger reference group. This already shows how it may be remarkably difficult in most cases to identify and involve an appropriate group of stakeholders that would be representative of the wider community of interest – whether this is a global, regional, national or local population. It seems that, in the absence of simple solutions, most authors who conducted the CHNRI exercises preferred not to include stakeholders in the process, rather than including an ill–defined and non–representative group and then having to adjust the final ranks based on their input. By not including input from stakeholders, the CHNRI exercises simply remained "unfinished" to an extent, though weights and thresholds could still be applied *post–hoc* should an appropriate group of stakeholders be identified at some later stage – unless the context changes substantially in the meanwhile.

Among the 12 CHNRI exercises that involved stakeholders and took their input into account, 5 were papers that belonged to the series of exercises related to addressing research priorities for the five major causes of child mortality globally – eg, pneumonia, diarrhoea, neonatal infections, preterm birth/low birth weight, and birth asphyxia [7–11]. All of these papers were co–ordinated by the World Health Organization (WHO) and they used the weights and thresholds defined above by Kapiriri et al. [3]. However, the remaining seven exercises made their own individual attempts, using guidelines for implementation of the CHNRI method, to identify appropriate stakeholders within their own contexts and involve them in the process. This section explores the experiences and results from these seven studies. **Table 1** summarizes the approaches to involving stakeholders in these seven exercises.

Two exercises were carried out at the global level. They were focused on mental health research and acute malnutrition in infants less than six months, respectively [12,13]. The remaining five exercises were conducted at the national level and focused on research in child health in South Africa [14], zoonotic disease in India [15], health policy and maternal and child health in China [16,17], and Prevention of Mother–to–Child Transmission of HIV (PMTCT) in Malawi, Nigeria and Zimbabwe [18]. Given that the large majority (over 80%) of the 50 CHNRI exercises were focused on either the global context, or on all low– and middle–income countries (LMIC), the high representation of national–level exercises among those CHNRI studies that

Table 1. Summary tables on the involvement of stakeholders

Reference	Profiles and mode of identification	Number of stakeholders	Responsibility	Criteria	Weights and thresholds applied to the criteria	Impact of stakeholders' involvement on the final scores
[12]	Psychiatrists (9), psychologists (4), social workers (2), government employees (3), non–governmental organization representatives (6), researchers (6), users of mental health services (6) and members of the public service (7), including those from low–and middle–income countries; No indication as to how they were identified and selected	43	They were asked to rank the five pre–defined criteria with range of 1 to 5 (1–highest rank to 5–lowest rank)	5 standard CHNRI criteria used [4]	Weights were assigned based on ranking: effectiveness (+21%), maximum potential for burden reduction (+17%), deliverability (+0%), equity (–9%), answerability (–19%); Thresholds not applied	There was no description whether the ranks significantly differed between non–weighted and weighted scores
[13]	Mostly researchers and policy makers; also included technical experts, senior practitioners in the area of nutrition and child health (including 9 members of "MAMI" groups: Management of Acute Malnutrition for Infant less than six month reference group). Above profiles included all the participants and there was no clear description of the profile of stakeholders. Identified from the participants at meetings, symposia related to the technical area of concern	64	They were asked to score the research questions against the pre–defined criteria, rather than place weights on the criteria	5 standard CHNRI criteria (two composite criteria split into two – 7 in total) [4]	Weights and thresholds not applied	See main text: the stakeholder group was used for scoring, rather than weighting
[14]	Researchers, academics, clinicians, government officials, clinical psychologists, and member of the public. Identified based on their availability and accessibility with an attempt to ensure diversity of the group	30	Same as reference [12]	5 standard CHNRI criteria used [4]	Weights were defined using the rank given to the 5 pre–defined criteria: equity (+30%), efficacy and effectiveness (+9%), deliverability, affordability and sustainability (+2%), maximum potential for disease burden reduction (–9%), answerability and ethics (–19%); Thresholds not applied	The paper presented both the weighted and non–weighted scores. The stakeholders' inputs changed the ranking of the research options somewhat, but the top 20 research options remained the same in both cases
[15]	Scientists, students and lay people. Identified from staff members of the Public Health Foundation of India (PHFI) and those identified through personal networks of authors	Not mentioned	They are asked to rank the pre–defined five criteria from most important (ranked 1) to least important (ranked 5) within the national context	5 standard CHNRI criteria used [4]	Weights were defined using the rank given to five pre–defined criteria: deliverability, affordability (+18%), maximum potential for disease burden reduction (+18%), efficacy and effectiveness (+13%), equity (–17%) and answerability and ethics (–18%); thresholds not applied	The final outcome was not affected by the stakeholders' inputs on the criteria in that the top 15 research options remained the same across weighted and non–weighted scores
[16]	Managers from medical institutions, doctors, patients, and representatives of public (5 representatives of each group). Method of identification not mentioned	20	They were asked to rank the and also provide the thresholds on the pre–defined five criteria. However it was unclear whether or not other participants also provided the ranking to the criteria	5 criteria used: potential to affect change, maximum potential for disease burden reduction, deliverability, economic feasibility and equity	Weights: Potential to affect change (0.1925), maximum potential for disease burden reduction (0.1925), deliverability (0.2160), economic feasibility (0.1890) and equity (0.2050); Thresholds: Potential to affect change (33.5%), maximum potential for disease burden reduction (29.7%), deliverability (27.0%), economic feasibility (28.0%) and equity (27.8%).	It was unclear whether any major differences in the ranks were observed after applying the weights and thresholds

Table 1. Continued

Reference	Profiles and mode of identification	Number of stakeholders	Responsibility	Criteria	Weights and thresholds applied to the criteria	Impact of stakeholders' involvement on the final scores
[17]	Obstetricians, gynaecologists, paediatricians, representatives of patients group, industry and international organizations; mode of identification was not mentioned	19	They were asked to rank the and also provide the thresholds on the pre–defined ten criteria	10 criteria used: answerability and ethics, efficacy and effectiveness, deliverability, maximum potential for disease burden reduction, equity, acceptability, sustainability, translation to policy, and economic feasibility and equity	Weights: answerability (0.11), efficacy and effectiveness (0.09), deliverability (0.10), maximum potential for disease burden reduction (0.14), equity (0.11) acceptability (0.07), sustainability (0.11), translation to policy (0.10), economic feasibility (0.10) and equity (0.07). Thresholds: answerability (33%), efficacy and effectiveness (38%), deliverability (28%), maximum potential for disease burden reduction (29%), equity (29%), acceptability (41%), sustainability (33%), translation to policy (33%), economic feasibility (40%) and equity (38%)	It was unclear whether any major differences in the ranks were observed after applying the weights and thresholds
[18]	The article addressed three country–led research prioritization exercises. In each country, stakeholders were researchers, academics, policy makers, district health workers, frontline health workers, implementing partners, people living with HIV/AIDS; mode of identification was not mentioned	40 to 70 participants each in Malawi, Nigeria and Zimbabwe	Stakeholders participated in the entire process ie, generation of research ideas and the scoring of research ideas. The weighting of scores was not applied in the exercise, because all stakeholders participated in the entire process.	6 criteria were used: answerability and ethics; potential maximum disease burden reduction on paediatric HIV infections; addresses main barriers to scaling–up; innovation and originality; equity; and likely value to policy makers	Weights and thresholds not applied	This exercise included diverse group of stakeholders. In this regard the relevance of the research ideas identified in the respective exercise to the national context was high.

used stakeholders input (5/12) is likely a reflection of the fact that it is much easier to involve stakeholders at the national or sub–national level than it is on a regional or global level.

In all exercises, the stakeholders involved were first given an induction course about the CHNRI process. Then, an opportunity for asking and sharing questions and concerns with respect to the CHNRI process was provided. In five of the seven exercises, stakeholders were asked to rank the relative importance of the pre–defined criteria from most important one ("1") to the least important ("5"), while considering the context of the research prioritization. The average score was calculated for each criterion and was then used to calculate the relative weights by dividing the average expected score of 3.0 (ie, the average expected rank if all criteria were valued the same) by the mean assigned rank. For example, a mean assigned rank for "answerability" criterion of 2.47 translates a relative weight of 1.21 (ie, 3.00/2.47 = 1.21). In this way, "answerability" will receive 21% greater weight than if all the criteria were weighted equally.

The concept of thresholds was very rarely used. Even when it was applied, it was clear that it wasn't properly explained to participating stakeholders. This is not surprising, because the thresholds really refer to a measure of "collective optimism" of the scorers, rather than a real computation of likelihood or probability that is rooted in any real–world parameters. It is very difficult to estimate what this measure of "collective optimism" could amount to for different criteria. This is why such attempts to set thresholds typically resulted in them being set at 25%–30%, much too low to have any discriminatory power and disqualify many research ideas, so that almost all research ideas passed all the thresholds.

In the remaining two exercises, the nature of stakeholder involvement was modified radically from that which was originally envisaged in the CHNRI exercises with reasonable justification [13,18]. Instead of using the group of stakeholders only to adjust the ranks that were derived from an expert–driven scoring process, the authors involved a broad range of stakeholders in the generation of research ideas [18] and/or scoring the research ideas [13,18]. We will now reflect on these experiences in a critical way, identify some lessons learnt and make recommendations for future exercises.

CRITICAL ASSESSMENT OF STAKEHOLDER INVOLVEMENT IN CHNRI EXERCISES

In the 7 studies that tried to develop a larger reference group of stakeholders that would be appropriate to their respective contexts, the number of stakeholders involved was disappointingly small: it ranged from 20 to 70. Although attempts were clearly made to ensure diversity of the stakeholders involved, such small sample sizes can hardly be considered sufficiently inclusive of many different groups of stakeholders and their representativeness. Although good representativeness of stakeholders can be ensured without necessarily requiring a very large number of participants – such as, eg, in many examples of national parliaments in democratic societies, who represent all the people of the nation through a relatively small number of their elected members – we still feel that bigger numbers would ensure more legitimacy to the process, or more relevance of the outcomes to the context of the exercise.

It would be difficult to consider the examples in the reviewed exercises as truly representative of the wider communities, let alone the nation or the world. This shows that despite the authors' best intentions to fully adhere to the guidelines and complete the CHNRI process, they didn't really manage to find a satisfactory solution to involving large and diverse group of stakeholders. In these papers, the profile of stakeholders often included researchers, who would have been better reserved for the scoring process. Other stakeholders included clinicians, government officials, and representatives of academia and professional organizations, which again are rare in the society and hardly representative of the wider community. The examples of the profiles of persons who we would expect included in the larger reference group are also laypersons, frontline health workers and direct beneficiaries of health services, such as patients who contracted disease of concern. We encourage the authors of the future CHNRI exercises to try to get as much feedback as possible from those groups, because they have their own specialised knowledge (including lived experience), which would not be captured by other participating groups in the process. They also have "stake", or interest, in the outcome of the exercise.

The small sample sizes and differences in approaches to ensure diversity and representativeness of the stakeholders led to large variations in stakeholders' input [12–18]. In the global exercise, the greatest relative importance was assigned to effectiveness, and the lowest to answerability, though these results should not be generalized. Stakeholders at the national level varied in their preferences, alternately supporting the criteria equity, deliverability (with affordability and sustainability), or the maximum potential for disease burden reduction (**Table 1**). Clearly, small sample sizes used in these exercises limit the generalizability of such preferences even within their local context, let alone more broadly.

It is also important to note that in all exercises that applied the "weights", this procedure didn't really have dramatic effects on the final rankings of the research ideas. Although a research idea might move a few places up or down the list following the weighting procedure, these shifts did not profoundly affect the non–weighted ranking order that was determined by the researchers and experts. Perhaps this is one of the additional reasons why so many groups conduct-

Photo: Meeting with a group of stakeholders at the maternity health clinic in Ghana (Courtesy of Dr Alice Graham, personal collection)

ing the CHNRI exercise did not place sufficient importance on involving stakeholders. From the exercises that involved stakeholders, one might conclude that the process of expert scoring is sufficient and the outcome of the exercise will not be greatly altered by the involvement of stakeholders. We believe that such a view is premature and would like to see more examples of the involvement of the stakeholders in the CHNRI process before such judgements could be made.

In two exercises that actively involved stakeholders, their involvement wasn't limited to weights or thresholds, but rather they were also involved in research idea generation and scoring [13,18]. In the exercise on PMTCT in three African countries [18], about 40–70 people took part in respective countries, and all participants contributed to all stages of the CHNRI process. This included academics/researchers, district health workers and implementing partners such as UN agencies, people living with HIV/AIDS, frontline health workers and policy makers. The authors' justification for including these diverse groups in all stages of the CHNRI process was to avoid discriminating within this diverse range of groups, but to truly engage the groups according to their technical expertise and to enhance inclusiveness and participation in similar priority–setting exercises across the nation. Eventually, the stakeholders' weighting of the scores was not even applied, possibly due to an assumption that it was no longer needed. This example represented a rather interesting deviation from the original CHNRI conceptual framework, but we can see a rationale for this modification, which makes it an illuminating exception.

The other exercise, on the management of acute malnutrition in infants in low– and middle–income countries, involved stakeholders only in the scoring process [13]. The stakeholder group included participants at meetings and symposia related to the topic area (**Table 1**). In this exercise, the core group of researchers ("management team") developed the list of research questions based on the review of the literature in this field that preceded the CHNRI exercise as the preparatory step. The final list of questions was then circulated for scoring to both researchers invited to the CHNRI process and also the conference participants, who were considered stakeholders. Equal weighting was given to all criteria. The management team justified this on the grounds that malnutrition was a new area of research in infants younger than 6 months and they therefore believed that unweighted estimates would be more suitable and interpretable by their intended policy–maker audience. However, the authors stated that the lack of weighting of criteria might have resulted in limited reflection of the values in the broader community. In this case, we can conclude that the borderline between the invited researchers and the "stakeholders" (who were likely to include unrelated researchers and any other people of similar profile who could be expected to attend an international conference in this topic), was blurred and not really clear. It is likely that this deviation from the suggested approach didn't really invalidate the conceptual framework, because all the scorers would still be expected to possess knowledge on the topic of interest. It would perhaps be more appropriate not to call the second group "stakeholders", but rather an additional, "convenience" sample of scorers that increased the number of scorers considerably.

PROPOSED SOLUTIONS AND WAY FORWARD

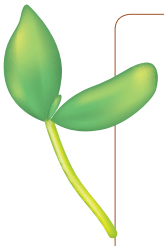
So far, there hasn't really been a good example of stakeholder involvement as originally envisaged by the CHNRI across the first 50 implementations, apart from perhaps the Kapiriri's priority–setting network involvement that was used in 5 child mortality papers [3,7–11]. This is certainly a shortcoming of all the previously conducted processes. This finding may also reinforce the initial concern that involving stakeholders in research priority setting processes is very challenging and that the solutions proposed in the original CHNRI method were quite difficult to implement as envisaged.

This is not to suggest that the results of previous CHNRI exercises are not useful, and the thresholds and weights can be applied later, if a good solution to obtain them can be found within the time scale during which the context described to scorers would still remain largely unchanged. The efforts conducted to date to perform the CHNRI exercises were not wasted and their results can be used. However, it must be acknowledged that most CHNRI exercises to date are, in fact, incomplete at least with respect to the original vision for them. To bridge this gap better definition is needed of who are the stakeholders at different levels (ie, global, regional, national and local) and how best to represent them.

For global exercises, we'll inevitably need a very large and inclusive crowd–sourcing exercise of many stakeholder representatives, who would place weights and thresholds on all 25 priority–setting criteria that were used to date across all 50 CHNRI exercises (5 "standard" and 20 new). The sample of stakeholders will need to be truly large, because we may later need several sub–samples that could provide us with region–specific stakeholders, or allow selecting specific groups of stakeholders and leaving others out of the exercise. In this way, the large "global" sample of stakeholders would also serve as a base for the regional samples of stakeholders. A major concern relating to this suggested approach would be how to avoid a strong urban bias in low–income settings and be inclusive of un-

developed and/or rural areas. In terms of national–level or local–level exercises, it is likely that highly targeted samples that aimed to include 500–1000 stakeholders would already be sufficient and representative of national or local context. The "targeting" component of the sampling strategy would define the profile of the stakeholders that would be most appropriate to the exercise, and then a person could be found in the community to fit each such profile.

How could these large sample sizes be achieved technically? How could we engage thousands of people globally, or hundreds nationally? With further attention to the development of the area of "crowd–sourcing" in the age of the internet and social networks (such as Facebook, Twitter, etc.), we should be able to do lot more in the future with respect to truly engaging the stakeholders in the process of setting priorities in health research investments at different levels of the human population.



Acknowledgements: We sincerely thank the representatives of the stakeholders for taking part in the previous CHNRI exercises and providing useful information. SY is the employee of the World Health Organization. Her views expressed here do not necessarily represent the views of the Organization.

Funding: None.

Authorship declaration: SY and KW wrote the article. KYC and SC revised the article and made useful comments.

Competing interests: All authors have completed the Unified Competing Interest form at www. icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare no conflict of interest.

REFERENCES

1 Rudan I, Yoshida S, Chan KY, Cousens S, Sridhar D, Bahl R, et al. Setting health research priorities using the CHNRI method: I. Involving funders. J Glob Health. 2016;6:010301. doi:10.7189/jogh.06.010301

2 Yoshida S, Cousens S, Wazny K, Chan KY. Setting health research priorities using the CHNRI method: II. Involving researchers. J Glob Health. 2016;6:010302. doi:10.7189/jogh.06.010302

3 Kapiriri L, Tomlinson M, Gibson J, Chopra M, El Arifeen S, Black RE, et al. Child Health and Nutrition Research Initiative (CHNRI): Setting priorities in global child health research investments: Addressing the values of the stakeholders. Croat Med J. 2007;48:618-27. Medline:17948948

4 Rudan I, Gibson JL, Ameratunga S, El Arifeen S, Bhutta ZA, Black M, et al. Child Health and Nutrition Research Initiative (CHNRI). Setting Priorities in Global Child Health Research Investments: Guidelines for implementation of the CHNRI Method. Croat Med J. 2008;49:720-33. Medline:19090596 doi:10.3325/cmj.2008.49.720

5 Rudan I, Gibson J, Kapiriri L, Lansang MA, Hyder AA, Lawn J, et al. Child Health and Nutrition Research Initiative (CHNRI). Setting priorities in global child health research investments: Assessment of principles and practice. Croat Med J. 2007;48:595-604. Medline:17948946

6 Rudan I. The complex challenge of setting priorities in health research investments. Indian J Med Res. 2009;129:351-3. Medline:19535827

7 Fontaine O, Kosek M, Bhatnagar S, Boschi–Pinto C, Chan KY, Duggan C, et al. Setting research priorities to reduce global mortality from childhood diarrhoea by 2015. PLoS Med. 2009;6:e1000041. Medline:19278292 doi:10.1371/journal.pmed.1000041

8 Rudan I, El Arifeen S, Bhutta ZA, Black RE, Brooks A, Chan KY, et al. Setting research priorities to reduce global mortality from childhood pneumonia by 2015. PLoS Med. 2011;8:e1001099. Medline:21980266 doi:10.1371/journal.pmed.1001099

9 Lawn JE, Bahl R, Bergstrom S, Bhutta ZA, Darmstadt GL, Ellis M, et al. Setting research priorities to reduce almost one million deaths from birth asphyxia by 2015. PLoS Med. 2011;8:e1000389. Medline:21305038 doi:10.1371/journal.pmed.1000389

10 Bahl R, Martines J, Ali N, Bhan MK, Carlo W, Chan KY, et al. Research priorities to reduce global mortality from newborn infections by 2015. Pediatr Infect Dis J. 2009;28:S43-8. Medline:19106763 doi:10.1097/INF.0b013e31819588d7

11 Bahl R, Martines J, Bhandari N, Biloglav Z, Edmond K, Iyengar S, et al. Setting research priorities to reduce global mortality from preterm birth and low birth weight by 2015. J Glob Health. 2012;2:010403. Medline:23198132 doi:10.7189/jogh.01.010403

12 Tomlinson M, Rudan I, Saxena S, Swartz L, Tsai AC, Patel V. Setting priorities for global mental health research. Bull World Health Organ. 2009;87:438-46. Medline:19565122 doi:10.2471/BLT.08.054353

REFERENCES

13 Angood C, McGrath M, Mehta S, Mwangome M, Lung'aho M, Roberfroid D, et al. Research Priorities to Improve the Management of Acute Malnutrition in Infants Aged Less Than Six Months (MAMI). PLoS Med. 2015;12:e1001812. Medline:25898252 doi:10.1371/journal.pmed.1001812

14 Tomlinson M, Chopra M, Sanders D, Bradshaw D, Hendricks M, Greenfield D, et al. Setting priorities in child health research investments for South Africa. PLoS Med. 2007;4:e259. Medline:17760497 doi:10.1371/journal.pmed.0040259

15 Sekar N, Shah NK, Abbas SS, Kakkar M; Roadmap to Combat Zoonoses in India (RCZI) Initiative. Research options for controlling zoonotic disease in India, 2010–2015. PLoS ONE. 2011;6:e17120. Medline:21364879 doi:10.1371/journal.pone.0017120

16 Yang M. Zhao K, Qi X, Li X. Determining the priority issues of health policy based on CHNRI. Chinese Journal of Health Policy. 2015;8:74-9.

17 Li X, Zhao Z, Qi X, Xiao Y, Guo W. Determining the priority issues of maternal and child health system in Western China based on CHNRI. Chinese Health Service Management. 2014;4:244-7.

18 Rollins N, Chanza H, Chimbwandira F, Eliya M, Nyasulu I, Thom E, et al. Prioritizing the PMTCT Implementation Research Agenda in 3 African Countries: INtegrating and Scaling up PMTCT Through Implementation REsearch (INSPIRE). J Acquir Immune Defic Syndr. 2014;67 Suppl 2:S108-13. Medline:25310115 doi:10.1097/QAI.0000000000000358

Correspondence to:
Sachiyo Yoshida
Department for Maternal, Newborn, Child and Adolescent Health
World Health Organization
Geneva, Switzerland
yoshidas@who.int

Addressing the growing burden of non–communicable disease by leveraging lessons from infectious disease management

Peter Piot[1], Aya Caldwell[2], Peter Lamptey[3], Moffat Nyrirenda[4], Sunil Mehra[5], Kathy Cahill[6], Ann Aerts[7]

[1] London School of Hygiene and Tropical Medicine, London, UK
[2] Innovative Healthcare Delivery Solutions, Novartis Foundation, Basel, Switzerland
[3] Family Health International 360, Accra, Ghana
[4] Malawi Epidemiology and Intervention Research Unit, Malawi
[5] MAMTA Health Institute for Mother and Child, India
[6] PATH, Seattle, Washington, USA
[7] Novartis Foundation, Basel, Switzerland

In recent decades, low– and middle–income countries (LMICs) have achieved decreased morbidity and mortality associated with infectious diseases and poor maternal– and child–health (MCH). However, despite these advances, LMICs now face an additional burden with the inexorable rise of non–communicable diseases (NCDs).

Deaths due to NCDs in LMICs are expected to increase from 30.8 million in 2015 to 41.8 million by 2030 [1]. While improvements in life expectancy, lifestyle and urbanisation go some way to explaining why more people in LMICs are affected by NCDs, it is less clear why these populations are contracting NCDs at a younger age and with worse outcomes than in high–income countries (HICs) [2]. Despite having a lower cardiovascular disease risk factor burden, LMIC populations have a four–fold higher mortality rate from cardiovascular events than HIC populations [3] in part due to a lack of access to quality, integrated health services and the poor availability of early interventions and effective NCD prevention programmes. The HIV/AIDS epidemic was the last time the world confronted a global health challenge that so disproportionately caused premature adult deaths in LMICs.

The conclusion is unavoidable: the time to act is now. Prevention of NCDs at a population and an individual level is key and requires policy and structural changes. We have a unique opportunity to learn from the successes of infectious disease control programmes in LMICs and leverage these to address the growing NCD burden. Translatable learnings include: 1) emphasizing primary prevention, particularly in those at highest risk; 2) targeting service delivery to high–risk populations; 3) enabling access to adequate, affordable care at community level; 4) engendering patient empowerment and involving people affected by chronic conditions; 5) enabling access to quality drugs and adherence programmes; 6) regularly measuring the effectiveness and impact of programmes to ensure their appropriateness and improvement; and 7) creating an environment of health financing for universal coverage.

> Despite advances in decreasing morbidity and mortality associated with infectious diseases and poor maternal– and child–health low– and middle–income countries now face an additional burden with the inexorable rise of non–communicable diseases.

Innovations to counter the emerging NCD epidemic must encompass both prevention and the delivery of

care. Infectious disease programmes have used task–shifting, where less skilled health workers and community members are involved in delivery of health services. In India, we have seen this used for NCDs in the Arogya Kiran model where the existing health workforce was overstretched. Volunteers and teachers successfully delivered diabetes and hypertension screening and management to over 600 000 people [4]. Patient empowerment, and community involvement in health care delivery and governance, will be critical in tackling NCDs, since most are chronic conditions, which initially present silently and require long–term management [5].

In Malawi, recognizing the close relationship of HIV infection and cardiovascular diseases has led to screening for hypertension being integrated into HIV care [6]. In Ghana, decentralised community–based hypertension care, using digital technology, is helping to empower patients to manage their own disease: a model that is again adapted from HIV management [7]. We are also starting to see examples in India of MCH care coupled with life–long NCD screening and awareness programmes [8].

While these examples of managing the dual burden of infectious diseases and NCDs are encouraging, more needs to be done. The largest gap is in NCD prevention. Tackling the obesity epidemic and wrestling with the issues around curbing tobacco sales and smoking are rightly high on the NCD prevention agenda. The greatest opportunity is preventing a tobacco–related epidemic in sub–Saharan Africa where smoking levels are still low. Health budgets and development assistance for health must allocate resources commensurate with the dual disease burden. Health spending of governments in LMICs has tripled over the past 20 years, but remains low [9]. In addition, more health care models should consider diversified revenue streams or hybrid financing (eg, tiered payment schemes) to ensure sustainability. If equity is to be improved, patients need access to quality health care, through sustainable health–financing systems for universal health coverage, while reducing out–of–pocket expenditure for the under–served population.

> We have a unique opportunity to learn from the successes of infectious disease control programmes and leverage these to address the growing non–communicable disease burden: 1) emphasizing primary prevention; 2) targeting service delivery to high–risk populations; 3) enabling access to adequate, affordable care; 4) engendering patient empowerment; 5) enabling access to quality drugs and adherence programmes; 6) regularly measuring the effectiveness and impact of; and 7) creating an environment of health financing for universal coverage.

Implementing such models will require strong government leadership and interventions, and partnerships across the public and private sectors. Some public–private partnerships (PPPs) in infectious diseases have demonstrated their potential to catalyze the delivery of, and access to, prevention and care through providing complementary strengths [10]. The private sector draws on its business and scientific expertise, focusing on strong results–based operations, whereas the public sector brings a wealth of expertise in implementation with equity, management and documentation.

The end–users of the services, including patients and health care providers, also need to be included from the outset to ensure that the models are people–centered, co–created, adapted to prevailing contextual nuances, and sustainable. If we build on what we have learnt from infectious disease management, we could have a transformational impact on the growing NCD burden.

Photo: © Nana Kofi Acquah/Novartis Foundation'

Funding: The Novartis Foundation, a philanthropic organisation, provided external agency support (services provided by 90TEN Healthcare) for literature searches and for the preparation of an outline of the paper.

Authorship declaration: The opinion expressed in this article was developed following a Novartis Foundation workshop held in December 2014–the first in a series of Basel Dialogues on global health issues. The expert faculty discussed lessons learned and experience gained in managing infectious diseases and how these might inform future strategies on the emerging dual burden of disease (infectious and non–communicable) in low– and middle–income countries. All authors contributed equally to the article. The work to develop the article content was that of the contributing authors.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). We declare no competing interests. The Novartis Foundation provided no funding support to the non-Novartis Foundation authors for the preparation of the article and did not influence the content of the article. Two of the authors Dr. Ann Aerts, Head Novartis Foundation, and Aya Caldwell, Head Innovative Healthcare Delivery Solutions, are employed by the Novartis Foundation. Peter Piot has received grants from Gates Foundation, IMI, DFID and UNAIDS outside the submitted work and is a Board member for the Novartis Foundation. Peter Lamptey receives support from the Novartis Foundation for the Community based control of Hypertension Improvement Project (Com-HIP) and is the Principal Investigator for the evaluation of the programme. Kathy Cahill is Vice President International Development at PATH, which received a grant from the Novartis Foundation for a project in Vietnam focused on prevention of CVD.

REFERENCES

1 World Health Organization. Projections of mortality and causes of death, 2015 and 2030. Available: http://www.who.int/healthinfo/global_burden_disease/projections/en/. Accessed: 1 July 2015.

2 Alwan A, MacLean DR. A review of non–communicable disease in low– and middle–income countries. Int Health. 2009;1:3-9. Medline:24036289 doi:10.1016/j.inhe.2009.02.003

3 Yusuf S, Rangarajan S, Teo K, Islam S, Li W, Liu L, et al. Cardiovascular risk and events in low–, middle– and high–income countries. N Engl J Med. 2014;371:818-27. Medline:25162888

4 Sogarwal R, Mehra R. Arogya Kiran model for early detection of diabetes and hypertension: an initiative for the community and by the community in India. BMC Health Serv Res. 2014;14 Suppl 2:113. Medline:24602215 doi:10.1186/1472-6963-14-S2-P113

5 The Economist Intelligence Unit. Sub–Saharan African healthcare: the user experience. A focus on non–communicable diseases. 2014. Available: http://www.economistinsights.com/sites/default/files/NCD%20Africa%20report.pdf. Accessed: 1 July 2015.

6 Muronya W, Sanga E, Talama G, Kumwenda JJ, van Oosterhout JJ. Cardiovascular risk factors in adult Malawians on long–term antiretroviral therapy. Trans R Soc Trop Med Hyg. 2011;105:644-9. Medline:21924753 doi:10.1016/j.trstmh.2011.07.016

7 FHI 360. Community–based Hypertension Improvement Project (ComHIP). Available: http://www.fhi360.org/projects/community-based-hypertension-improvement-project-comhip. Accessed: 1 October 2015.

8 MAMTA Health Institute for Mother and Child Health. Community engagement and health systems. Available: hhttp://mamta-himc.org/ncd/. Accessed: 1 October 2015.

9 Institute for Health Metrics and Evaluation. Financing global health 2013: transition in an age of austerity. Seattle: IHME, 2014.

10 Dholakia YN. TB/HIV coordination through public private partnership: lessons from the field. Indian J Tuberc. 2013;60:23-7. Medline:23540085

Correspondence to:
Aya Caldwell
Novartis Foundation
Novartis Campus
Forum 1-3.92
4002 Basel
Switzerland
aya.caldwell@novartis.com

Pro–poor pathway towards universal health coverage: lessons from Ethiopia

Kesetebirhan Admasu[1], Taye Balcha[1], Tedros Adhanom Ghebreyesus[2]

[1] Ministry of Health, Addis Ababa, Ethiopia
[2] Ministry of Foreign Affairs, Addis Ababa, Ethiopia

Protection from care–related catastrophic expenditures through equitable access to affordable health services is the hallmark of a pro–poor health policy [1]. Over the past two decades, the Government of Ethiopia has implemented policies with a clear intent of reducing poverty and improving the daily lives of its citizens, especially the poor [2]. Guided by these cross cutting pro–poor government policies and spurred by the United Nations Millennium Development Goals (MDGs), the health sector has implemented multi–pronged initiatives towards ensuring every citizen an access to affordable health services without catastrophic expenditures (**Table 1**). The health sector initiatives have been guided by evolution of the innovative health programs nationally introduced as well as the needs of the community in each village across the country [3,4]. Primary focus on the poor and ownership of new health initiatives by the community have been the linchpin for investment at scale.

All–inclusive, community–led primary health care is the bedrock of the health services in Ethiopia. Notwithstanding its low–income status, the country has gradually but radically expanded access to a spectrum of health services and essential medicines. More than 38 000 all-female Health Extension Workers (HEWs) have been deployed to more than 16 500 community health posts across the country to lead a novel primary health care– Health Extension Program (HEP). Launched in 2003, the HEP has brought simple, cost–effective and locally–desired health interventions [4] close to where the country's majority, rural citizens live. The service package includes maternal and child health (family planning, antenatal services, immunization, nutrition services, treatment of infectious childhood con-

Table 1. Major pro–poor policies and initiatives in Ethiopia

Pro–poor initiatives	Year implemented	Objective	Key outcomes
Introduction of Health Extension Program	2003	To achieve universal primary health care coverage which mainly benefits low–income households	More than 38 000 health extension workers have been deployed in 16 500 villages of the country; and universal primary health care coverage has been achieved.
Establishment of Pharmaceuticals Fund and Supply Agency	2006	To ensure accessibility and affordability of essential medicines and laboratory investigations	Medicines and laboratory investigations for key health conditions have been provided free of charge; out–of–pocket expenditures have reduced; health services utilization has improved; and health MDG targets have been met.
Health Development Army Program with community soolidarity fuding	2012	To disseminate health information and facilitate uptake of critical health services and finance priority challenges identified by the community	Procured more than 200 ambulance vehicles for medical referral; constructed health posts and maternity waiting homes at rural health centers; and Health Development Armies have actively involved in health facility governance to improve the quality of health services.
Scaling up Community–based Health Insurance scheme	2015	To provide quality health care without financial hardship to the poor in informal sector	By the end of 2016, 50% of citizens in informal sector are expected to be covered.
Implementation of Social Health Insurance scheme	2016	To deliver quality health care and ensure financial protection to citizens employed in formal sector and achieve universal health coverage	All employees of formal sector are expected to be covered by the end of 2016.

ditions), prevention, and sometimes treatment of communicable diseases (tuberculosis, HIV and malaria) and environmental sanitation. All health services provided by HEWs at health post level are free of charge. The government carries the brunt of the financial costs for these services but it is commonly supplemented by innovative community financing. For example, while the government has assigned trained and salaried HEWs as civil servants in each village and provided ambulance vehicles at district level, the community has constructed health posts and maternity waiting homes at health centers, and has procured additional ambulance vehicles for medical referral through Health Development Army [3]. The new government–community partnership has improved service uptake and health outcomes. A study in Northern Ethiopia attributed a substantial reduction in pregnancy–related mortalities to a wide availability of ambulance services for obstetric referral [5].

Eyeing Universal Health Coverage (UHC), Ethiopia has intensified its implementation of pro–poor initiatives. It is currently rolling out the second generation HEP to meet the growing needs of the community. The second generation program adds more interventions targeted on emerging infectious diseases, common non–communicable diseases and, mental health. It also includes deepening the partnership with the community even beyond primary care level. The government has put a mechanism to stimulate innovations improving access to and utilization of health services by the poor [6].

Whereas the HEP remains country's priority, considerable efforts have been made to improve the higher levels of health care. For example the number of health centers and hospitals in the country has sharply increased by 350% and 150%, respectively between 2004 and 2015. Correspondingly, an enormous development has been made in human resources for health through increasing the number of medical universities. Tuition fees and lodging expenses have been covered by the government in all public medical universities and colleges; and a pay–by–service strategy has been implemented whereby health professionals eventually return the cost by rendering health services at public health facilities. As a consequence, the number of medical doctors graduated annually in the country greatly increased, from about 150 in 2004 to 3000 in 2016. Similarly, unprecedented increase has been seen in the number of other critical cadres including mid–wives and specialized nurses. To achieve the right balance between public health and clinical interventions, the country trained innovative cadres like public health officers and integrated emergency surgical officers. Integrated emergency surgical officers perform major surgeries for emergency obstetric and surgical conditions, close to where the rural poor live [7].

> Protection from care–related catastrophic expenditures through equitable access to affordable health services is the hallmark of a pro–poor health policy.

> All–inclusive, community–led primary health care is the bedrock of the health services in Ethiopia. Notwithstanding its low–income status, the country has gradually but radically expanded access to a spectrum of health services and essential medicines.

In 2006, Ethiopia established Pharmaceutical Fund and Supply Agency (PFSA) with the primary mission of expanding access to medicines, vaccines and laboratory services. Following its launch, PFSA has steered a comprehensive health supply chain management in the country. The agency has used pooled procurement strategy as a vantage to gain economies of scale. It has procured refrigerated trucks, constructed 17 cold room–installed hubs, at least one within 180 km radius of each health facility across the country to maximize the efficiency gains in pharmaceuticals distribution. These gains in efficiency have led to a huge decline in retail prices of medicines with consequential drop in the out–of–pocket expenditure [8], benefiting particularly the poor. Further, the total value of PFSA's products has sharply grown from US$ 100 million in 2008 to US$ 1.3 billion in 2015; it is projected to be more than US$ 2 billion by 2017.

A complete set of care for priority maternal and child health interventions (family planning, abortion care, labour and delivery, immunization and nutrition services), infectious diseases (tuberculosis, malaria and HIV) and diseases of poverty (onchocerciasis, podoconiosis and trachoma) are provided free of charge at all public service delivery points. Community needs, the ability to pay and potential population level impacts were factored in defining program services and medicines. The strategic removal of financial barrier to care for maternal and child conditions and major infectious diseases has significantly contributed to Ethiopia's recent achievements including meeting all health MDG targets.

Improved access to affordable health care embraces the budding non–communicable diseases including diabetes mellitus and cardio–vascular diseases. For instance, mark–up has been removed from insulin and medicines used for common malignant conditions. To address the growing demand and expectations of the public and the limited outlets for medicines and supplies for non–prioritised interven-

tions, Ethiopia is currently rolling out community pharmacies in cities and big towns across the country. Similar to health facility–embedded pharmacies, community pharmacies provide medicines at substantially reduced prices.

Ethiopia has established a stringent regulatory system to guard the poor against low standard and counterfeit medicines. Evidence shows that the comprehensive supply management and robust regulation has effectively blocked entry of counterfeit medicines into the market and significantly reduced stock–outs and wastage of medicines [9]. Furthermore, it has improved the affordability of medicines with consequential upsurge in health care utilization. In 2013, the availability of essential medicines at public health facilities in Ethiopia was 76%. The average price of generic medicines significantly reduced between 2004 and 2013, positively impacting the poor. During the same period, a considerable drop in price of medicines at private for–profit outlets was reported [10], implying the strategy's proxy impact on price regulation in the country's health care market, further contributing to the mitigation of catastrophic expenditures on the poor.

Although access to health care has been expanded, significant geographic disparities persist in regards to health care utilization and health outcomes. The inequity in health care is largely attributed to the lifestyles of communities. Pastoralist communities in Ethiopia are generally lagging behind in major health indicators. Further, differences in service uptake exist within communities, slowing Ethiopia's pro–poor, pro–equity route towards UHC.

Photo: Children at a vaccinations clinic near Sululta, Ethiopia. Yasmin Abubeker/DFID [CC BY-SA 2.0 (http://creativecommons.org/licenses/by-sa/2.0)], via Wikimedia Commons

The health sector is currently providing differentiated support to communities and regions left behind. More broadly, to ensure the universality of health coverage and prevent financial impoverishment, it is scaling up the successful Community–Based Health Insurance pilot for citizens in informal sector, where the poor predominate. Also, the country is introducing Social Health Insurance targeted at employees of formal sector for full coverage. Both schemes offer a package of clinical services without cost ceiling upon services provided at any domestic health facility. The two schemes are expected to cover 80% of the population within the next 5 years and will be consolidated as a single payer system within the next 10 years.

In conclusion, Ethiopia advances locally–tailored, multi–faceted pro–poor approaches to ensure UHC building on its successful transformation of the health sector in the last two decades and the achievement of key health MDG targets. Broader plan for inclusive economic development, effective implementation of primary health care, expansion of access to medicines and introduction of health insurance is the pathway towards UHC for all citizens. Concurrently, quality improvement initiatives and pushing non–communicable diseases to the forefront of the agenda are under way. More importantly, levelled partnership of the government with the community across the spectrum of health care ensures sustainability and community ownership of the system. We believe that these approaches to health care could propel Ethiopia to expedite its efforts to achieve the sustainable development goals including UHC within a short time.

Funding: No funding was obtained for this work.

Authorship declaration: All authors were involved in drafting the manuscript. All authors approved the final manuscript.

Competing interest: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). The authors declare no competing interests.

REFERENCES

1 DFID Health Systems Resource Centre. Health financing: designing and implementing pro–poor policies. London: DFID HSRC, 2001. Available: http://www.heart–resources.org/wp–content/uploads/2012/10/Health–financing.pdf. Accessed: 8 March 2016.

2 Ministry of Finance and Economic Development of Ethiopia. Building on Progress: A Plan for Accelerated and Sustained Development to End Poverty (PASDEP). Addis Ababa: Ministry of Finance, 2006. Available: http://www.afdb.org/fileadmin/uploads/afdb/Documents/Policy–Documents/Plan_for_Accelerated_and_Sustained_(PASDEP)_final_July_2007_Volume_I_3.pdf. Accessed: 8 March 2016.

REFERENCES

3 Admasu K, Balcha T, Getahun H. Model villages: a platform for community–based primary health care. Lancet Glob Health. 2016;4:e78-9. Medline:26718537 doi:10.1016/S2214-109X(15)00301-0

4 Ministry of Health of Ethiopia. Health extension program in Ethiopia. Addis Ababa: Ministry of Health, 2007.

5 Godefay H, Byass P, Kinsman J, Mulugeta A. Understanding maternal mortality from top–down and bottom–up perspectives : Case of Tigray. J Glob Health. 2015;5:010404. Medline:25674351 doi:10.7189/jogh.05.010404

6 Balcha T, Getahun H, Admasu K. Local innovations and country ownership for sustainable development. Bull World Health Organ. 2015;93:742. Medline:26549897 doi:10.2471/BLT.15.164483

7 Ghosh B. Non–doctor emergency surgeons are saving thousands of lives in rural Ethiopian hospitals. The Lancet Global Health Blog Available: http://globalhealth.thelancet.com/2016/01/13/non-doctor-emergency-surgeons-are-saving-thousands-lives-rural-ethiopian-hospitals. Accessed: 8 March 2016.

8 Pharmaceuticals Fund and Supply Agency of Ethiopia, US Agency for International, Development. National survey of the integrated pharmaceutical logistics system. Addis Ababa: Pharmaceuticals Fund and Supply Agency of Ethiopia, 2014.

9 Ethiopian Food, Medicines and Health Administration and Control Authority, World Health Organization. Assessment of substandard/counterfeit drugs in the Ethiopian pharmaceutical market. Addis Ababa: Ethiopian Food, Medicines and Health Administration and Control Authority, 2013.

10 United Nations. Taking stock of the global partnership for development. New York: UN, 2015.

Correspondence to:
Kesetebirhan Admasu
Ministry of Health
Addis Ababa, Ethiopia
moh@ethionet.et

Global disease burden due to antibiotic resistance – state of the evidence

Mark Woolhouse[1,2], Catriona Waugh[1], Meghan Rose Perry[1,3], Harish Nair[2]

[1]Centre for Immunity, Infection and Evolution, University of Edinburgh, Edinburgh, UK
[2]Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK
[3]Regional Infectious Diseases Unit, Western General Hospital, Edinburgh, UK

BURDEN OF ANTIBIOTIC RESISTANCE

Antimicrobial resistance is widely regarded as one of the major public health concerns of the 21st century [1,2], but there are no good estimates of the net global health burden due to resistance of bacteria to antibiotics. Although numerous studies have provided estimates of the burden of resistance of specific combinations of clinical disease, bacterial agent, antibiotic and health care setting (primarily hospitals in developed countries), metrics vary, coverage is patchy and methodologies are inconsistent. Such data have been used to obtain partial estimates of resistance–related mortality and other outcomes for Europe [3], the USA [4] and the world [5], but because of huge information gaps and the need to extrapolate from small–scale studies these estimates, though helpful, should be regarded as tentative at best.

Multiple metrics are used to quantify the "burden" of infectious diseases, including mortality, morbidity, disability adjusted life years, length of stay in hospital, or cost of care. Here we focus on mortality, although similar considerations apply to other metrics. An essential first step is to provide a clear definition of the burden of antibiotic resistance. We consider the most appropriate definition to be: *the number of deaths attributable to the failure of antibiotic therapy due to antibiotic resistance*. Importantly, this is not equivalent to the total number of deaths among patients with antibiotic resistant infections and may be much less than this for two main reasons: not all patients who may have resistant infections are treated with clinically indicated antibiotics and, for those that are, the measurable difference in outcome for patients with resistant vs susceptible infections may be relatively small.

More formally, this definition of burden can be expressed as a population attributable fraction (PAF, also referred to as the aetiological fraction), ie, the number of deaths that would not occur if antibiotic resistance were eliminated. As set out in the Box, to calculate PAF for mortality due to antibiotic resistance requires data not only on the number of patients with resistant infections and the number that die but also enumeration of the population of interest, which includes patients who survived and/or had susceptible infections. Enumeration of the population of interest in turn requires information on the incidence of the relevant clinical condition, its aetiology, and coverage of the antibiotic therapy of choice. Because such information is rarely available, PAF is rarely used to estimate the global burden of resistance; one recent example considered neonatal sepsis [2], but had to extrapolate key parameters from estimates obtained from a single hospital.

INCIDENCE AND AETIOLOGY

The main clinical conditions where antibiotic therapy can reduce mortality (**Table 1**) fall into three groups: commu-

The absence of comprehensive and reliable estimates of the global health burden due to antibiotic resistance makes it difficult to assess trends and harder to justify the allocation of adequate resources to deal with the problem.

Table 1. Common clinical conditions for which antibiotic therapy reduces the risk of mortality

Category of condition	Condition
Communicable diseases	Tuberculosis Sexually transmitted bacterial infections Respiratory bacterial infections (especially of the lower respiratory tract) Diarrhoea caused by bacteria* Healthcare associated bacterial infections
Endogenous infections	Urinary tract infections Skin and soft tissue infections Infective endocarditis Sepsis
Prevention of infection	Burns, wounds Caesarean sections Joint replacements Cancer therapy Organ transplants

*Antibiotics are not necessarily indicated for diarrhoeal cases.

nicable diseases, endogenous infections and prophylaxis to prevent endogenous infections in high risk patients. The number of patients in these categories defines the population potentially at risk of mortality attributable to antibiotic resistance. Relatively good incidence estimates are available for only some of these categories, notably tuberculosis and health care associated infections [4].

Of the clinical conditions listed in the Table only tuberculosis has a specific aetiology. The remainder are associated with multiple kinds of bacteria and several, such as sexually transmitted infections, diarrhoea and respiratory infections, may also be caused by viral and/or fungal agents.

A restricted set of both gram negative and gram positive bacterial agents, plus *Mycobacterium tuberculosis*, are commonly highlighted in the context of antibiotic resistance (eg, [4]). Some of these are of particular concern in hospital settings, such as *Acinetobacter* spp, *Enterobacteriaceae* spp, *Enterococcus* spp, *Pseudomonas aeruginosa*, *Staphylococcus aureus* and *Streptococcus* spp. Others are associated with communicable diseases typically acquired outside hospitals, such as *Campylobacter* spp, *Neisseria gonorrhoeae*, *Salmonella typhi*, non–typhoidal *Salmonella* spp, *Shigella* spp, and *Streptococcus pneumoniae*. Several of these contribute to multiple clinical conditions of interest.

Quantification of the burden of resistance requires data on the incidence of clinical conditions appropriately treated with antibiotics, the frequency of treatment failures due to resistance and their impact on clinical outcome. Treatment failures in turn depend on the level of resistance in the aetiological agent to the antibiotic used. These data are not easily obtained. One obstacle is that global health statistics as currently collected do not provide the necessary information.

Possible ways forward include making some categories of resistance notifiable, modifying the International Classification of Diseases, use of sentinel sites, and structured polling of clinicians.

ANTIBIOTIC USAGE

Global consumption of antibiotics has recently been estimated at more than 70 billion doses per annum [6]. By volume, antibiotic usage in 2010 was dominated by penicillins, cephalosporins, macrolides, fluoroquinolones, trimethoprim and tetracyclines.

These data refer to sales by pharmacies; they do not link antibiotic consumption to the treatment of patients with specific clinical conditions. The WHO last published generic guidelines for the therapeutic use of antibiotics in 2001 [7] but these and more current national and international guidelines tend not to be prescriptive, emphasizing the need to account for local circumstances, not least local patterns of antibiotic resistance. Usage profiles can thus vary considerably between locations. For some countries antibiotic usage data are available at hospital level; again however, these data are not routinely linked to information on the conditions that were being treated [8].

Current antibiotic usage profiles are, of course, influenced by current patterns of antibiotic resistance. Resistance patterns mean that, for example, aminopenicillins alone may not be used to treat serious gram negative bacterial infections, alternative drugs would be used additionally where available. In this scenario, aminopenicillin resistance does not contribute to the population attributable fraction as defined above, although it is arguably an element of the overall burden of antibiotic resistance.

ANTIBIOTIC RESISTANCE

The most comprehensive data on global levels of antibiotic resistance come from a recent WHO survey [9]. Even so, for most combinations of bacterial species and antibiotic the countries providing the minimum data required (testing of 30 isolates) accounted for less than half the world's population. A major contribution of this exercise was to

highlight significant variations in the kinds of isolates tested and in resistance testing protocols.

Moreover, bacteria–antibiotic combinations were not explicitly linked to clinical condition, so it is unclear when the resistances tested were clinically relevant and when they were not. This, together with the lack of data relating antibiotic usage to clinical condition, makes it difficult to estimate the relevant component of the PAF calculation, the fraction of patients with bacterial infections that are resistant to the antibiotic used to treat them (**Box 1**).

Box 1. Population attributable fraction (PAF) of mortality due to antibiotic resistance.

PAF calculations are a standard method of quantifying disease burden associated with a specified risk exposure [2], in this case bacterial infections resistant to the antibiotic used to treat them. The first step is to enumerate the population of interest. For current purposes, this would be the incidence (number per unit time) of patients with one of the clinical conditions of concern (see **Table 1**) and for whom antibiotic therapy is clinically indicated and is provided. The incidence of such patients is denoted *I*.

PAF calculation is routinely expressed in terms of the proportion of population exposed to the risk factor (here, patients with antibiotic–resistant infections) and the risk ratio for mortality standardised to the unexposed group (patients with antibiotic–susceptible infections) [2]. An equivalent, easily understood version is: PAF = (*IF*–*DR*)/(*ID*–*DR*), where *I* is the overall incidence (number of patients per unit time); *F* is the number of patients with resistant infections that die; *D* is the number of patients that die; *R* is the number of patients with resistant infections. If all deaths are associated with resistance (*F* = *D*) then PAF = 1; if deaths are not disproportionately associated with resistance (corresponding to *F* = *DR*/*I*) then PAF = 0. Importantly, PAF = 0 does not equate *F* = 0.

Intuitively, it seems natural to equate F with treatment "failures". However, some care is required because, given PAF<1, it is implicit that some of these patients (estimated as DR/I) would have died anyway, even if they had not had a resistant infection (this number reflecting the 'background' level of mortality observed in patients who were appropriately treated and had a susceptible infection). Similarly, of patients with susceptible infections who survive, some would have survived anyway, even had they had a resistant infection; that is, not all positive outcomes can be attributed to successful antibiotic therapy.

As detailed in the main text, although there is sometimes information available on *F*, *D* and/or *R*, there is often insufficient information to determine *I*. To do so requires additional data on one or more of the following: i) the total number of patients of interest that survive; ii) the number with susceptible infections; or iii) the number with susceptible infections that survive.

Obtaining a single global estimate of mortality attributable to antibiotic resistance presents the additional challenges of combining and extrapolating estimates of PAF for given combinations of clinical condition, antibiotic, aetiological agent and location.

TREATMENT FAILURE AND CLINICAL OUTCOME

Two key quantities for estimating the burden of antibiotic resistance are the frequency and clinical impact of failures of antibiotic therapy. Treatment failure is a complex phenomenon that may well be attributable to factors other than antibiotic resistance, including misdiagnosis. Treatment failure can also occur in patients with antibiotic–susceptible infections. Central to the calculation of burden is the distinction between the death of a patient who has an antibiotic resistant infection and the death of a patient that is *attributable* to having an antibiotic resistant infection (see Box).

Data on treatment failures are not routinely recorded. One source of data on mortality is the ICD–10 (International Classification of Disease, version 10) codes used by the WHO [10]. ICD–10 covers many, though not all, of the clinical conditions listed in **Table 1.** However, ICD–10 submissions do not usually include treatment failures associated with antibiotic resistant infections (reference to which is confined to the rarely used "Codes for Special Purposes"). Nor does the Institute for Health Metrics and Evaluation's Global Burden of Disease cause list have categories linked to antibiotic resistance [11].

RECOMMENDATIONS

Information currently collected at global or multi–national scales is not sufficient to generate estimates of the disease burden attributable to antibiotic resistance. As a result, current knowledge of the burden of antibiotic resistance is still based largely on the collation of one–off, small–scale, individual studies that vary greatly in setting, scope, sampling frame and methodology, and often requires bold extrapolations to be made from very limited data sets. For estimation of the global burden of antibiotic resistance and, even more, for monitoring changes in burden over time more systematic approaches would be helpful. There are several possibilities.

ICD–10 is due to be replaced by ICD–11 in 2017 [10]. This provides an opportunity to create routinely used categories that record treatment failures, or at least linking treatments with outcomes, the most direct ways to estimate the burden of antibiotic resistance. Specific concerns, such as XDR–TB or carbapenem–resistant *Enterobacteriaceae*, might be prioritised for inclusion.

ICD facilitates passive reporting. An alternative is active reporting by recruiting sentinel sites. For example, 660 hospitals from 67 countries responded to an internet survey on antimicrobial stewardship in 2012 [8]. Monitoring treatment failures due to antibiotic resistance in these hospitals using standardised protocols would generate valuable data.

Making selected, high priority antibiotic resistant infections 'notifiable' at national level could further improve data capture, extending existing mandatory reporting for specific conditions (for example in the UK for scarlet fever or invasive streptococcal group A disease). Another possibility is a more qualitative approach of recruiting a global panel of individual clinicians who are polled to determine trends in the impact of antibiotic resistance on their patients. Polling has been used successfully in other clinical contexts [12].

As well as estimating the global burden of antibiotic resistance another useful exercise would be to estimate the global burden due to lack of access to suitable antibiotics. For some clinical conditions, this may be a substantially greater burden at the present time [13]. The two issues potentially overlap where there is a lack of knowledge of local resistance profiles (perhaps due to lack of testing facilities) and alternative drugs would have been effective.

CONCLUSIONS

Estimation of the global burden of antibiotic resistance is extremely challenging and arguably not an attainable objective with currently available health data. We stress that this conclusion does not contradict the generally accepted view that antibiotic resistance is a major public health problem of global significance. There is a large number of studies documenting levels of resistance and its clinical impact, and well–founded concerns that both will rise, perhaps dramatically, in the foreseeable future. However, as reviewed here, the valuable insights provided by such studies do not sum to a comprehensive, coherent picture of the global antibiotic resistance burden and how it is changing. Improving this situation will require changes to the ways in which global health statistics are collected; existing approaches are not up to the task. The primary benefit will be more accurate assessment of the global disease burden due to antibiotic resistance and its forward trajectory, helping make the case for investment in combating the problem, and allowing assessment of future trends.

Photo: Courtesy of Kit Yee Chan, personal collection

Acknowledgements: We thank Ana Clayton–Smith for assistance with literature searches and Margo Chase–Topping for help with statistical analysis. This work was funded by the Wellcome Trust (grant number 093724) and by a European Union project grant (EvoTAR, grant number HEALTH–F3–2011–282004).

Funding: This work was funded by the Wellcome Trust (grant number 093724) and by a European Union project grant (EvoTAR, grant number HEALTH-FS-2011-282004).

Authorship declaration: MW conceived the study and led manuscript writing. CW carried out literature searches and collated data. MRP advised on clinical practice. HN advised on disease burden estimation. All authors contributed to manuscript preparation.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). All authors declare no competing interests.

REFERENCES

1 Woolhouse M, Farrar J. An intergovernmental panel on antimicrobial resistance. Nature. 2014;509:555-7. Medline:24877180 doi:10.1038/509555a

2 Laxminarayan R, Matsoso P, Pant S, Brower C, Rr˜ttingen JA, Klugman K, et al. Access to effective antimicrobials: a worldwide challenge. Lancet. 2016;387:168-75. Medline:26603918 doi:10.1016/S0140-6736(15)00474-2

3 European Centre for Disease Prevention and Control (ECDC), European Medicines Agency (EMEA). ECDC/EMEA joint technical report. The bacterial challenge: time to react. Stockholm: ECDC, 2009. Available: http://ecdc.europa.eu/en/publications/Publications/0909_TER_The_Bacterial_Challenge_Time_to_React.pdf. Accessed: 30 March 2015.

4 US Centers for Disease Control and Prevention. Antibiotic resistance threats in the United States, 2013. Available: http://www.cdc.gov/drugresistance/pdf/ar–threats–2013–508.pdf. Accessed: 31 March 2015.

REFERENCES

5 Review on Antimicrobial Resistance Antimicrobial resistance: Tackling a crisis for the health and wealth of nations. 2014. Available: http://amr–review.org/sites/default/files/AMR%20Review%20Paper%20–%20Tackling%20a%20crisis%20for%20the%20health%20and%20wealth%20of%20nations_1.pdf. Accessed: 11 December 2014.

6 Van Boeckel TP, Gandra S, Ashok A, Caudron Q, Grenfell BT, Levin SA, et al. Global antibiotic consumption 2000 to 2010: an analysis of national pharmaceutical sales data. Lancet Infect Dis. 2014;14:742-50. Medline:25022435 doi:10.1016/S1473-3099(14)70780-7

7 World Health Organisation. WHO model prescribing information: Drugs used in bacterial infections. 2001. Available: http://apps.who.int/medicinedocs/pdf/s5406e/s5406e.pdf. Accessed: 20 July 2015.

8 Howard P, Pulcini C, Hara GL, West RM, Gould IM, Harbarth S, et al. An international cross–sectional survey of antimicrobial stewardship programmes in hospitals. J Antimicrob Chemother. 2015;70:1245-55. Medline:25527272

9 World Health Organisation. Antimicrobial resistance: Global report on surveillance 2014. Available: http://apps.who.int/iris/bitstream/10665/112642/1/9789241564748_eng.pdf?ua=1. Accessed: 18 September 2014.

10 World Health Organisation. International Classification of Diseases, ICD–10 Version: 2015. Available: http://apps.who.int/classifications/icd10/browse/2015/en#/A40.0. Accessed: 8 April 2015.

11 GBD 2013 Mortality and Causes of Death Collaborators. Global, regional, and national age–sex specific all–cause–specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease study 2013. Lancet Infect Dis. 2014;385:117-71. Medline: 25530442

12 Eisinger F, Blay J-Y, Morère J-F, Rixe O, Calazel–Benque A, Cals L, et al. Cancer screening in France: subject's and physician's attitudes. Cancer Causes Control. 2008;19:431-4. Medline:18085414 doi:10.1007/s10552-007-9102-9

13 O'Brien KL, Wolfson LJ, Watt JP, Henkle E, Deloria–Knoll M, McCall N, et al. Burden of disease caused by Streptococcus pneumoniae in children younger than 5 years: global estimates. Lancet. 2009;374:893-902. Medline:19748398 doi:10.1016/S0140-6736(09)61204-6

Correspondence to:
Prof Mark Woolhouse
Centre for Immunity, Infection and Evolution
University of Edinburgh
Kings Buildings
Charlotte Auerbach Rd
Edinburgh EH9 3FL, UK
mark.woolhouse@ed.ac.uk

Electronic supplementary material:
The online version of this article contains supplementary material.

journal of global health

PAPERS

Effectiveness of mHealth interventions for maternal, newborn and child health in low– and middle–income countries: Systematic review and meta–analysis

Siew Hwa Lee[1], Ulugbek B Nurmatov[1], Bright I Nwaru[1], Mome Mukherjee[1,2], Liz Grant[3], Claudia Pagliari[1,3]

[1] eHealth Research Group, Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh Medical School, Edinburgh, UK
[2] Edinburgh Health Services Research Unit, Edinburgh, UK
[3] Edinburgh Global Health Academy, The University of Edinburgh Medical School, Edinburgh, UK

Objective To assess the effectiveness of mHealth interventions for maternal, newborn and child health (MNCH) in low– and middle–income countries (LMIC).

Methods 16 online international databases were searched to identify studies evaluating the impact of mHealth interventions on MNCH outcomes in LMIC, between January 1990 and May 2014. Comparable studies were included in a random–effects meta–analysis.

Findings Of 8593 unique references screened after de–duplication, 15 research articles and two conference abstracts met inclusion criteria, including 12 intervention and three observational studies. Only two studies were graded at low risk of bias. Only one study demonstrated an improvement in morbidity or mortality, specifically decreased risk of perinatal death in children of mothers who received SMS support during pregnancy, compared with routine prenatal care. Meta–analysis of three studies on infant feeding showed that prenatal interventions using SMS/cell phone (vs routine care) improved rates of breastfeeding (BF) within one hour after birth (odds ratio (OR) 2.01, 95% confidence interval (CI) 1.27–2.75, $I^2 = 80.9\%$) and exclusive BF for three/four months (OR 1.88, 95% CI 1.26–2.50, $I^2 = 52.8\%$) and for six months (OR 2.57, 95% CI 1.46–3.68, $I^2 = 0.0\%$). Included studies encompassed interventions designed for health information delivery (n = 6); reminders (n = 3); communication (n = 2); data collection (n = 2); test result turnaround (n = 2); peer group support (n = 2) and psychological intervention (n = 1).

Conclusions Most studies of mHealth for MNCH in LMIC are of poor methodological quality and few have evaluated impacts on patient outcomes. Improvements in intermediate outcomes have nevertheless been reported in many studies and there is modest evidence that interventions delivered via SMS messaging can improve infant feeding. Ambiguous descriptions of interventions and their mechanisms of impact present difficulties for interpretation and replication. Rigorous studies with potential to offer clearer evidence are underway.

Correspondence to:
Dr. Claudia Pagliari
eHealth Research Group
Usher Institute of Population Health Sciences and Informatics
The University of Edinburgh Medical School
Teviot Place
Edinburgh EH8 9AG, UK
claudia.pagliari@ed.ac.uk

Mortality in children under the age of five has fallen from an average rate of 90 per 1000 live births in 1990 to 43 in 2015, while maternal mortality has declined by 45% [1]. Despite these improvements, progress in

achieving Millennium Development Goals 4 and 5 fell short of expectations, and low– and middle–income countries (LMIC) still account for nearly all cases of maternal and neonatal mortality worldwide [2,3]. The availability and quality of maternal health care varies widely in different parts of the world and in LMIC women continue to die each year from preventable causes [4–6]. This is further compounded by limited resources and poor information infrastructures, which act as barriers to care coordination and quality, and hinder the effective management and governance of health systems [7–11].

mHealth, or mobile health, refers to the use of wireless, portable Information and Communication Technologies (ICT) to support health and health care [12]. There are numerous examples of mHealth interventions being used to support mothers through safe pregnancy and childbirth and to facilitate neonatal and infant health. Although scaled programmes do exist, the majority of mHealth projects in LMIC have tended to be small–scale donor–funded initiatives, which have taken place without the benefit of an adequate evidence–base [13].

A number of efforts have attempted to map the state of the evidence relating to mHealth for maternal, newborn and child health (MNCH) in LMIC, but no rigorous systematic reviews exist on this specific topic [14–16]. Philbrick's 'gap analysis', for the mHealth Alliance, combined literature review and stakeholder interviews [17], whilst literature reviews by Noordam et al. and Tamrat and Kachowski addressed the topic using simple search terms and a subset of available databases [18,19]. Free et al. reported two broader systematic reviews of interventions for patient behavior change and for health care service delivery processes and, while studies from LMIC were not excluded, the focus was higher income country settings [20,21]. In another mHealth report, Labrique et al. reviewed existing research for the purposes of developing a taxonomy of interventions [22]. While all of these provided valuable insights and recommendations, the World Health Organization (WHO) recognised the need for a rigorous systematic review when commissioning the current study. As we move on from the Millennium Development Goals and plan forward strategies for improving MNCH, mHealth is likely to play an increasing important role in light of continuing health needs and the growing global penetration of mobile technologies.

This study synthesized the evidence on the effect of mHealth interventions on MNCH in LMIC, with a particular focus on studies reporting impacts on patient outcomes.

METHODS

A detailed protocol was registered with the International Prospective Register for Systematic Reviews (PROSPERO) CRD42014008939 (http:///www.crd.yourk.ac.uk/prospero) and has been published [23]. The review is reported according to the requirements of the Preferred Reporting Items for Systematic Reviews and Meta–Analyses (PRISMA) [24]. We assessed studies that have investigated the effectiveness of mHealth interventions for improving MNCH in LMIC. LMIC were identified in accordance with World Bank country classifications [25]. The target groups were women in the antenatal, intranatal, and postnatal periods; newborns; children aged 0–5 years; and health workers through which interventions aimed at these groups are mediated. Men, non–pregnant women or those not recently having given birth, and children over the age of 5 years were excluded. We included studies evaluating interventions delivered through mobile ICT and considered the various delivery modes through which this might be achieved (**Box 1**). We excluded related ICT–based interventions delivered via fixed line internet or standard telephone line, interventions labeled 'mobile' which did not involve cellphones, such as Mobile Maternal Health Clinics which are touring buses staffed by health care professionals.

The primary outcomes were estimates of maternal, newborn and child mortality and morbidity. Secondary outcomes included number of planned antenatal and postnatal visits; number of unscheduled care visits and emergency care; quality of life; quality of care (delivery by skilled birth attendants, appropriate use of evidence–based medical and obstetric interventions); self–efficacy; cost–effectiveness; immunisation cover; child developmental milestones; and other process indicators.

Search strategy and study selection

16 international electronic databases were interrogated (**Box 2**) using highly sensitive search strategies implemented in OVID MEDLINE and then adapted to other databases (see Tables s1 and s2 in **Online Supplementary Document**). Searches were limited to articles published between

Box 1. Mobile ICT and delivery modes

Mobile ICT includes: cell–phones, smart–phones, satellite phones, personal digital assistants, enterprise digital assistants, tablet computers, laptops, portable media players and gaming consoles, Radio Frequency Identification Device (RFID) tags, Global Positioning System (GPS) trackers and digital diagnostic devices.

Mobile delivery modes includes: voice calling, Voice over Internet Protocol (VoIP), text messaging via Short Message Service (SMS), transfer of still or moving images via Multimedia Message Service (MMS), multimedia downloads, and live video.

Box 2. Sources of literature included in this systematic review and meta–analyses

Databases:

- Cochrane Library (Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Methodology Register),
- MEDLINE
- EMBASE
- CINAHL
- PsycINFO
- AMED
- Global Health
- TRIP
- ISI Web of Science (Science and Social Science Index)
- WHO Global Health Library
- IndMed
- PakMediNet
- KoreaMed
- NHS Health Technology Assessment Database
- African Index Medicus (encompassed in the WHO Global Health Library)
- POPLINE

Clinical trials registry for on–going studies and trial protocols:

- WHO International Clinical Trials Registry platform
- Clinical trials.gov
- Controlled–trials.com
- Australian New Zealand Clinical Trials Registry

Reference tracking:

- References list of all included studies

January 1990 and May 2014, acknowledging the emergence of digital cellular networks in the early 1990s [14]. The search strategies were piloted in order to optimise sensitivity and specificity. The decision was taken to dispense with country restrictions after finding that limiting searches to the LMIC countries specified in the World Bank's classification scheme had resulted in the omission of a highly relevant study from Zanzibar. (Although Zanzibar is part of Tanzania, which is listed, the word Tanzania did not appear in the title or abstract, hence the article was ignored.) There were no restrictions on language of publication. We included randomised controlled trials (RCTs), variations of RCTs, controlled before and after studies, interrupted time series studies and observational studies (cohort, case–control). We excluded cross–sectional and qualitative studies, expert opinions, reports, discussion papers, case reports, and studies from developed countries. Authors were contacted for access to unpublished research.

At least two reviewers independently screened the titles and abstracts of identified studies, assessed the full text of potentially eligible studies against the inclusion and exclusion criteria, and abstracted relevant study data onto a customised data extraction sheet. Country classification was undertaken by hand. Due to the large number of articles, and annual fluctuations in the World Bank index, a pragmatic decision was taken to include countries classed as LMIC at any time during the search period, or otherwise described using a phrase such as "developing country" (as described in the protocol).

Assessment of risk of bias

The methodological quality of intervention studies was assessed independently by at least two reviewers, following the recommendations of the Cochrane Effective Practice and Organization of Care Group [26]. Observational studies were assessed using the Effective Public Health Practice Project quality assessment tool [27]. Discrepancies were resolved by team consensus.

Meta–analysis

There was substantial heterogeneity between studies with regards to the mHealth interventions and study outcomes, except for the studies on breastfeeding (BF) and infant feeding [28–44]. Consequently, we performed a random–effects meta–analysis using the inverse variance method for three comparable studies, which had all used SMS/cell phone as the intervention vs routine prenatal care and had assessed breastfeeding as the primary outcome [30,33,42]. The study by Sellen et al. compared cell phone–based peer support, monthly peer support group and standard existing routine care for BF [42]. However in the meta–analysis we compared only the cell phone group with the routine care group as the relevant intervention for the review. The estimates of effect in the study by Sellen et al. were given as percentages [42], but we recalculated these into odds ratios with their 95% confidence intervals (95% CI) before the pooled analysis. Given the small number of studies in each meta–analysis, we did not explore reasons for the observed heterogeneity. For the same reason, we did not investigate the influence of publication bias or undertake possible sensitivity analyses. Meta–analyses were undertaken using STATA 11 (Stata Corp, College Station, Tx) [45].

RESULTS

Study selection and characteristics

Initial searches identified 12078 titles. After removing duplicates, 8593 papers were included for initial screening. Of these, 8401 papers were excluded after screening by title and abstract, leaving 192 papers, which were considered in more detail. A further 168 papers were subsequently excluded for not meeting the relevant criteria. 24 papers remained, and one additional paper was identified through

searching the reference lists of these papers. Of the 25 full–text papers, 17 met the inclusion criteria and were included in the final review (**Figure 1**). These were based on 15 primary studies [28-34,36–41,43,44], of which two were only available as conference abstracts [35,42].

Twelve of the eligible studies were intervention studies, comprising eight RCTs [28,30,32,34,36,37–39,42,43], two quasi–RCTs [33,44], one controlled clinical trial (CCT) [29], and one uncontrolled before and after study [41]. Two studies were cohort studies [31,35] and one was a case–control study (**Table 1**) [40]. Seven studies were undertaken in Sub–Saharan Africa (Kenya [31,42], Mali [44], Nigeria [30,40], Tanzania [37–39], and Zambia [41]), five in East Asia (China [33,36], Taiwan [28,29], and Thailand [32]), two in South Asia (Bangladesh and India) [35,43] and one in the Middle East (Iran) [34]. All the studies were published between 2008 and 2014. The study population comprised pregnant women in ten studies [28,29,30,32,33,34,35,37–39,40, 42], children in five studies [31,36,41,43,44], and village elders in one (**Table 1**) [31].

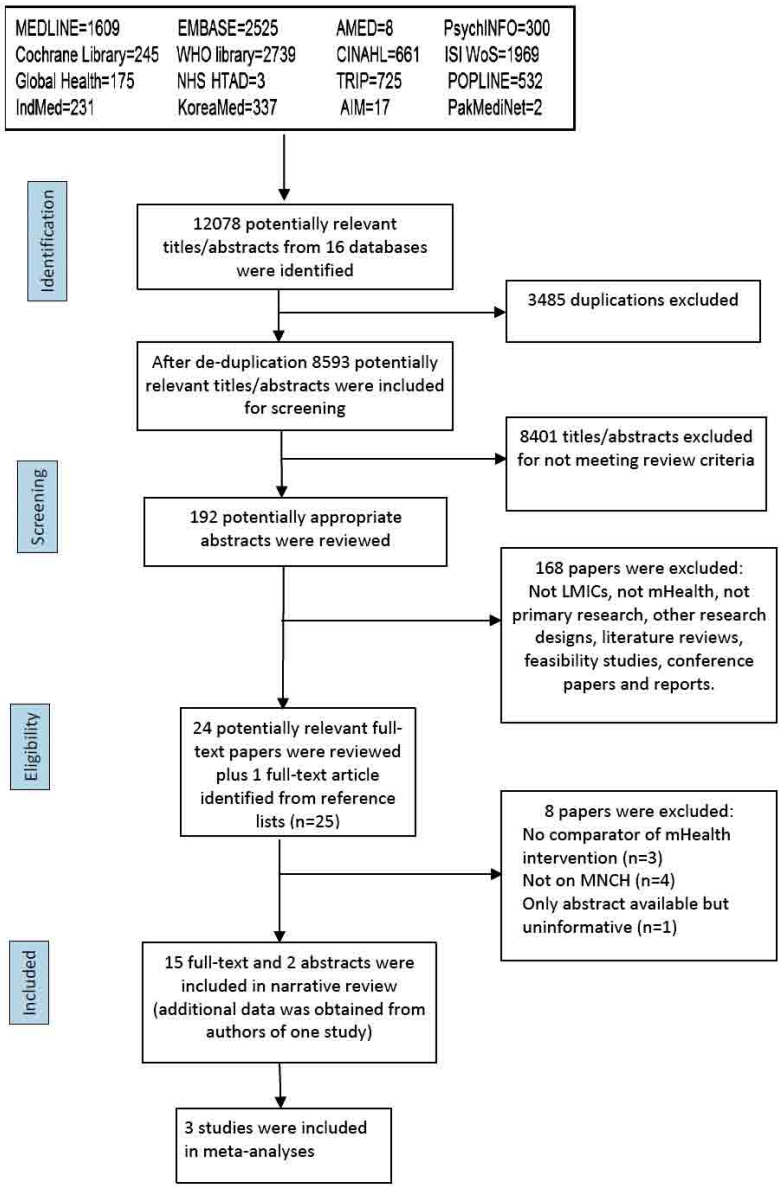


Figure 1. PRISMA flow diagram for database search of studies on mHealth interventions for maternal, newborn and child health in low– and middle–income countries, 1990–2014.

Assessment of risk of bias

Risk of bias grading for the different components of each study is shown in Tables s3 and s4 in the **Online Supplementary Document**. Only two of the intervention studies were graded as being at low risk of bias [36,42], seven as moderate [29,30,32,34,37–39] and four at high [28,33,43,44] risk of bias (see Table s3 in **Online Supplementary Document**). One cohort study was graded high risk of bias [31], while a case–control and a before–and–after study were graded moderate risk of bias (see Table s4 in **Online Supplementary Document**) [40,41]. Two of the studies included in our review were available only as conference abstracts [35,42]. Both sets of authors were contacted for further information and one replied, providing additional data that enabled us to better assess that study [42].

Mobile delivery media

The delivery modes used were mobile phones with SMS (n = 11) [28,32–34,36–39,41–43], SMS and voice messaging (n = 1) [30] and voice calls (n = 2) [35,40]. Two studies used mobile applications to collect data [31,44] and one study used MP3 players to deliver audio recordings [29].

Types of interventions

We classified the interventions according to our interpretation of their aims, based on the descriptions provided in the study reports, having first assessed existing taxonomies and found them to be not ideally suited to our purposes [20–22]. Studies were included in more than one category if the intervention was multi–faceted. The categories were health information delivery (n = 6) [30,32,33,34,37–39,43], reminders (n = 3) [34,36,37–39], communication platform (n = 2) [35,40], data collection platform (n = 2) [31,44], test result turnaround (n = 2) [28,41], peer/group support (n = 2) [30,42], and psychological intervention (n = 1) [29]. The results of this classification exercise are shown in **Figure 2**.

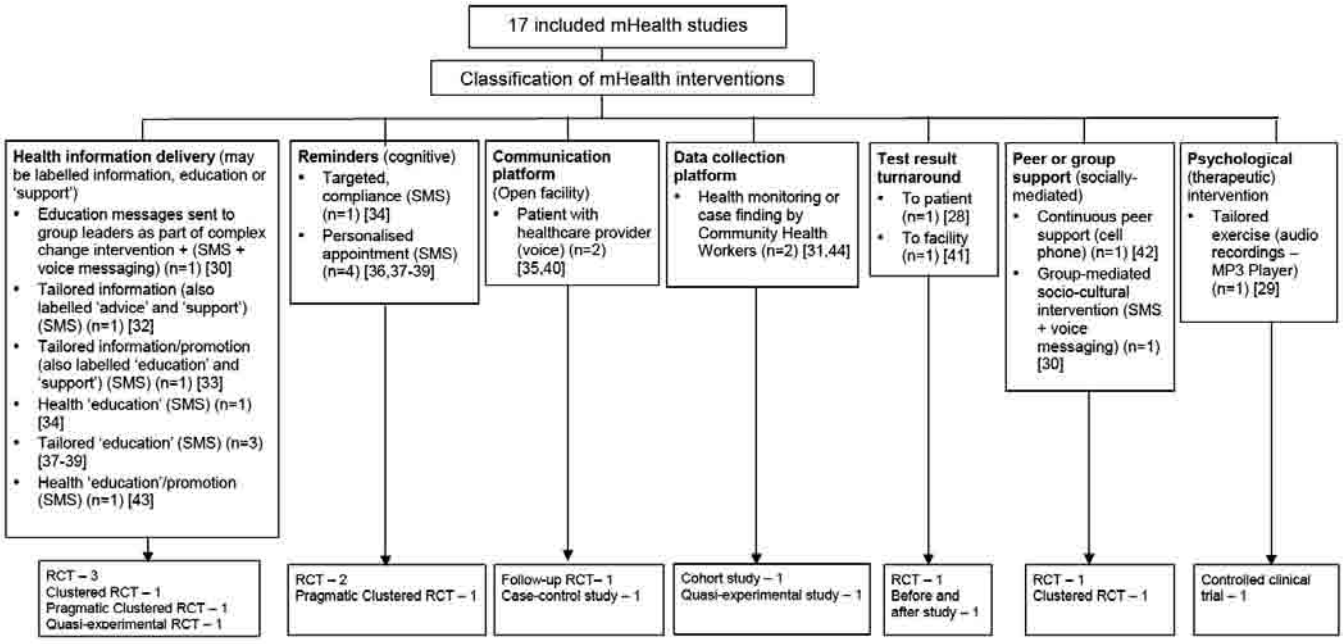


Figure 2. Classification of mHealth interventions of included studies. Categories are as interpreted by the reviewers, based on study descriptions. The authors may label studies somewhat differently. For example, the word 'support' may be used to describe informational messages, such as where it is theorized that these may confer psychological support in addition to knowledge support (eg, knowing that it is normal to experience morning sickness), although rarely do the authors elaborate on this. Studies are included in more than one category if the intervention is multi–faceted.

Types of outcomes examined

Eight studies examined indicators of maternal, newborn and child morbidity and mortality [28,29,31,32,34,37,40,44]. These covered maternal death [37], indicators of anaemia [34], duration of gestation at birth or preterm delivery [29,32], perinatal death and stillbirth [29,37], birth weight [30,31], Apgar score [28], hospitalization [29], route of delivery [29,31], infectious diseases [40,44], and oral health [43]. Other outcomes included indicators of infant feeding and breastfeeding [30,33,42], utilisation of antenatal, intrapartum, and postnatal care [31,35,37–40,44], quality of care [36,38], recording and collection of study data [31,40], indicators of self–efficacy [28,33], and compliance with recommended practices, such as micronutrient intake and uptake of immunization [34,36-38]. We did not find any study evaluating the cost–effectiveness of mHealth. The results are organised below according to the types of outcomes examined in each study.

Effects on maternal, newborn and child morbidity and mortality

A Taiwanese CCT compared pregnancy outcomes in women at risk of pre–term labour who had received daily 13–minute relaxation therapy sessions delivered via mp3 player, as compared with routine prenatal care (**Table 1**) [29]. Women in the intervention group had longer pregnancies, but there was no difference in the rate of pre–term birth, birth weight, perinatal mortality or Apgar score.

In a RCT from Thailand [32], the duration of gestation, birth weight, preterm delivery and caesarean section were comparable in pregnant women receiving SMS prenatal support via mobile phone to those who received routine prenatal care (**Table 1**). Similar results were seen in a pragmatic cluster RCT from Zanzibar, Tanzania [37–39], in which women receiving SMS prenatal support were comparable to those who received routine prenatal care, however, the risk of perinatal death decreased by half in the SMS group compared to the routine care group (odds ratio (OR) 0.50, 95% confidence interval (CI) 0.27–0.93) (**Table 1**). An Iranian RCT evaluated a 12–week programme of SMS reminders encouraging compliance with iron supplementation. While self–reported compliance was greater in the intervention group than in a control group not receiving the SMS reminders, there was no difference between the groups in objective measures of serum iron [34] (**Table 1**).

A Nigerian case–control study [40] compared rates of facility utilization and maternal morbidity in health care facilities where pregnant women had received mobiles as a communication platform. No measurable differences were observed between the two samples (**Table 1**).

A quasi–experimental study from Mali of children aged 0–72 months [44] did not reveal differences in the inci-

Table 1. Characteristics and results of studies investigating the effectiveness of mHealth interventions for maternal, newborn and child health in low– and middle– income countries during January 1990 – May 2014

Study and country	Study design and setting	Study population	Intervention/ Exposure	Outcomes	Results	Overall risk of bias grading	Classification of interventions
Cheng et al. (2008), Taiwan [28]	Randomised controlled trial (RCT), Hospital	Pregnant women at 14–18 weeks of gestation. Total N = 2782 Intervention group = 1422 Control Group = 1360	Report of results of Down Syndrome via SMS vs report at the time of routine clinic appointment	*Primary outcomes:* Anxiety levels of women as measured by Trait–and Stat–anxiety scores	*Negative results for Down Syndrome:* Trait–anxiety score ($P = 0.69$): SMS group mean 39.8 ± 11.2; Control group mean 38.4 ± 10.9 State–anxiety score Before screening ($P = 0.51$): SMS group mean 38.9 ± 9.9; Control group mean 37.8 ± 11.3 After screening ($P = 0.02$): SMS group mean 33.8 ± 7.9; Control group mean 39.1 + 10.1 *Positive results for Down Syndrome:* Trait–anxiety score ($P = 0.57$): SMS group mean 38.7 ± 8.8; Control group mean 40.1 ± 13.2 State–anxiety score Before screening ($P = 0.66$): SMS group mean 39.2 ± 11.4; Control group mean 39.9 ± 9.4 After screening ($P = 0.21$): SMS group mean 44.1 ± 13.4; Control group mean 42.9 ± 11.5	High	Test result turnaround
Chuang et al. (2012), Taiwan [29]	Controlled Clinical Trial, Hospital	Women diagnosed with preterm labour at 20–34 weeks of gestation Total N = 129. Intervention group = 68 Control group = 61	13–minute relaxation audio program via mp3 player vs no mp3 player (routine prenatal care)	*Primary outcomes:* Gestation at birth; new–born birth weight; Apgar score; perinatal mortality; admission to neonatal intensive care unit; number of days of prolongation of pregnancy	*Gestational weeks at birth ($P = 0.217$):* Mp3 player group mean 35.2 ± 4.4; Control group mean 34.2 ± 4.5 *Birth weight in grams ($P = 0.296$):* Mp3 player group mean 2389.2 ± 828 Control group mean 2266.6 ± 898 *Apgar score at 1 min ($P = 0.782$):* Mp3 player group mean 7.9 ± 2.0 Control group mean 7.8 ± 2.0 *Apgar score at 5 min ($P = 0.732$):* Mp3 player group mean 9.2 ± 1.9 Control group mean 9.0 ± 1.9 *Route of delivery ($P = 0.918$):* Normal: mp3 player group 52.9%; control group 54.2% Caesarean section: mp3 player group 47.1% control group 45.8% *Perinatal mortality ($P = 0.337$):* Mp3 player group 1.5%; control group 5.1%	Moderate	Psychological (therapeutic) intervention – Tailored exercises (audio recordings)
Flax et al. (2014), Nigeria [30]	Cluster RCT, General population	Pregnant women aged between 15–45 y. Total N = 461 Intervention group = 229 Control group = 232	Breastfeeding (BF) learning sessions and SMS and songs/dramas vs none of these (routine care)	*Primary outcomes:* –Exclusive BF to 1, 3, and 6 months –Initiation of BF within 1 h of delivery –Use of colostrum or breast milk within the first 3 d of life.	*Exclusive BF at 1 months:* Intervention group 73%; Control group 61%; OR 1.6 (95% CI 0.6–1.8) *Exclusive BF at 3 months:* Intervention group 71%; Control group 58%; OR 1.8 (95% CI 1.1–3.0) *Exclusive BF at 6 months:* Intervention group 64%; Control group 43%; OR 2.4 (95% CI 1.4–4.0) *Initiated BF within 1 h of delivery:* Intervention group 70%; Control group 48%; OR 2.6 (95% CI 1.6–4.1) *Gave only colostrum/breast milk during the first 3 d:* Intervention group 86%, Control group 71%; OR 2.6 (95% CI 1.4–5.0)	Moderate	Health Information delivery – Education messages sent to group leaders as part of complex change intervention (SMS+Voice messaging) – Group–mediated socio–cultural intervention (SMS±Voice Messaging)

Table 1. Continued

Study and country	Study design and setting	Study population	Intervention/ Exposure	Outcomes	Results	Overall risk of bias grading	Classification of interventions
Gisore et al. (2012), Kenya [31]	Cohort study General population	Village elders Total N = 474	Use of mobiles by village elders for pregnancy case finding and reporting birth weights	*Primary outcomes:* –% change in birth weights reported by mobile phones compared to previous national estimates –% of women enrolled after delivery	Recorded birth weights increased from 43 ± 5.7% to 97 ± 1.1% % of women enrolled after delivery decreased from 30.4% to 25%, $P < 0.0001$	High	Data collection (health monitoring or case finding by Community Health Workers)
Jareethum et al. (2008), Thailand [32]	RCT Hospital	Pregnant women at <28 weeks gestation Total N = 61 Intervention group = 32 Control group = 29	SMS via mobile phone for prenatal support vs no SMS (routine prenatal care)	*Primary outcome:* Mothers' level of satisfaction with antenatal care *Secondary outcomes:* –Mothers' confidence level at prenatal care –Mothers' anxiety level at prenatal care –Gestational weeks at delivery –Foetal birth weight –Route of delivery –Preterm delivery	*Mothers' level of satisfaction with prenatal care (P=<0.001):* SMS group mean 9.3 ± 0.7; Control group mean 8.0 ± 1.1 *Mothers' confidence level at prenatal care (P=0.001):* SMS group mean 8.9 ± 0.9; Control group mean 7.8 ± 1.5 *Mothers' anxiety level at prenatal care (P=0.002):* SMS group mean 2.8 ± 2.1; Control group mean 4.9 ± 2.9 *Gestational weeks at delivery (P=0.340):* SMS group mean 38.7 ± 1.1; Control group mean 38.6 ± 1.1 *Foetal birth weight in grams (P=0.350):* SMS group mean 3051 ± 636; Control group mean 3188 ± 456 *Preterm delivery (P=0.220):* SMS group 0%; Control group 6.9% *Route of delivery (P=1.00):* Normal vaginal delivery: SMS group 81.3%; Control group82.8% Caesarean section: SMS group18.7%; Control group17.2%	Moderate	Health Information Delivery – Tailored information (also labelled 'advice' and 'support') (SMS)
Jiang et al. (2014), China [33]	Quasi–RCT Community Health Centres	Pregnant women at <13 weeks gestation Total N = 582 Intervention group = 281 Control group = 301	Text via SMS vs no SMS (routine prenatal care)	*Primary outcome:* Duration of exclusive BF *Secondary outcomes:* –Rate of exclusive BF at 6 months –Duration of any BF –Timing of intro. solid foods –Rate of BF at 12 months –Rates of other infant feeding behaviours	*Exclusive BF at 4 months:* SMS group 46.4%; Control group 39.9%; OR 1.4 (95% CI 1.0–2.0) *Exclusive BF at 6 months:* SMS group 15.1%; Control group 6.3%; OR 2.7 (95% CI 1.5–4.9) *BF at 12 months:* SMS group 20.2%; Control group 19.2%; OR 1.0 (95% CI 0.7–1.6) *Introduction of solid foods before 4 months:* SMS group 1.5%; Control group 3.8%; OR 0.3 (95% CI 0.1–0.9) *Introduction of solid foods before 6 months:* SMS group 67.5%; Control group 61.3%; OR 1.3 (95% CI 0.9–1.8) *Drinking from a cup at 12 months:* SMS group 53.6%; Control group 46.5%; OR 1.3 (95% CI 0.9–2.0) *Receiving food as a reward:* SMS group 45.5%; Control group 33.6%; OR 1.5 (95% CI 1.0–2.3) *Taking a bottle to bed:* SMS group 51.9%; Control group 49.8%; OR 1.1 (95% CI 0.7–1.6)	High	Health Information Delivery – Tailored information/ promotion (also labelled 'education' and 'support') (SMS)

Table 1. Continued

Study and country	Study design and setting	Study population	Intervention/ Exposure	Outcomes	Results	Overall risk of bias grading	Classification of interventions
Khorshid et al. (2014), Iran [34]	RCT Public Health Centres	Pregnant women at gestational 14–16 weeks Total N = 116 Intervention group = 58 Control group = 58	A 12–week SMS reminders in addition to usual care vs no SMS reminders (only usual care) on compliance with intake of iron supplements	*Primary outcome:* Compliance with intake of iron supplements *Secondary outcomes:* Measures of blood indices for anaemia (haemoglobin, haematocrit, ferritin)	*Compliance with intake of iron supplements (P = 0.003):* High compliance: SMS group 94%; Control group 66% Moderate compliance: SMS group 4%; Control group 18% Low compliance: SMS group 2%; Control group 16% *Measures of blood indices for anaemia:* Haemoglobin in g/dL (*P* = 0.960): SMS group mean 11.2 ± 0.5; control group mean 11.2 ± 0.9 Haematocrit, % (*P* = 0.670): SMS group mean 33.9 ± 1.7; control group mean 34.0 ± 2.6 Ferritin in ng/dL (*P* = 0.630): SMS group mean 24.4 ± 35.0; control group mean 22.5 ± 19.7	Moderate	Health Information Delivery – Health 'education' (SMS)
Labrique et al. (2011), Bangladesh [35]*	Follow–up analysis of RCT General population	Pregnant women interviewed at 1 month postpartum to collect information on complications of labour and delivery Total *N*>100 000)	Use of mobile phones to report obstetric emergencies	*Primary outcomes:* –Reported use of mobile phones during intrapartum	55.2% of women reported using a mobile phone for obstetric emergencies. Of these: 57.0% to receive medical advice 71.7% to call a health care provider 32.6% to arrange for transportation 20.9% to ask for financial support.	N/A	Communication Platform (one way or two way interpersonal communication) – Patient with Health Care Providers (Voice)
Lin et al. (2012), China [36]	RCT Hospital	Parents of children with diagnosis of cataract aged <18 years Total N = 258. Intervention group = 135 Control group = 123	Text messaging via SMS vs standard follow–up appointments	*Primary outcome:* Rate of attendance at scheduled study appointments *Secondary outcomes:* –Additional procedures (surgeries, laser treatments for posterior capsular opacification, or changes in eyeglass prescription) –Occurrence of secondary ocular hypertension	*Attendance rates (P = 0.003):* SMS group 91.3%; Control group 62.0%; RR: 1.47 (95% CI 1.16–1.78) *Secondary outcomes:* Surgeries (*P* = 0.03): SMS group 43.0% ; Control group 27.6%) ; RR 1.55 (95% CI 1.10–2.20) Laser for capsular opacification (*P* = 0.008): SMS group 46.0%; Control group 18.7%; RR 2.46 (95% CI 1.63–3.71) Prescription of new glasses (*P* = <0.001): SMS group 71.1%; Control group 52.8%; RR 1.35 (95% CI 1.10–1.64) Treatment for ocular hypertension (*P* = 0.04): SMS group 23.0%; Control group 9.8%; RR 2.35 (95% CI 1.27–4.38)	Low	Reminders (Cognitive) – Personalised, appointment (SMS)

Table 1. Continued

Study and country	Study design and setting	Study population	Intervention/ Exposure	Outcomes	Results	Overall risk of bias grading	Classification of interventions
Lund et al. 2012, 2014a, 2014b, Zanzibar, Tanzania [37–39]	Pragmatic Cluster–RCT General population	Pregnant women at first prenatal care attendance Total N = 2637 Intervention group = 1351 Control group = 1286	Mobile phone vouchers and SMS vs no mobile phones (routine care)	*Primary outcomes:* –Skilled delivery attendance –Number of women receiving four or more antenatal care visits *Secondary outcomes:* –Home delivery assisted by skilled birth attendants –Quality of care in terms of content and timing of antenatal care –Stillbirth –Perinatal death –Death of a child within 42 d of life	*Skilled delivery attendance:* SMS group 60%; Control group 47% *Four or more antenatal visits:* SMS group 44%; Control group 31%; a OR 2.39 (95% CI 1.03–5.55) *Secondary outcomes:* –Tetanus toxoid vaccination at first antenatal care visit: SMS group 96%; Control group 94%; aOR 1.58 (95% CI 0.41–6.01) –Tetanus toxoid vaccination at least 4 weeks after first antenatal care visit: SMS group 72%; Control group 56%; aOR 1.62 (95% CI 0.81–3.26) –Intermittent preventive treatment in pregnancy at first prenatal visit: SMS group 91%; Control group 86%; aOR 1.10 (95% CI 0.35–3.43) –Intermittent preventive treatment in pregnancy at least 4 weeks after first prenatal visit: SMS group 65%; Control group 52%; aOR 1.97 (95% CI 0.98–39.4) –Gestational age 36 or more at last antenatal care visit: SMS group 28%; Control group 20%; aOR 1.48 (95% CI 0.89–2.45) –Antepartum referral: SMS group 10%; Control group 5%; aOR 1.66 (95% CI 0.68–4.06) –Stillbirth: SMS group 17 per 1000 births; Control group 26 per 1000 births; aOR 0.65 (95% CI 0.34–1.24) –Perinatal mortality: SMS group 19 per 1000 births; Control group 36 per 1000 births; aOR 0.50 (95% CI 0.27–0.93) –Death of child <42 d after birth: SMS group 14 per 1000 births; Control group 15 per 1000 births; aOR 0.79 (95% CI 0.36–1.74)	Moderate	Health Information Delivery – Tailored education (SMS) – Reminders (Cognitive) – Personalised, appointment (SMS)
Oyeyemi and Wynn (2014), Nigeria [40]	Case–control study General Population	Pregnant women Cases = 1429 Controls = 1801	Giving mobile phones to pregnant women to increase primary health facility utilisation (cases) vs no mobile phones (controls)	*Primary outcome:* Facility utilisation rate *Secondary outcome:* Frequency of occurrence of 5 major causes of maternal deaths (severe bleeding, hypertensive disorder of pregnancy with fits, infection, obstructed labour, unsafe abortion)	*Facility utilisation:* Cases 43.4%; Controls 36.6; OR 1.32 (95% CI 1.15–1.53 *Number of illness cases:* Cases 1.6%; Controls 1.6%; OR 1.00 (95% CI 0.58–1.74)	Moderate	Communication Platform – One– or two–way interpersonal communication (Voice)

Table 1. Continued

Study and country	Study design and setting	Study population	Intervention/ Exposure	Outcomes	Results	Overall risk of bias grading	Classification of interventions
Seidenberg et al. (2012), Zambia [41]	Before and after study General population	All infants who came for antenatal care Before program = 1009 After program = 406	Notification of blood results of infant diagnosis of HIV infection through SMS vs postal notification	*Primary outcomes:* –Mean turnaround time (time from sample collection to delivery of test result to either the relevant point–of–care health facility or a caregiver of the tested infant) –Result error rate (per cent discordance between the results recorded on paper and the corresponding results sent by SMS)	*Turnaround (days) at relevant health facility:* Before program: mean 44.2 ± 28.0 After program: mean 26.7 ± 31.8 Difference in mean days: –17.5 (95% CI –14.1 to –20.9) *Turnaround (days) to a caregiver:* Before program: mean 68.8 ± 38.8 After program: 35.0 ± 31.2 Difference in mean: –33.8 (95% CI –28.7 to –38.9) *Per cent discordance:* Number of samples agreed by paper and SMS = 336 Number of discrepancies = 2 Error rate 0.5%	Moderate	Test result turnaround – To facility
Sellen et al. (2013), Kenya, [42]*	RCT Hospital	Pregnant women from late pregnancy –3rd trimester (32–36 weeks) to 3 months postpartum n = 530 CPS = 223 PSG = 267 SOC = 263	Pregnant women were randomised to 3 groups A. Continuous cell phone based peer support (CPS) B. Monthly peer support group (PSG) C. Standard of care (SOC)	*Primary outcome:* Exclusive BF at 3 months	*BF initiated within 1 h:* CPS 73.0%, PSG 70.2%, SOC 67.2%, *P* = 0.519; OR for CPS vs SOC 1.32 (95% CI 0.82–2.12) *Onset of lactation >3days:* CPS 10.3%, PSG 8.9%, SOC 11.2%, *P* = 0.764; OR for CPS vs SOC 1.09 (95% CI 0.55–2.19) *Exclusive BF at 3 months:* CPS vs SOC: 90.9%vs 78.2% (Chi–square 9.8201, *P* = 0.0017) CPS vs PSG: 90.9% vs 82.8% (*P* = 0.032) OR for CPS vs SOC 2.77 (95% CI 1.44–5.32)	Low	Peer or group support (socially–mediated) – Continuous peer support (Cell phone)
Sharma et al. (2011), India [43]	RCT	Preschool children and their mothers Total N = 143 Intervention group = 71 Control group = 72	Oral health education via SMS vs pamphlet	*Primary outcomes:* Mothers' knowledge, attitude, and practice of child's oral health; Visible Plaque Index (VPI)	*Mean KAP scores for knowledge:* Pre–intervention: SMS group 8.2 ± 1.2; pamphlet group 7.8 ± 1.5 Post–intervention: SMS group 9.4 ± 0.8; pamphlet group 8.8 ± 1.1 Differences between groups: –0.43 (95% CI –0.33 to 0.51) *Mean KAP scores for attitude:* Pre–intervention: SMS group 8.8 ± 1.3; pamphlet group 7.8 ± 1.8 Post–intervention: SMS group 9.4 ± 0.7; pamphlet group 8.8 ± 1.3 Differences between groups: –0.37 (95% CI –0.61 to –0.13) *Mean KAP scores for practices:* Pre–intervention: SMS group 11.3 ± 1.8; pamphlet group 11.1 ± 1.8 Post–intervention: SMS group 12.1 ± 1.3; pamphlet group 11.5 ± 1.7 Differences between groups: –0.44 (95% CI –0.77 to –0.12) *Mean KAP scores for VPI:* Pre–intervention: SMS group 45.0 ± 21.2; pamphlet group 45.4 ± 20.5 Post–intervention: SMS group 33.5 ± 17.0; pamphlet group 35.6 ± 16.2 Differences between groups: 1.81 (95% CI –1.39 to 5.01)	High	Health Information Delivery – Health 'education'/ promotion (SMS)

Table 1. Continued

Study and country	Study design and setting	Study population	Intervention/ Exposure	Outcomes	Results	Overall risk of bias grading	Classification of interventions
Simonyan et al. (2013), Mali [44]	Quasi–experimental study, General population	0–72 months old with no diagnosed chronic diseases Total N = 188 Intervention group = 99 Control group = 89	Diagnosis, collection and transfer of health care data using mobile phone via a JAVA applet to a central server vs usual care	*Primary outcome:* Healthcare utilisation *Secondary outcomes:* Child morbidity indicated by number of episodes of cold, cough, diarrhoea, fever, infection, pain, teething, vomiting, wounds, and others	*Healthcare utilisation:* Mobile phone group 93.4%; control group 31.5%; OR 2.2 (95% CI 1.3–3.9) Total number of disease episodes: Mobile phone group 236; Control group 168 Episodes for specific disease are given in the paper. These were not statistically significantly different from the two groups.	High	Data collection (health monitoring or case finding by Community Health Workers)

ASR – adjusted standardized residuals, CPS – continuous cell phone based peer support, DS – Down Syndrome, EBF – effective breastfeeding, ICC – interclass correlation coefficient, IYCF – infant and young child feeding, NICU – Neonatal Intensive Care Unit, OR – odds ratio, aOR – adjusted odds ratio, PSG – peer support group, RCT – randomized controlled trial, RR – relative risk, SD – standard deviation, SES – socio–economic status, SMS – short message service, SOC – standard of care

* Only abstract available.

dence of childhood diseases between those whose health care data and diagnosis were recorded and transferred using mobile phone compared to children whose data were not recorded using mobile phone (**Table 1**).

Effects on infant feeding

Flax et al. [30], Jiang et al. [33], and Sellen et al. [42], compared the effect of SMS/cell phone vs no SMS (routine prenatal care) on breastfeeding in Nigeria, China and Kenya, respectively. The results of each trial showed that the rate of exclusive breastfeeding (EBF) for three or four months was higher in the SMS/cell phone group than in the non–SMS/cell phone group (**Table 1**). We undertook meta–analyses of the effect of SMS/cell phone vs routine prenatal care on the initiation of breastfeeding within one hour after birth [30,42], giving colostrum or breast milk within three days after birth [30,42], and EBF at three/four months [30,33,42], and at six months [30,33]. The pooled estimates showed that the rates of initiating breastfeeding within one hour after birth (OR 2.01, 95% CI 1.27–2.75, $I^2 = 80.9\%$, **Figure 3**) were higher in the groups given a SMS/cell phone prenatal intervention than in groups not given the SMS/cell phone intervention. The evidence for giving colostrum or breast milk within three days after birth was not strong (OR 1.90, 95% CI 0.86–2.94, $I^2 = 77.0\%$, **Figure 4**). The rates of EBF for three/four months (OR 1.88, 95% CI 1.26–2.50, $I^2 = 52.8\%$, **Figure 5**) and EBF for six months (OR 2.58, 95% CI 1.44–3.71, $I^2 = 0.0\%$, **Figure 6**) were also higher in the groups given a SMS/cell phone prenatal intervention than in groups not given the SMS/cell phone intervention.

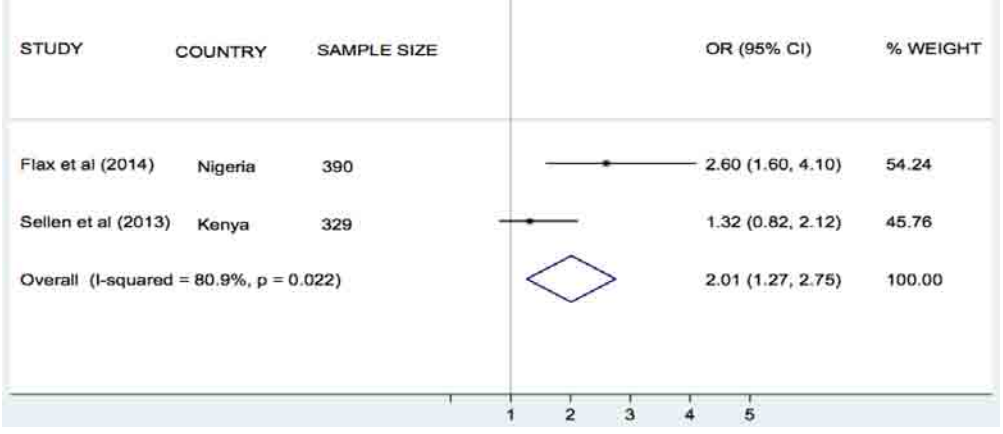


Figure 3. Meta–analysis of the effect of SMS/cell phone intervention vs routine prenatal care on initiation of breastfeeding within one hour after birth based on two RCT undertaken in Nigeria and Kenya: OR represents the odds ratio of effect.

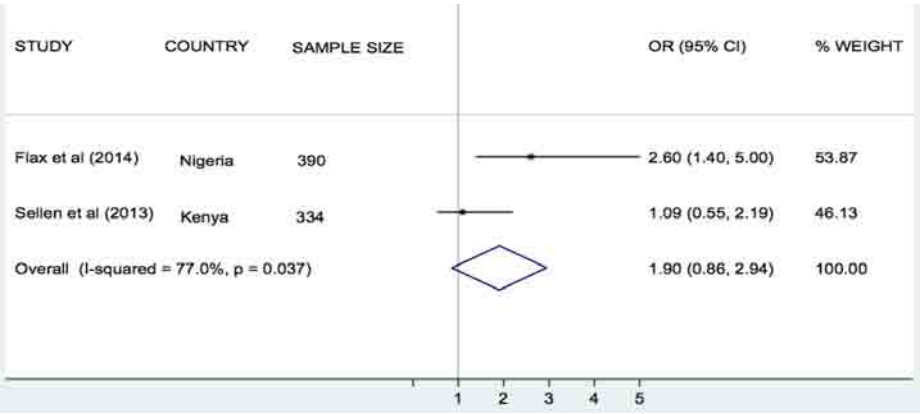


Figure 4. Meta–analysis of the effect of SMS/cell phone intervention vs routine prenatal care on onset of lactation within three days after birth based on two RCT undertaken in Nigeria and Kenya: OR represents the odds ratio of effect.

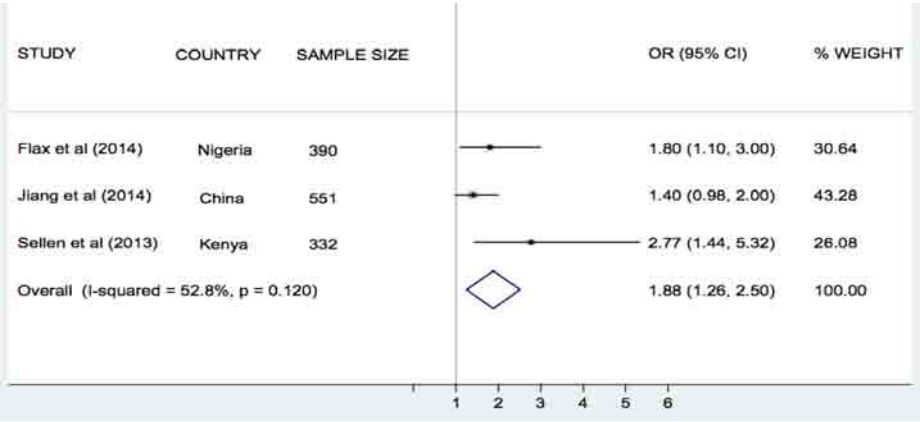


Figure 5. Meta–analysis of the effect of SMS/cell phone intervention vs routine prenatal care on exclusive breastfeeding for three or four months based on three RCT undertaken in Nigeria, China, and Kenya: OR represents the odds ratio of effect.

Effect on *health care utilisation and quality of care*

In a follow–up study of a RCT in Bangladesh, Labrique et al. assessed the level of use of mobile phones by pregnant women in reporting obstetric emergencies [35]. 55% of pregnant women reported having used the mobile phones to obtain medical advice, call a health care provider, arrange for transportation or ask for financial support.

A Chinese RCT evaluated the effects of SMS–based appointment reminders for parents with children 0–18–years diagnosed with cataract and attending the paediatric clinic of a specialist eye hospital [36]. Attendance at follow–up clinics was higher in the group receiving SMS reminders than in those with standard appointments (91% vs 62%). This was associated with more surgeries, laser treatment for capsular opacification, prescription of new glasses and treatment for ocular hypertension in the intervention group, compared with those not receiving these reminders (**Table 1**). No subgroup analysis was reported for the under 5s.

Among pregnant Tanzanian (Zanzibar) women [37–39], those given mobile phones in order to receive SMS information about antenatal care were more likely to attend four or more antenatal care clinics (OR 2.39, 95% CI 1.03–5.55) and have skilled attendance at delivery (OR 5.73, 95% CI 1.51–21.81) than those who received routine prenatal care. No strong evidence of differences regarding tetanus vaccination, intermittent preventive treatment during pregnancy and antepartum referral were found (**Table 1**).

Healthcare utilisation was higher in pregnant Nigerian women from health facilities receiving mobile phones compared to women from health facilities without mobile phones (OR 1.32, 95% CI 1.15–1.53) [39].

Finally, Simonyan et al. found that health care utilisation was higher in Malian children whose health care data and diagnosis were collected and transferred using mobile phones compared with children whose data were collected and transferred using standard methods (OR 2.20, 95% CI 1.3–3.9) [44].

Ongoing studies

Seven ongoing studies assessing the influence of mHealth interventions on maternal and child health outcomes in LMIC were identified in the course of the review. Three of these are being undertaken in Kenya, one each in Cameroon, Ethiopia, India and Mozambique. Six studies involve

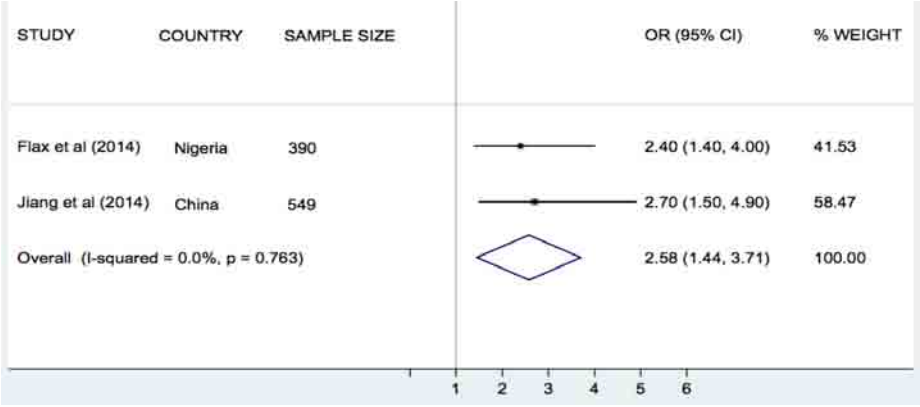


Figure 6. Meta–analysis of the effect of SMS/cell phone intervention vs routine prenatal care on exclusive breastfeeding for six months based on three RCT undertaken in Nigeria and China: OR represents the odds ratio of effect.

pregnant women and one involves children as participants. The mHealth interventions in all studies involve SMS or voice calls via mobile phones. (see Table s5 in **Online Supplementary Document**)

DISCUSSION

The current evidence base contains many studies describing the use of mHealth for supporting MNCH in LMIC but comparatively few have robustly evaluated the impacts of these interventions on health outcomes in these groups.

The majority of included studies took place in Sub–Saharan Africa and East Asia, while a few were undertaken in South Asia and the Middle East. Most studies were at moderate risk of bias. Although heterogeneity between studies precluded the calculation of a pooled estimate, mHealth interventions did not improve indicators of maternal, newborn, and child morbidity and mortality, except in one study from Tanzania that reported a decreased risk of perinatal death with use of SMS for prenatal support during pregnancy. However, a meta–analyses of three studies judged to be sufficiently homogenous showed that delivering prenatal breastfeeding interventions using SMS/cell phone (vs routine prenatal care) improved rates of initiation of BF within one hour after birth and increased the likelihood of EBF for up to six months, although there was no strong evidence regarding the giving of colostrum or breast milk within three days after birth.

mHealth technologies are increasingly being used to enhance health care utilisation, improve the quality of pre– and post–pregnancy care, and as a means of collecting pregnancy and child health data. Some studies showed that mHealth interventions, particularly those delivered using SMS, were associated with increased utilisation of health care, including uptake of recommended prenatal and postnatal care consultation, skilled birth attendance, and vaccination.

Most authors did not fully explain the basis of their intervention, in terms of its components or the mechanisms through which it would deliver the intended outcomes, and overall the studies lacked a common taxonomy for describing the type and purpose of the intervention. For example, the term 'support' was sometimes identified with health information delivery whereas elsewhere with a more psychosocial intervention. To aid interpretation and comparison we developed a framework for classifying the interventions according to their purpose, as previously described (see **Figure 2**). Based on our interpretation, the most common use of mHealth was for *health information delivery,* such as nutritional advice [30, 32, 33,34,37–39,43]. This was followed by *reminders*, chiefly for clinic attendance [34,36,37–39]. The other observed categories were mHealth as a *communication platform*, mainly to access support from care providers [35,40]; as a *data collection platform*, to enable birth registration or reporting of health indicators [31,44]; for accelerating *test result turnaround* times through by–passing the need for physical transportation [28,41]; part of *peer*–support [30,42]; and as a means by which to deliver *psychological (therapeutic)* interventions [29].

This systematic review draws on a comprehensive, inclusive and highly sensitive literature search strategy, analyses both health and health care utilization indicators; includes all legitimate mHealth technologies, covers the full spectrum of maternal and infant health and was not restricted by language. It has successfully captured the body of quantitative comparative studies on mHealth for MNCH through analysing a very large initial corpus of studies, and not simply those specified by the World Bank list of LMIC which, our pilot searches revealed, would have excluded key trials that we were aware of.

Comparable reviews have lacked such a robust search strategy [18], or have focused on the operational functions of mobile technologies rather than their outcomes [19,46]. In addition to those described in our introduction, new reviews arising after the publication of our protocol have similar limitations: Aranda–Jan and colleagues reviewed a range of mHealth studies carried out in Africa using only two data-

bases [47], while Hall et al.'s review assessing 'what interventions work' for a range of conditions, was limited to two databases and grey literature [48]. As already noted, although Free et al.'s review covered a broad range of mHealth interventions, the majority of the trials revealed were from high income countries [20,21], whilst a systematic review on mHealth for LMIC, mentioned in Philbrick's broader scoping review, is not available for comparison [17].

As with many systematic reviews in the field of eHealth, this analysis is limited by the difficulty of interpreting and synthesizing complex intervention studies and the variable description of interventions across studies. Although Labrique et al. developed a taxonomy for categorising different types of eHealth interventions [22], which we considered at the protocol stage, it did not fit our specific requirement to describe the interventions in terms of their purposes, for which the framework in Figure 2 was developed. Further work is needed to refine and test this with a larger body of interventions and to establish how best to integrate it with the various other published frameworks that exist. Due to the heterogeneity of the interventions and study outcomes we were unable to undertake meta–analyses, except in the case of the studies on infant feeding interventions, although this should be interpreted with caution due to the small number of studies analysed.

Our inclusion of studies from Taiwan is debatable, given its relatively high GDP but official status as part of China, which, although classified as 'upper middle' since 2012, is still a developing country. This, and our need to drop country restrictions from the search strategy due to labeling effects (eg, Zanzibar vs Tanzania), indicates taxonomic and socio–political challenges for systematic reviews of global research that warrant further methodological study.

Overall, the quality of studies included in the review was moderate, highlighting the importance of improving the methodological rigor of future research. For randomised trials, there is need for allocation concealment and adequate blinding of outcomes, while the quality of observational studies will be improved through prospective–designs and adjustment for confounding variables.

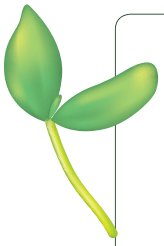
Departure from protocol

World Bank Country Classification [25] was used instead of United Nations Human Development Index, due to our focus on income level rather than other aspects of development. The outcomes remain unchanged.

CONCLUSIONS

There is a growing body of research indicating the potential of mHealth interventions for improving MNCH in LMIC, but overall the available evidence is weak and the results, in most cases, are too inconsistent to enable robust conclusions to be drawn about impacts on patient health outcomes. However supportive evidence exists with respect to the use of SMS/cell phones for improving infant feeding. Further research, using rigorous methodologies, is needed to better establish the effectiveness of mHealth interventions in MNCH initiatives in LMIC. In particular, trials with quantifiable economic, clinical and long–term patient–centred health outcomes are warranted. A number of in–progress trials are set to supplement this literature, while new research investments hold great promise for the development and evaluation of mHealth innovations for MNCH and other health priorities [49]. As low–cost smartphones begin to penetrate in these regions, a new generation of mobile Apps is now emerging, which will also require evidence–based methods to establish their safety, efficacy and societal impacts [15,50]. Innovative methods of integrating real–time evaluation into these deployments will also be essential if the potential evidence to be gained from them is to be effectively captured.

Our experience of engaging with this literature during the review also supports the common assertion that mHealth research projects are typically under–theorised, poorly specified and vaguely described. This creates challenges for effective evidence synthesis, risks unintended consequences that cannot be explained, makes replication and scaling difficult and hinders the effective translation of research to practice. We recommend that mHealth researchers, sponsors, and publishers prioritise the transparent reporting of interventions in terms of their aims, contexts, modes of delivery and presumed mechanisms of impact. Although anecdotal evidence of the benefits of mHealth for MNCH in LMIC is compelling, without this level of specification it will be difficult to develop robust evidence–based recommendations for policymakers and planners wishing to make informed choices about mHealth investments in these regions.



Acknowledgements: We thank medical librarians Mrs Marshall Dozier the University of Edinburgh and Mr Tomas Allen at the World Health Organization, for valuable and useful feedback on our search strategy, and Dr Joan Dzenowagis for commissioning the research.

Ethical approval: Not required.

Funding: World Health Organization, eHealth Unit.

Authorship declaration: CP conceived the study; UBN and CP prepared the initial protocol draft. UBN, LG and CP were involved in study design, protocol and manuscript development. UBN and CP designed the search strategy, which was refined by all authors (UBN, SHL, BIN, MM, LG and CP). All authors screened papers. SHL, UBN, BIN and MM extracted data, critically appraised papers. UBN, SHL, BIN and CP wrote the first draft of manuscript, BIN undertook the meta–analysis. SHL, CP, and BIN synthesised and finalised the manuscript. All authors read and approved the final version of the protocol manuscript.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). The authors declare no competing interests.

REFERENCES

1 United Nations. The Millennium Development Goals Report 2015. New York: UN DESA, 2015.

2 Liu L, Johnson HL, Cousens S, Perin J, Scott S, Lawn JE, et al; The Child Health Epidemiology Reference Group of WHO and UNICEF. Global regional and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. Lancet. 2012;379:2151-61. Medline:22579125 doi:10.1016/S0140-6736(12)60560-1

3 WHO. UNICEF, UN Population Fund and the World Bank. Trends in maternal mortality: 1990 to 2010. WHO, UNICEF, UNFPA and the World Bank estimates. Geneva: World Health Organisation, 2012.

4 Wulf SK, Johns N, Lozano R. Non–fatal burden of maternal conditions: country–level results from the GBD 2010 Study. Lancet. 2013;381 Suppl 2:S149. doi:10.1016/S0140-6736(13)61403-8

5 Mwaniki MK, Atieno M, Lawn JE, Newton CR. Long–term neurodevelopmental outcomes after intrauterine and neonatal insults: a systematic review. Lancet. 2012;379:445-52. Medline:22244654 doi:10.1016/S0140-6736(11)61577-8

6 Say L, Raine R. A systematic review of inequalities in the use of maternal health care in developing countries: examining the scale of the problem and the importance of context. Bull World Health Organ. 2007;85:812-9. Medline:18038064 doi:10.2471/BLT.06.035659

7 UNICEF. State of the world's children 2008: Maternal and newborn health. New York: UNICEF, 2008. Available: http://www.unicef.org/publications/index_47127.html#. Accessed: 3 January 2014.

8 Children WHO. reducing mortality. Fact sheet 178. 2012. Available: http://www.who.int/mediacentre/factsheets/fs178/en/index.html. Accessed: 3 January 2014.

9 Ronsmans C, Graham WJ. Maternal mortality: who, when, where, and why. Lancet. 2006;368:1189-1200. Medline:17011946 doi:10.1016/S0140-6736(06)69380-X

10 Van Malderen C, Van Oyen H, Speybroeck N. Contributing determinants of overall and wealth–related inequality in under–5 mortality in 13 African countries. J Epidemiol Community Health. 2013;67:667-76. Medline:23704052 doi:10.1136/jech-2012-202195

11 Kerber KJ, de Graft–Johnson JE, Butta ZA, Okong P, Starrs A, Lawn JE. Continuum of care for maternal, newborn, and child health: from slogans to service delivery. Lancet. 2007;370:1358-69. Medline:17933651 doi:10.1016/S0140-6736(07)61578-5

12 Istepanian RSH, Laxminarayan S, Pattichis CS, editors. M–Health: Emerging Mobile Health Systems. New York: Springer, 2005.

13 Tomlinson M, Rotheram–Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: Where is the evidence? PLoS Med. 2013;10:e1001382. Medline:23424286 doi:10.1371/journal.pmed.1001382

14 Consulting VW. mHealth for Development: The Opportunity of Mobile Technology for Healthcare in the Developing World. Washington, D.C. and Berkshire, UK: UN Foundation–Vodafone Foundation Partnership, 2009. Available: http://www.globalproblems-globalsolutions-files.org/unf_website/assets/publications/technology/mhealth/mHealth_for_Development_full.pdf. Accessed: 24 January 2014.

15 Rotheram–Borus MJ, Tomlinson M, Swendeman D, Lee A, Jones E. Standardized Functions for Smartphone Applications: Examples from Maternal and Child Health. Int J Telemed Appl. 2012;2012:973237. Medline:23304136

16 Wallace AS, Ryman TK, Dietz V. Experiences Integrating Delivery of Maternal and Child Health Services with Childhood Immunization Programs: Systematic Review Update. J Infect Dis. 2012;205 Suppl 1:S6-19. Medline:22315388 doi:10.1093/infdis/jir778

REFERENCES

17 Philbrick WC. mHealth and MNCH: State of the Evidence. Trends, Gaps, Stakeholder Needs, and Opportunities For Future Research on the Use of Mobile Technology to Improve Maternal, Newborn, and Child Health. mHealth Alliance, UN Foundation Jan 2013. Available: http://www.mhealthworkinggroup.org/sites/mhealthwg.org/files/philbrick_mnch_final.pdf. Accessed: 3 January 2014.

18 Noordam AC, Kuepper BM, Stekelenburg J, Milen A. Improvement of maternal health services through the use of mobile phones. Trop Med Int Health. 2011;16:622-6. Medline:21342374 doi:10.1111/j.1365-3156.2011.02747.x

19 Tamrat T, Kachnowski S. Special Delivery: An Analysis of mHealth in Maternal and Newborn Health Programs and Their Outcomes Around the World. Matern Child Health J. 2012;16:1092-101. Medline:21688111 doi:10.1007/s10995-011-0836-3

20 Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of Mobile–Health Technology–Based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med. 2013;10:e1001362. Medline:23349621 doi:10.1371/journal.pmed.1001362

21 Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. The effectiveness of Mobile–Health Technologies to improve health care service delivery processes: a systematic review and meta–analysis. PLoS Med. 2013;10:e1001363. Medline:23458994 doi:10.1371/journal.pmed.1001363

22 Labrique AB, Vasudevan L, Kochi E, Fabricant R, Mehl G. mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. Glob Health: Sci Pract. 2013;1:160-71. Medline:25276529

23 Nurmatov UB, Lee SH, Nwaru BI, Mukherjee M, Grant L, Pagliari C. The effectiveness of mHealth interventions for maternal, newborn and child health in low– and middle–income countries: protocol for a systematic review and meta–analysis. J Glob Health. 2014;4:010407. Medline:24976966 doi:10.7189/jogh.04.010407

24 Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta–analysis: the PRISMA statement. PLoS Med. 2009;6:e1000097. Medline:19621072 doi:10.1371/journal.pmed.1000097

25 The World Bank. Updated income classifications. 2014. Available: http://data.worldbank.org/news/2015-country-classifications. Accessed: 25 January 2014.

26 Higgins JPT. Green S eds. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration 2011. Available: http://www.cochrane.org/training/cochrane-handbook. Accessed: 29 January 2014.

27 Effective Public Health Practice Project (EPHPP). Quality assessment tool for quantitative studies. Available: www.ephpp.ca. Accessed 20 December 2013.

28 Cheng PJ, Wu TL, Shaw SW, Chueh HY, Lin CT, Hsu JJ, et al. Anxiety levels in women undergoing prenatal maternal serum screening for Down syndrome: the effect of a fast reporting system by mobile phone short–message service. Prenat Diagn. 2008;28:417-21. Medline:18395869 doi:10.1002/pd.1988

29 Chuang LL, Lin LC, Cheng PJ, Chen CH, Wu SC, Chang CL. The effectiveness of a relaxation training program for women with preterm labour on pregnancy outcomes: a controlled clinical trial. Int J Nurs Stud. 2012;49:257-64. Medline:21968280 doi:10.1016/j.ijnurstu.2011.09.007

30 Flax VL, Negerie M, Usman A, Leatherman S, Daza E, Bentley M. Integrating group counselling, cell phone messaging, and participant–generated songs and dramas into a microcredit program increases Nigerian women's adherence to International Breastfeeding Recommendations. J Nutr. 2014;144:1120-4. Medline:24812071 doi:10.3945/jn.113.190124

31 Gisore P, Shipala E, Otieno K, Rono B, Marete I, Tenge C, et al. Community based weighing of newborns and use of mobile phones by village elders in rural settings in Kenya: a decentralised approach to health care provision. BMC Pregnancy Childbirth. 2012;12:15. Medline:22429731 doi:10.1186/1471-2393-12-15

32 Jareethum R, Titapant V, Tienthai C, Viboonchart S, Chuenwattana P, Chatchainoppakhun J. Satisfaction of healthy pregnant women receiving short message via mobile phone for prenatal support: a randomized controlled trial. J Med Assoc Thai. 2008;91:458-63. Medline:18556852

33 Jiang H, Wen LM, Hu Q, Yang D, He G, Baur LA, et al. Effect of short message service on infant feeding practice. Findings from a community–based study in Shanghai, China. JAMA Pediatr. 2014;168:471-8. Medline:24639004 doi:10.1001/jamapediatrics.2014.58

34 Khorshid MR, Afshari P, Abedi P. The effect of SMS messaging on the compliance with iron supplementation among pregnant women in Iran: a randomized controlled trial. J Telemed Telecare. 2014;20:201-6. Medline:24803276 doi:10.1177/1357633X14533895

35 Labrique AB, Paul R, Sikder S, Wu LS, Jahan N, West KP Jr, et al. Mobile phones as disruptive agents in the pathway to mortality during emergency obstetric crises in rural Bangladesh. American Journal of Tropical Medicine and Hygiene. Conference: 60th Annual Meeting of the American Society of Tropical Medicine and Hygiene, ASTMH 2011 Philadelphia, PA United States. Conference Publication: (var.pagings). 85 (6 SUPPL. 1; page 36).

36 Lin H, Chen W, Luo L, Congdon N, Zhang X, Zhong X, et al. Effectiveness of a short message reminder in increasing compliance with pediatric cataract treatment: A randomized trial. Ophthalmology. 2012;119:2463-70. Medline:22921386 doi:10.1016/j.ophtha.2012.06.046

37 Lund S, Hemed M, Nielsen BB, Said A, Said K, Makungu MH, et al. Mobile phones as a health communication tool to improve skilled attendance at delivery in Zanzibar: a cluster–randomised controlled trial. BJOG. 2012;119:1256-64. Medline:22805598 doi:10.1111/j.1471-0528.2012.03413.x

REFERENCES

38 Lund S, Rasch V, Hemed M, Boas IM, Said A, Said K, et al. Mobile phone intervention reduces perinatal mortality in Zanzibar: secondary outcomes of a cluster randomized controlled trial. JMIR Mhealth Uhealth. 2014;2:e15. Medline:25098184 doi:10.2196/mhealth.2941

39 Lund S, Nielsen BB, Hemed M, Boas IM, Said A, Said K, et al. Mobile phones improve antenatal care attendance in Zanzibar: a cluster randomized controlled trial. BMC Pregnancy Childbirth. 2014;14:29. Medline:24438517 doi:10.1186/1471-2393-14-29

40 Oyeyemi SO, Wynn R. Giving cell phones to pregnant women and improving services may increase primary health facility utilization: a case–control study of a Nigerian project. Reprod Health. 2014;11:8. Medline:24438150 doi:10.1186/1742-4755-11-8

41 Seidenberg P, Nicholson S, Schaefer M, Semrau K, Bweupe M, Masese B, et al. Early infant diagnosis of HIV infection in Zambia through mobile phone testing of blood test results. Bull World Health Organ. 2012;90:348-56. Medline:22589568 doi:10.2471/BLT.11.100032

42 Sellen D, Mbugua S, Webb GA, Kalungu S, Sarange C, Lou W, et al. A randomized controlled trial indicates benefits of cell phone based peer counseling to support exclusive breastfeeding in Kenya. Annals of Nutrition and Metabolism 2013; Conference: 20th International Congress of Nutrition Granada Spain. Conference Start: 20130915 Conference End: 20130920. Conference Publication: (var.pagings). 63:751.

43 Sharma R, Hebbal M, Ankola AV, Murugabupathy V. Mobile–phone text messaging (SMS) for providing oral health education to mothers of preschool children in Belgaum City. J Telemed Telecare. 2011;17:432-6. Medline:22025742 doi:10.1258/jtt.2011.110416

44 Simonyan D, Gagnon MP, Duchesne T, Roos–Weil A. Effects of a telehealth programme using mobile data transmission on primary healthcare utilisation among children in Bamako, Mali. J Telemed Telecare. 2013;19:302-6. Medline:24163292 doi:10.1177/1357633X13503429

45 Corp S. Stata Statistical Software: Release 11. College Station, TX: Stata Corp LP, 2009.

46 Higgs ES, Goldberg AB, Labrique AB, Cook SH, Schmid C, Cole CF, et al. Understanding the role of mHealth and other media interventions for behaviour change to enhance child survival and development in low– and middle–income countries: an evidence review. J Health Commun. 2014;19 Suppl 1:164-89. Medline:25207452 doi:10.1080/10810730.2014.929763

47 Aranda–Jan CB. Mohutsiwa–Dibe, Loukanova S. Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. BMC Public Health. 2014;14:188. Medline:24555733

48 Hall CS, Fottrell E, Wilkinson S, Byass P. Assessing the impact of mHealth interventions in low– and middle–income countries: what has been shown to work? Glob Health Action. 2014;7:25606. Medline:25361730

49 National Institutes of Health. Mobile Health: Technology and Outcomes in Low and Middle Income Countries (R21) June 4th 2014. Available: http://grants.nih.gov/grants/guide/pa-files/PAR-14-028.html. Accessed: 8 April 2015.

50 Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR MHealth and UHealth. 2015;3:e20. Medline:25803705 doi:10.2196/mhealth.3335

journal of global health

PAPERS

Management of childhood diarrhea among private providers in Uttar Pradesh, India

Christa L. Fisher Walker[1], Sunita Taneja[2], Laura M. Lamberti[1], Amnesty Lefevre[1], Robert Black[1], Sarmila Mazumder[2]

[1] Johns Hopkins Bloomberg School of Public Health, Department of International Health, Baltimore, MD, USA

[2] Society for Applied Studies, New Delhi, India

Background In Uttar Pradesh (UP), India, a new initiative to introduce zinc and reinvigorate ORS for diarrhea treatment in the public and private sectors was rolled out in selected districts. We conducted an external evaluation of the program that included assessing the knowledge and practices of private sector providers 6 months after the initial program rollout.

Methods We conducted interviews and direct observations among a randomly selected group of formal and informal private sector providers in 12 districts of UP. We calculated summary statistics for reported provider characteristics, diarrhea treatment knowledge and preferred treatments, as well as the treatments advised during consultation with a child with diarrhea.

Results We interviewed 232 providers, of whom 67% reported receiving a diarrhea treatment training/drug detailing visit. In the interview, 14% of providers reported prescribing zinc to all children with diarrhea and 36% reported prescribing zinc to more than half of diarrhea cases. During direct observation, ORS and zinc were prescribed by 77.3% and 29.9% of providers, respectively. Treatments other than zinc and ORS were also commonly prescribed, including antibiotics (61.9%) and antidiarrheals (17.5%).

Conclusion Adequate treatment of childhood diarrhea with zinc and ORS remains a challenge among private sector providers in rural UP, India. Additional training and knowledge transfer activities are needed to curb the overprescription of antibiotics and antidiarrheals and to increase the confidence of private providers in advising zinc and ORS. In addition, policymakers and program implementers must ensure collaborative efforts to target and meaningfully engage informal private providers who play a major role in childhood diarrhea treatment in hard–to–reach areas.

Correspondence to:
Sunita Taneja
Center for Health Research and Development
Society for Applied Studies
New Delhi, India
sunita.taneja@sas.org.in

Diarrhea is a leading cause of morbidity and mortality among children less than 5 years of age in low– and middle–income countries [1]. Most diarrheal deaths can be prevented by the simple and effective treatment regimen currently recommended by the World Health Organization, which includes Oral Rehydration Salts (ORS) to prevent and treat dehydration, zinc supplementation for 10–14 days, and continued feeding [2]. The availability of ORS has been widespread in India since the 1980s, and yet it is only used to treat one quarter of diarrheal episodes [3]. Similarly, de-

spite the endorsement of zinc by the Indian Academy of Pediatrics (IAP) in 2004 and 2006 and adoption by the Indian government [4,5], it was not widely available in the public or private sectors in UP prior to implementation of the Diarrhea Alleviation through Zinc and ORS Therapy (DAZT) program in 2011.

In UP, caregivers of children with diarrhea more commonly seek care from private as opposed to public sector providers [6]. The private sector is comprised of providers with formal medical degrees and those practicing in the informal sector, many of whom do not have a license to practice and thus are not recognized by government. Evaluations of private sector providers in India have found that even informal providers are capable of delivering services of relatively high quality for basic medical care if knowledge and competency are high [7]. However, a recent provider assessment using surveys and patient vignettes in Bihar, India found a considerable gap between knowledge and practice with regard to the treatment of pediatric diarrhea [8]. Compared to ORS, zinc is a relatively new addition to the advised childhood diarrhea treatment protocol, and there is a dearth of available evidence on the acceptability of zinc among practitioners in rural areas, many of whom are removed from formal training resources and the influences of pediatric associations promoting national guidelines. In rural India, where up to 80% of children are brought to the private sector for care [9] and informal providers may outnumber qualified physicians 10 to 1, it is difficult to access high quality and consistent care without addressing the importance of the private sector in scaling–up adequate diarrhea treatment [7].

We present the results of a private sector provider assessment conducted in Uttar Pradesh after the rollout of the DAZT program, which aimed to improve diarrhea treatment for children under 5 years of age. The aim of this assessment was to characterize the childhood diarrhea treatment knowledge and practices of both formal and informal private sector providers approximately one year after project roll–out.

METHODS

DAZT program description

DAZT was a 4–year project (2011 – 2014) supported by the Bill and Melinda Gates Foundation, which aimed to enhance the uptake of zinc and ORS in rural UP, India. The DAZT project aimed to increase the coverage of ORS and zinc for treatment of diarrhea among children <5 years of age in 12 selected districts. The Johns Hopkins School of Public Health (JHSPH), in collaboration with the Society for Applied Studies, conducted an independent evaluation of the project activities in both the public and private sectors.

The overall goal of the private sector activities, which were led by FHI360, was to increase ORS and zinc prescribing by both formal and informal private sector providers. Formal providers included those who had completed government–recognized medical degrees (MD/MBBS). Informal providers included those with no medical training or certificates in traditional *Ayush* (Ayurveda, Yoga & Naturopathy, Unani, Siddha and Homeopathy) medicine. To target both groups, FHI360 used a two prong "push – pull" strategy. The "push" focused on changing diarrhea prescription practices among key opinion leaders through routine drug detailing for informal sector providers and formal training sessions for practicing physicians. To execute the "pull" component of the strategy, FHI360 recruited and trained staff from non–governmental organizations (NGOs) and private pharmaceutical companies in adequate childhood diarrhea treatment with ORS and zinc. The trained staff visited the practices of rural informal private sector providers to promote and sell ORS and zinc. By providing face–to–face meetings that included a short information video and information materials, FHI360 created demand for appropriate diarrhea treatment.

DAZT evaluation

The external evaluation of the DAZT program aimed to assess the coverage, quality and cost–effectiveness of implementation efforts through population–based household surveys and provider assessments. Additional details of the program and the results of the household coverage surveys have been published previously [10]. The provider assessment was timed to provide critical information on provider knowledge and behavior early enough to inform and enable programmatic adjustments.

Sample size

We generated a probability proportional to size (PPS) random sample of 29 tehsils (~50% of all tehsils) across the 12 districts in UP. By PPS sampling, the proportion of tehsils sampled from each district was equal to the proportion of the private provider population operating in that district relative to the total informal provider population across the 12 districts. Using a sampling frame of providers identified by FHI360 during implementation of the "push – pull" strategy, we randomly selected 8 private providers per tehsil to achieve the required sample size of 232. The sample size requirement was calculated assuming zinc prescribing of 20% at the time of the survey and accounting for 10% margin of error, a design effect of 1.365 (personal communication, S. Taneja) and 15% refusal. Formal and informal providers were sampled as one unit.

Data collection

We conducted the provider assessment from June –July, 2012, one year after roll–out of the program in June 2011. The survey instruments were developed based on previous surveys conducted by the investigators in similar populations. The survey was carried out by Mindfield Research, an experienced research firm based in New Delhi. The team was comprised of interviewers with previous interview experience who were fluent in both English and Hindi and comfortable interviewing providers. We conducted training of the survey team in New Delhi over the course of 3 days. Training included a thorough review of the survey protocol and survey tools, as well as pilot testing of the interview and observation forms using mock interviews and mock observations.

The trained survey teams visited the selected tehsils (administrative regions denoting sub–districts) according to a preset schedule. Interviewers identified the location of selected providers' shops/clinics and visited multiple times on the given day in an effort to find the provider. In the case where a provider was not located, the interviewer asked at least 3 other providers and/or community leaders about the whereabouts of the specified provider. If the interviewer was unable to locate the provider by the end of the day, the identified provider was dropped and replaced by the next randomly selected provider on the list who was from the same tehsil but not the same village; this methodology was employed to avoid biasing the sample towards easily accessible providers at the village–level.

Interviewers informed selected providers that the purpose of the visit was to observe the provider treating a child with diarrhea and to subsequently interview the provider about the diarrhea treatment practices typically provided. Interviewers obtained informed written consent and then waited for a caregiver of a child 2–59 months of age to seek care for diarrhea. Standard practice was to conduct the observation before the interview so questions asked during the interview would not bias the treatment provided during the observation. Prior to the observation, the interviewer also obtained verbal consent from the caregiver of the sick child. The interviewer remained a silent observer during the provider's interaction with the caregiver and child and used a standardized form to log questions on the history of the episode and treatments recommended and/or treatment referrals. The observation lasted approximately 30 minutes.

The interview portion of the assessment lasted approximately one hour and included short vignettes describing children with a range of diarrheal episode symptoms indicative of varying degrees of severity and comorbidities. For each vignette, the provider was asked whether he would refer or treat the child and, if he would treat, he was asked to describe the advised treatment regimen in detail.

The survey form was designed to collect categorical responses, including the option of "other," in which case the interview recorded the exact detailed response given by the provider. The interview also consisted of questions on diarrhea treatment knowledge and practices and access to routinely available ORS and zinc supplies. To ensure confidentiality, all interviews and observed treatment exercises took place in private locations in the presence of survey team member(s) and the provider alone.

After each day of fieldwork, the survey forms were double checked by the supervisor and incomplete entries or logical errors were corrected by contacting the provider immediately. This process ensured that all final forms were complete and free of logical errors prior to photocopying and data entry. The completed surveys were photocopied, with one copy sent via a secure courier to the data entry team at the Society for Applied Studies in New Delhi and one copy remaining with the survey team in the field. If inconsistencies were identified during the data entry process, the survey team attempted to clarify the issue by revisiting the provider if possible.

We received ethical approval from the Johns Hopkins University Bloomberg School of Public Health Institutional Review Board (IRB) and the Society for Applied Studies Ethics Review Committee. All providers signed a full informed consent document. All caregivers of observed children gave verbal consent.

Statistical methods

We included private sector providers from both the formal and informal sectors, representing the breadth of the private sector DAZT program in UP. The survey was not designed to detect differences between informal and formal sector providers. However, we conducted *t* tests and χ^2 tests to determine if there were differences in the sociodemographic characteristics of the two groups. We also compared exposure to drug seller informational visits about diarrhea and recall of having seen a video on diarrhea treatment and/or posters promoting zinc and ORS in the last 6 months.

Although interviews were conducted for all participating providers, observations were not available for those who were not visited by a child with diarrhea on the day of the assessment. To assess the generalizability of providers who completed both observations and interviews to those who only completed interviews, we conducted *t* tests and χ^2 tests of basic sociodemographic characteristics and program exposure and found no differences (data not shown). We therefore combined both sets of providers for all subsequent analyses on interview data. We generated summary statistics for providers' responses to the clinical vignettes during the interview and adherence to the current WHO

and IAP guidelines for childhood diarrhea treatment during the observed consultation [2,5]. All statistical analyses were conducted using Stata 12.0 software [11].

RESULTS

Of the 295 private sector providers included in the sample, 232 (78.6%) agreed to participate in the assessment. Of these, 97 were included in both the interview and the observation and 135 provided information in the interview alone (**Figure 1**). In **Table 1**, we present an overview of the demographics and training of the providers included in our assessment stratified by the informal and formal sector. We also present an overview of the demographics and training of the providers included in our assessment stratified by informal vs formal sector. Among all interviewed providers, 30.6% had a formal medical degree (including MBBS/MD, pharmacy, and nursing) and 69.4% were practicing as part of the informal sector with either Ayurvedic/homeopathic training (21.1%) or no recognized degree or training program (48.3%).

Ninety–five percent of providers reported prescribing drugs/medications but only 89% reported providing diagnosis or patient consultation; the remaining 11% only sold drugs/medications. In the previous 6 months, 155 providers (67%) reported they had been visited by someone who provided training/information about diarrhea treatment. In addition, 84% of private providers reported seeing posters advertising zinc for diarrhea treatment. There were no differences between formal and informal sector providers with regard to having been visited by a drug seller or having seen a video about diarrhea treatment in the last 6 months, nor were there any difference between the two groups in the proportion who had seen a billboard advertising zinc and ORS. Because the survey was not designed to compare provider types and given there were no differences in these programmatic indicators, we did not stratify by formal/informal in subsequent analyses.

In **Table 2**, we summarize provider responses on the diarrhea treatments typically recommended during five case vignettes. For simple acute diarrhea (5 loose, watery stools/d for 3 days), providers reported recommending ORS (68.1%) and antibiotics (65.9%) most often. However, the proportion reportedly prescribing ORS decreased for the remaining four scenarios, each of which represented more severe diarrhea. For simple acute diarrhea, 35.8% of providers reported prescribing zinc. In all five scenarios, providers reported advising antibiotics more frequently than zinc. As the severity of the case vignette increased with either signs of dehydration, duration of illness, or accompanying signs and symptoms, providers more frequently reported referral in lieu of any type of treatment, including life–saving ORS.

When providers were asked specifically about willingness and frequency of prescribing zinc and ORS, a small percentage reported prescribing zinc to all patients (14%), and an additional 36% reported prescribing zinc to at least 50% of cases [data not shown]. When asked what circumstance would lead a provider to NOT recommend zinc, the most common responses were not fully understanding zinc treatment (57.9%) and not having zinc in stock (30.8%). Providers were asked to recall the duration of zinc treatment; 15.6% reported duration of at least 10 days, in adherence to the current WHO guidelines [2]. Eighty–six percent of private providers reported routinely recommending ORS

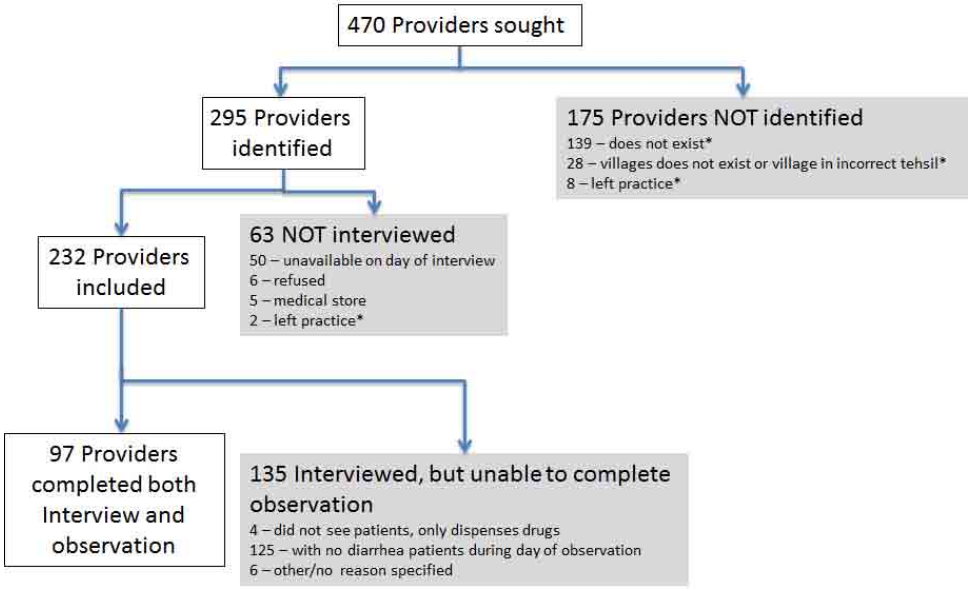


Figure 1. Study sample. Asterisk indicates that the information collected from village administrative heads, other providers, and local shops.

Table 1. Comparison of the characteristics of informal and formal providers

Characteristics	Informal providers, N = 161 (%)	Formal providers, N = 71 (%)	P–value
Proportion male	161 (100.0)	71 (100.0)	–
Mean age in years (SD)	43.6 (12.2)	38.9 (11.4)	0.006
Mean years of education (range)	13.1 (3.0)	15.6 (1.4)	<0.001
Proportion who sell drugs/ medicines	150 (93.2)	71 (100.0)	0.024
Proportion who consult and provide a diagnosis for patients	146 (90.7)	60 (84.5)	0.169
Mean years working as a private provider	14.5 (10.6)	11.7 (10.4)	0.059
Mean number of days worked in the last week (SD)	6.4 (1.4)	6.5 (1.1)	0.597
Proportion who recalled an informational visit in past 6 months by any source who spoke about pediatric diarrhea treatment	111 (68.9)	44 (62.0)	0.299
Proportion who recalled seeing a video about diarrhea treatment in last 6 months	56 (34.8)	21 (29.6)	0.438
Proportion who have seen posters advertising the use of zinc for the treatment of diarrhea in last 6 months	136 (84.5)	58 (81.7)	0.598

SD – standard deviation

Table 2. Treatment practices reported by private providers in interview (n = 232)

Description of a child brought to you with:	Treatments typically prescribed (No, %)*							Only refer (not treat)
	ORS	Sugar – salt solution	Increased fluids	Zinc	Antibiotics	Antidiarrheals	IV	
5 loose/watery stools per day for 3 days with no signs of dehydration	158 (68.1)	14 (6.0)	16 (6.9)	83 (35.8)	153 (65.9)	50 (21.6)	1 (0.4)	43 (18.5)
5 loose/watery stools per day for 3 days with sunken eyes and lethargic	107 (46.1)	7 (3.0)	9 (3.9)	64 (27.6)	103 (44.4)	24 (10.3)	54 (23.3)	106 (45.7)
4 loose/watery stools per day for 15 days	97 (41.8)	9 (3.9)	7 (3.0)	61 (26.3)	82 (35.3)	25 (10.8)	17 (7.3)	122 (52.6)
The presence of blood in the stools	98 (42.2)	5 (2.2)	10 (4.3)	59 (25.4)	130 (56.0)	40 (17.2)	11 (4.7)	91 (39.2)
Fever, fast breathing, and 3 loose, watery stools per day for 3 days	79 (34.1)	5 (2.2)	11 (4.7)	49 (21.1)	93 (40.1)	21 (9.1)	5 (2.2)	128 (55.2)

ORS – Oral rehydration salts

*Multiple responses accepted.

as part of their practice; yet 42.7% of those could not correctly describe how to prepare ORS. Zinc and ORS were in–stock (ie, providers able to show the stocked products) for 38% and 69% of providers at the time of the interview, respectively.

A direct observation of the provider treating a child 2–59 months of age with diarrhea was conducted for 97 (42%) private providers (**Table 3**). Of the children treated during the observation session, 71% were male children and the median age was 24 months. Among the 97 providers who participated in the observations, 98% asked at least one question with regard to the history of the diarrheal episode. Zinc was sold in 16.5% of cases and caregivers were recommended to go elsewhere for zinc in an additional 13.4% of cases. Among the 28 providers who recommended zinc, 50% gave no reason as to why zinc would benefit the child and only three administered the first dose during the treatment session. Providers prescribing zinc gave variable instructions on the duration of zinc treatment; 50% recommended zinc for 10–14 days; 18% for 7 days, and 32% either recommended zinc for <7 days or did not mention duration of treatment. ORS was sold as a treatment in 56.7% of cases with an additional 20.6% instructed to purchase ORS elsewhere. Of the providers who recommended ORS, 50% gave correct preparation instructions. Providers commonly prescribed treatments other than zinc and ORS, including antibiotics (61.9%) and antidiarrheals (17.5%). Twenty–six providers (26.8%) were in compliance with current WHO/IAP guidelines and either sold or prescribed both zinc and ORS. Only 2 providers sold or prescribed zinc without ORS.

DISCUSSION

We conducted a provider assessment of 232 providers from both the formal and informal private sectors in UP, India. Private providers are the mostly widely sought sources of careseeking for pediatric diarrhea in UP; yet there is a dearth of information on private, especially informal, providers and thus it is critical to understand their knowledge and willingness to provide appropriate diarrhea treatment, including zinc and ORS. We carried out this study using short vignettes and direct observations one–year after roll–out of program implementation, by which time a pharma-

Table 3. Treatment behaviors of private providers observed during the treatment of a child 2–59 months of age with diarrhea (n = 97)

Provider behavior	No. (%)
Proportion who asked at least 1 question about the diarrhea episode:	95 (97.9)
Frequency of diarrhea	53 (55.8)
Character of stool	62 (65.3)
Duration of diarrhea	76 (80.0)
Blood in stool	10 (10.5)
Vomiting	10 (10.5)
Proportion of children given zinc	16 (16.5)
Proportion of caregivers recommended to obtain zinc elsewhere	13 (13.4)
Place from where caregiver told to get zinc:	
Chemist	12 (92.3)
Did not specify a particular place to go	1 (7.7)
Among providers who advised zinc (gave or referred, n = 28)	
At least 1 benefit of zinc told to caregiver*	14 (50)
Reduces duration of diarrhea	1 (3.6)
Reduces frequency of stool	6 (21.4)
Reduces stool volume	2 (7.1)
Good for diarrhea	8 (28.6)
Zinc acts a tonic after diarrhea	3 (10.7)
Proportion who demonstrated administering 1st zinc dose	3 (10.7)
Advised to give zinc for:	
2–5 days	3 (10.7)
7 days	5 (17.9)
10 days	2 (7.1)
14 days	12 (42.9)
Did not advise duration	6 (21.4)
Proportion who sold ORS during the observation	55 (56.7)
Number of packets sold (n = 55):	
1	49 (89.1)
2	6 (10.9)
Proportion who advised caregivers to purchase ORS	20 (20.6)
Benefits of ORS as told by providers to caregivers among those advised ORS (n = 71)*	
Good for rehydration/prevention of dehydration	36 (50.7)
Good for diarrhea	12 (16.9)
Benefits not told	26 (26.8)
Correct method of preparation advised to caregivers	47 (66.2)
Providers who sold or prescribed zinc and ORS	26 (26.8)
Number of children given any feeding advice	40 (41.2)
Proportion advised medications other than zinc and ORS	87 (89.7)

ORS – oral rehydration salts

*Multiple responses.

ceutical or NGO representative promoting and selling zinc and ORS should have visited all providers [12].

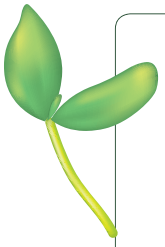
It is not surprising that we found providers were more likely to report their intention to prescribe ORS and zinc in the short vignette portion of the interview than in the observation portion of the study. This same discrepancy has been observed in Bihar, India in an evaluation of a similar program promoting zinc and ORS. Mohanan et al. found poor overall knowledge of diarrhea treatment among the 340 included providers, and none were found to treat diarrhea correctly when tested by an unannounced standard patient [8].

Despite IAP Guidelines that do not recommend antibiotics for the treatment of acute diarrhea, antibiotics remain the first line treatment for most private providers [5]. We observed consistently higher rates of antibiotic prescribing compared to zinc. In previous effectiveness studies, zinc has been shown to displace the unnecessary overuse of antibiotics [6,12], but this was not observed in our evaluations. The pharmaceutical training provided, as part of the DAZT program did not address the unnecessary use of antibiotics nor suggest that zinc could and should replace antibiotics for simple acute diarrhea. Pharmaceutical representatives may have been promoting other treatments for diarrhea in addition to zinc and ORS, such as antibiotics and antidiarrheals. Providers who prescribed zinc typically did so in addition to antibiotics and ORS. Only two providers recommended zinc in the absence of ORS, lessening concerns that NGO and pharmaceutical representatives might place more attention on zinc than life–saving ORS during drug detailing visits. Program planners employed the drug detailing approach to build rapport with providers before challenging their current practices. The visits were designed to increase confidence among informal providers, the majority of whom do not receive routine or formal training. Our results illustrate that the tendency to overprescribe antibiotics will not be changed quickly, even with the introduction of zinc. More research is warranted to determine how to best influence lasting change in prescribing practices; tactics might include group training sessions, incentives, and/or community promotion.

There were several limitations to our study, including the failure to conduct observations of all providers. It is possible that the prescribing knowledge and practices of observed providers are not representative of providers for whom an observation was not possible due to logistical constraints. We conducted the provider assessment during the dry season to facilitate the logistics of data collection, but as a consequence of this timing, diarrhea prevalence was low in some villages, resulting in less diarrhea care-seeking. Our data collectors waited at each provider's practice for one day but proceeded with the interview alone if no child eligible for observation sought care during that time frame. We also recognize that there is no gold standard method for observing providers in practice. Direct observation may bias providers to give higher quality care—a phenomenon known as the Hawthorne effect. As such, we expect the observation results to represent the best–case scenario, such that the care given to a child during an observed treatment interaction is the highest quality care the provider is able to give [10]. Lastly, our assessment is also limited by its cross–sectional nature and the failure to draw upon repeated measurements across the du-

ration of implementation. The evaluation was originally designed to assess the knowledge and practices of providers early in the intervention and then again at the end of the program, but the plan was changed mid–course by the broader project steering committee. We maintain, however, that the results of this one–time assessment are critical because they shed light on the challenges of changing the behavior of rural providers even early in the program implementation period when training is fresh and stocks are largely in place.

The DAZT private sector model mimics the reality of how drug information is currently delivered to private providers through drug detailing visits. The expansion of drug detailing to include zinc and ORS thus represents a sustainable method, especially for reaching the informal sector with critical information and products for adequately treating childhood diarrhea. The DAZT model has also highlighted the challenges in identifying and working with informal private providers who often operate underground to avoid government penalties for practicing medicine without recognized credentials. However, despite these challenges, informal private providers are often the first choice of caregivers and thus a worthwhile target of diarrhea management programming. If substantive improvements in diarrhea treatment are to be made in rural India, improvements in the treatment practices of the informal private sector will need to be addressed by increasing access to current guidelines and providing formal and informal training opportunities. Policymakers often resist developing programs targeting the informal sector, but denying the role of informal private providers in treating the rural poor is detrimental to the overall goal of improved diarrhea treatment for young children. Programs like DAZT aim to demonstrate that informal providers, including those with little–to–no education, can be taught to adequately treat childhood diarrheal episodes with ORS and zinc and to recognize the signs and symptoms requiring referral. The low–level training provided to public sector community health workers in many countries could be expanded to cover the private sector in areas where the latter is already providing the majority of care and treatment, such as rural UP. Diarrhea remains an important cause of morbidity and mortality among young children, especially the poorest and least–served, in low– and middle–income countries. Policymakers and program planners may not be able to influence the source through which children receive care, but collaborative and targeted efforts can improve the quality of care children receive through the most frequently utilized sources, in turn improving outcomes for all young children with diarrhea.



Acknowledgements: The authors would like to thank the providers and caregivers who participated in the survey. The authors thank FHI 360 for their feedback with regard to interpretation of the final results. This work was supported by a grant from the Bill and Melinda Gates Foundation. The funder played no role in the analysis or final interpretation and presentation of the data.

Funding: This work was supported by a grant from the Bill and Melinda Gates Foundation.

Authorship declaration: CLFW conceptualized and designed the study, conducted the analysis, drafted the initial manuscript and approved the final manuscript as submitted. ST and SM conceptualized and designed the study, oversaw the field operations and data entry, contributed to the manuscript and approved the final manuscript as submitted. LMLand AL contributed to the manuscript and approved the final manuscript as submitted. REB conceptualized and designed the study, contributed to the manuscript and approved the final manuscript as submitted.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). The authors declare no competing interests.

REFERENCES

1 Liu L, Johnson HL, Cousens S, Perin J, Scott S, Lawn JE, et al; Child Health Epidemiology Reference Group of WHO and UNICEF. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. Lancet. 2012;379:2151-61. Medline:22579125 doi:10.1016/S0140-6736(12)60560-1

2 WHO/UNICEF. Joint Statement: clinical management of acute diarrhoea (WHO/FCH/CAH/04.07). Geneva and New York: World Health Organization, Department of Child and Adolescent Health and Development, and United Nations Children's Fund, Programme Division, 2004.

3 International Institute for Population Sciences (IIPS) and MACRO International. National Family Health Survey (NFHS-3) 2005-06. Mumbai: International Institute for Population Sciences (IIPS) and MACRO International, 2007.

REFERENCES

4 Bhatnagar S, Bhandari N, Mouli UC, Bhan MK; IAP National Task Force. Consensus Statement of IAP National Task Force: status report on management of acute diarrhea. Indian Pediatr. 2004;41:335-48. Medline:15123862

5 Bhatnagar S, Lodha R, Choudhury P, Sachdev HP, Shah N, Narayan S, et al; Indian Academy of Pediatrics. IAP Guidelines 2006 on management of acute diarrhea. Indian Pediatr. 2007;44:380-9. Medline:17536143

6 Baqui AH, Black RE, El Arifeen S, Yunus M, Zaman K, Begum N, et al. Zinc therapy for diarrhoea increased the use of oral rehydration therapy and reduced the use of antibiotics in Bangladeshi children. J Health Popul Nutr. 2004;22:440-2. Medline:15663177

7 Radwan I. India - private health services for the poor. Washington, DC: The World Bank, 2005.

8 Mohanan M, Vera-Hernández M, Das V, Giardili S, Goldhaber-Fiebert JD, Rabin TL, et al. The know-do gap in quality of health care for childhood diarrhea and pneumonia in rural India. JAMA Pediatr. 2015;169:349-57. Medline:25686357 doi:10.1001/jamapediatrics.2014.3445

9 International Institute for Population Sciences. District level household and facility survey (DLHS-3), 2007-08: Gujarat, India. Mubai: IIPS, 2010.

10 Lamberti LM, Taneja S, Mazumder S, LeFevre A, Black RE, Walker CL. An external evaluation of the Diarrhea Alleviation through Zinc and ORS Treatment program in Gujarat and Uttar Pradesh, India. J Glob Health. 2015;5:020409. Medline:26682045 doi:10.7189/jogh.05.020409

11 StataCorp. LP. Stata 12.0. College Station, TX: StrataCorp, 2012.

12 Bhandari N, Mazumder S, Taneja S, Dube B, Agarwal RC, Mahalanabis D, et al. Effectiveness of zinc supplementation plus oral rehydration salts compared with oral rehydration salts alone as a treatment for acute diarrhea in a primary care setting: a cluster randomized trial. Pediatrics. 2008;121:e1279-85. Medline:18450870 doi:10.1542/peds.2007-1939

journal of global health

Electronic supplementary material:
The online version of this article contains supplementary material.

PAPERS

Integration of antenatal care services with health programmes in low– and middle–income countries: systematic review

Thyra E de Jongh[1], Ipek Gurol–Urganci[2], Elizabeth Allen[3], Nina Jiayue Zhu[3], Rifat Atun[3]

[1] Technopolis Group, Amsterdam, the Netherlands
[2] London School of Hygiene and Tropical Medicine, London, UK
[3] Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Harvard University, Boston MA, USA

Background Antenatal care (ANC) presents a potentially valuable platform for integrated delivery of additional health services for pregnant women–services that are vital to reduce the persistently high rates of maternal and neonatal mortality in low– and middle–income countries (LMICs). However, there is limited evidence on the impact of integrating health services with ANC to guide policy. This review assesses the impact of integration of postnatal and other health services with ANC on health services uptake and utilisation, health outcomes and user experience of care in LMICs.

Methods Cochrane Library, MEDLINE, Embase, CINAHL Plus, POPLINE and Global Health were searched for studies that compared integrated models for delivery of postnatal and other health services with ANC to non–integrated models. Risk of bias of included studies was assessed using the Cochrane Effective Practice and Organisation of Care (EPOC) criteria and the Newcastle–Ottawa Scale, depending on the study design. Due to high heterogeneity no meta–analysis could be conducted. Results are presented narratively.

Findings 12 studies were included in the review. Limited evidence, with moderate– to high–risk of bias, suggests that integrated service delivery results in improved uptake of essential health services for women, earlier initiation of treatment, and better health outcomes. Women also reported improved satisfaction with integrated services.

Conclusions The reported evidence is largely based on non–randomised studies with poor generalizability, and therefore offers very limited policy guidance. More rigorously conducted and geographically diverse studies are needed to better ascertain and quantify the health and economic benefits of integrating health services with ANC.

Correspondence to:
Rifat Atun
Department of Global Health and Population
Harvard T.H. Chan School of Public Health
665 Huntington Avenue
Building I, Room 1104A
Boston, MA 02115
ratun@hsph.harvard.edu

Since 2005, antenatal care (ANC) coverage has risen considerably worldwide [1]. The World Health Organization (WHO) estimates suggest that during 2005–2012 approximately 80.5% of pregnant women globally, including 71.8% of women in low–income countries, had at least one ANC visit during pregnancy [1]. ANC provides an opportunity for women to access effective interventions that reduce risks associated with pregnancy and improve their health and well–being, as well as that of their progeny. However, while there was considerable progress towards the Millennium Development Goals 4 (to reduce child mortality) and 5 (to improve maternal health), maternal and neonatal mortality from prevent-

able pregnancy– and birth–related complications remain high, particularly in low– and middle–income countries (LMICs) [2]. In 2013, around 289 000 women died during and following pregnancy and childbirth–the vast majority in low–resource settings [3]. Between one–third and one–half of these pregnancy–related deaths are due to preventable complications, such as eclampsia and haemorrhage, directly related to inadequate care [4]. Additionally, nearly three million newborns died during their first month of life, in large part due to insufficient provision of postnatal care (PNC) [2,5–8]. Lack of PNC not only affects neonatal mortality, but also has long–term negative impacts on the development of children who survive, as opportunities for promoting healthy home behaviours are missed. The unacceptably high maternal and neonatal mortality rates in LMICs suggest new approaches are needed to expand access to ANC, improve the quality of services provided during ANC contact, and strengthen continuity and quality of care through to the postnatal period.

In most LMICs pregnancy often marks a woman's first encounter with formal health services, and ANC can serve as an effective platform for a broad range of health interventions [9], including for the provision of services for conditions that increase the risk of complications during pregnancy (eg, malaria, sexually transmitted infections (STIs), HIV/AIDS, tuberculosis (TB), tetanus, and malnutrition). Integrating ANC with malaria, STIs, HIV/AIDS and TB services can also expand the reach of these programmes to a broader population [10]. In settings where the prevalence of such conditions is high, integrating ANC with cost–effective services like prevention of mother to child transmission (PMTCT) of HIV [11], intermittent preventative treatment in pregnancy for malaria, and provision of insecticide treated nets [9] would likely improve maternal and child health outcomes. The WHO has identified integration of ANC with other health services, including PNC, as a key strategy for reducing missed opportunities for patient contact and for effectively and comprehensively addressing the health and social needs of pregnant women and their children, thereby improving maternal and child health [5,8,9].

Integration in health systems is variously defined [12–15], referring to establishing joint systems for organisation, financing, management, planning and evaluation of health programmes at different levels of the health system (from health facilities to ministry of health level) to improve the efficiency and effectiveness of health systems [16]. Integrated care has also been defined by WHO as "bringing together inputs, delivery, management and organization of services related to diagnosis, treatment, care, rehabilitation and health promotion" in order to "improve services in relation to access, quality, user satisfaction and efficiency" [17,18]. The rationale for integrating health services is to improve user access to health services across the care continuum to meet users' health needs over time [19,20] and to create positive synergies among investments in health programmes [21].

However, 'injudicious integration' may also have harmful consequences for already constrained health systems [22]. For example, provision of multiple services during a single point of contact requires that health care providers be sufficiently trained in all aspects of the services concerned to ensure high quality care. But, in resource constrained systems training can take away health staff from frontline services [23]. Furthermore, provision of multiple services could stretch the already limited capacity, thus leading to long waiting times and hindering access for women who have to travel far to reach health facilities. In an attempt to reduce workload providers may reduce the time spent on consultations, thus compromising service quality.

To date few studies have systematically examined how integration of ANC with other services could influence health outcomes, service access, efficiency, or patient satisfaction [19,24–26]. Evidence to guide policy on the best ways to integrate ANC with PNC and other health services for pregnant women and integration impact is limited. This review examines the evidence on how integration of ANC services with PNC or other health services in LMICs affects health outcomes for women and children, health care provision (including processes, outputs, service quality) and costs. The review analyses ways in which the quality of ANC can be improved through integration with PNC and other health services. Specifically, the review focuses on the impact of integrated provision of ANC services, which can take different forms, such as co–location of ANC and PNC or other health services with a single point of access, through a well–connected referral system [27,28], or by merger of services within or across a domain of care [29].

METHODS

Criteria for considering studies for this review

We followed Cochrane guidelines for systematic reviews [30] and included both randomised controlled trials (RCT), where randomisation could be at individual or cluster level, and non–randomised studies (NRS). Non–randomised studies are defined in the Cochrane Handbook as quantitative studies that do not use randomisation to allocate units to comparison groups, but where allocation occurs in the course of usual treatment decisions or peoples' choices [30]. The NRS that were eligible for inclusion in this review were non–randomised controlled trials (NRCT), controlled before and after studies (CBA), interrupted time

series analyses (ITS), historically controlled studies, cohort and case–control studies.

Type of participants

We included studies focusing on pregnant women of all ages utilizing ANC services in LMICs.

Type of interventions

We considered any study that described a change from 'routine practice' with the intention to integrate provision of ANC services with i) PNC or ii) other health services. Integrated service provision models included:

- Co–location of services, using a single point of access;
- Collaboration between different service providers involved in a woman's care (eg, in integrated care teams);
- A well–organised referral system, with follow–up and feedback among different service providers.

We considered strategies promoting horizontal integration (ie, linking services at the same level of care domain), as well as vertical integration (ie, linking services across different levels of care) [29]. For inclusion, however, studies had to compare outcomes of the intervention against a control situation in which a similar set of services was delivered in a non–integrated way (ie, additional services were available to pregnant women, but were not routinely integrated into ANC).

Type of outcome measures

We explored the impact of ANC integration on health outcomes (including health behaviour and health status for mother and child, and user experience, such as user satisfaction) as well as health care outputs (including utilisation of services, access, coverage, quality, efficiency and cost) for all relevant users and providers, and including any adverse outcomes.

Search methods for identification of studies

We searched the Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Database of Systematic Reviews (Cochrane Reviews), Cochrane Database of Abstracts of Reviews of Effects (Other Reviews), MEDLINE (Ovid), Embase (Ovid), CINAHL Plus (EBSCO), Global Health (Ovid) and POPLINE on January 21, 2014. We used a comprehensive search strategy with no language or publication date restrictions. The search string for MEDLINE, which was tailored to each of the databases, is provided in **Online Supplementary Document**. The "integration" block was adapted from the search strings used in the Cochrane EPOC review of integration of PHC services [31] and the "LMIC" block was adapted and expanded from the Medline LMIC filter.

We checked the reference lists of all included studies and examined the bibliographies of relevant systematic reviews and meta–analyses identified during the search.

Data collection and analysis

We performed the selection of potentially eligible studies through a staged process. At every stage of the process, two authors independently assessed publications for their relevance and adherence to inclusion criteria. TdJ, EA and IGU first piloted and refined the selection process in a random sample of 100 studies to ensure high inter–rater agreement. In the first stage, the authors (TdJ, EA) evaluated publications for their potential relevance based on titles. Any title judged as potentially relevant by either of the authors was next assessed for eligibility on the basis of the abstract. All abstracts considered potentially eligible by both authors were retained for further scrutiny. Due to the large number of abstracts, those on which the authors disagreed were independently reviewed by a third author (IGU) who decided on its inclusion into the final round of screening. When no abstract was available, the publication was also retained in the selection until the full text was acquired and screened. In the final stage of screening, two authors (TdJ, EA) reviewed the full text of each retained publication to determine relevance and whether the publication met our inclusion criteria. If a study was published only as an abstract (eg, conference abstracts where full manuscript was not yet available), we only included the study if there was sufficient information presented in the abstract to demonstrate that it met the review's inclusion criteria and was of an acceptable methodological standard. In the case of disagreement between the authors, a third author (IGU) acted as an arbiter to decide upon the final inclusion.

Data extraction and management

For studies that were deemed eligible for inclusion, we extracted the data to a standardised form including key information such as administrative data (title, author, year of publication, country, setting, funding etc.); methods (stated study design, data relevant for risk of bias assessment, duration and completeness of follow–up); and information on participants, interventions and comparisons. Quantitative results for each study were separately extracted to an Excel™ spreadsheet for further analysis; and grouped by outcome measures as defined in the included studies. Two separate authors (EA, NZ) extracted the quantitative results, with independent verification by a third author (IGU).

Assessment of risk of bias in included studies

To assess the risk of bias in the included studies, we used standardised tools appropriate to different study designs. For RCT/NRCT/cRCT/CBA we used the criteria formulated

by the Cochrane Effective Practice and Organisation of Care (EPOC) Group, which rate each study on nine dimensions, namely: sequence generation; allocation concealment; baseline outcome measurement; baseline characteristics of participants; blinding of participants, personnel and outcome assessors; contamination; selective outcome reporting; and other sources of bias [32]. Each category was rated as low–risk, high–risk or unclear.

For cohort designs, case–control studies and historically controlled trials, we assessed risk of bias using the Newcastle–Ottawa scale, which contains only eight items and is simpler to apply than other checklists for NRS [33]. The scale uses a 'star' rating system with a maximum of nine stars, with ratings assigned in three categories: the selection of the study groups (four stars), the comparability of the groups (two stars) and the ascertainment of outcome of interest (three stars) (**Box 1**).

Box 1. The Newcastle Ottawa Scale

In the study group category one star could be awarded for each of the following 4 criteria: a) if the exposed group was representative of the average woman seeking antenatal care services and, where applicable, additional health services; b) if the control group was selected from the same community as the integrated services group, c) if the delivery of individual health services was ascertained from secure records or structured interviews, and d) if there was sufficient evidence that the outcome of interest was not present at the start of the study. In the group comparability category, one star was awarded if the study reported no significant differences in baseline characteristics. Two stars were awarded if there was statistical evidence of no baseline differences across groups or if the results were risk–adjusted (by minimum of maternal age). In the outcome category, three stars could be awarded if: a) the assessment of outcome was done by independent blind assessment or determined from secure records, b) follow–up was sufficiently long; and c) either loss to follow up was small (<5%) or if it could be sufficiently demonstrated that loss to follow–up was unlikely to have affected findings.

Assessment of heterogeneity and data synthesis

We considered whether it was appropriate to combine the studies in a meta–analysis by investigating heterogeneity in the methodologies (eg, type of service integration, study design setting and outcomes) and results of the included studies. As there was significant heterogeneity in the included studies and the study results were not expressed using consistent effect measures, we narratively summarise the findings. We also present the results of included studies in a Forest plot, but suppressed the pooled estimate, as recommended by the Cochrane Handbook [30]. We used the Forest plot to facilitate visualisation of the results, particularly to highlight the varied quality of the evidence and heterogeneity of results.

We used odds ratios (OR) as measures of effect for dichotomous outcomes. We had planned to use standardised mean differences (SMD) for continuous outcomes and where the study reported medians, to convert the medians to means using the methods proposed in Hozo and others [34]. However, for the three studies that reported continuous outcomes, either the standard deviation for the means or the ranges for medians were missing, therefore we present continuous outcomes as reported in original studies. The analysis only used data published in the studies.

RESULTS

In database searches, we identified 6416 unique citations. Of these, 922 titles were considered potentially relevant to this review. Of these citations, 842 included abstracts that were subsequently reviewed. Among the abstracts, 120 were considered potentially relevant. For an additional 80 citations no abstracts were available. These citations were all carried forward to the next stage of the screening process, in which the full text of the potentially eligible studies was reviewed. We retrieved the full text for 177 out of 200 citations. After screening against the inclusion criteria, we identified 14 citations, presenting data for 12 separate studies, that met all conditions and that were included in this review. One article that did not meet the inclusion criteria on its own was subsequently added, as it provided additional information on an already included study. The process of screening and selection is presented in **Figure 1**.

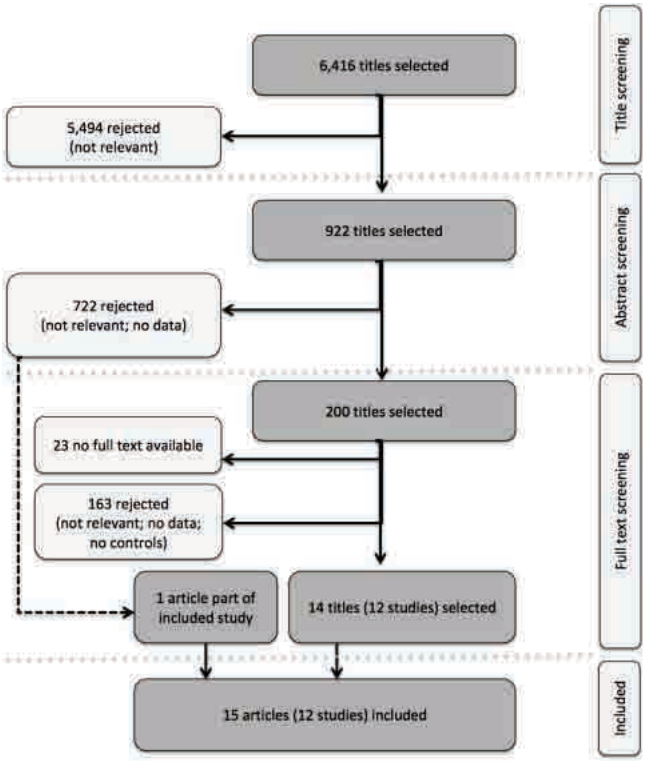


Figure 1. Flowchart showing process of screening and selection of studies for inclusion.

Description of included studies

Of the 12 studies included in this review, 10 were set in Sub–Saharan Africa: three in Kenya [35–39], three in South Africa [40–43], two in Mozambique [44,45], one in Zambia [46], and one in Malawi [47]. The other two studies were set in Asia, namely in Bangladesh [48] and Mongolia [49] (**Table 1** provides a summary of the included studies). All included studies had pregnant women, either with or without their newborns, as the principal study participants. Additionally, one article described the impact of integrated services from the point of view of health care providers [37]. Excluding the latter and one other study in which the number of participants was not specified [45], the included studies represented a total of 87 755 participants, with study sizes ranging from 164 [40] to 31 526 [46] participants.

Table 1. Summary of included studies

Study	Services integrated	Setting (participants)	Study design	Intervention description	Control description	Outcome measures
HIV						
Geelhoed 2013 [44]	ART, PMTCT	Mozambique (376)	Controlled before–and–after study	MCH nurses provided all recommended health interventions applicable to both mother and child, including follow–up of HIV–exposed infants and early infant diagnosis of HIV, during the antenatal, postnatal, family planning, growth monitoring, high–risk child and vaccination consultations.	In the health care facilities of the control group, the same services were provided separately, one type of services after another, as is routine in the Mozambican public health care system.	Follow–up of HIV–exposed infants (registration, follow–up visits, serological testing); MCH attendance; Acceptability of integrated services to health care providers.
van't Hoog 2005 [39]	PMTCT	Kenya (8231)	Historically controlled trial	HIV pre– and post–test counselling from an ANC nurse–counsellor; HIV testing at an on–site facility. The same counsellor also provided routine ANC preventive interventions like tetanus toxoid and sulfadoxine–pyrimethamine.	Opt–in HIV counselling was provided in a separate location within the hospital complex. HIV testing was conducted in an off–site laboratory.	Uptake of HIV counselling, testing and uptake of NVP.
Kasenga 2009 [47]	PMTCT	Malawi (1259)	Historically controlled trial	HIV testing and counselling services, and later on also management of sexually transmitted infections, were integrated within ANC.	Voluntary counselling and testing services were offered through a separate VCT unit at the outpatient department, through an opt–in approach.	Uptake of HIV testing
Killam 2010 [46]	ART	Zambia (31 536)	Stepped–wedge cluster non–randomised trial	Eligible women received ART in ANC until 6 weeks postpartum and then were referred to the general ART clinic.	Women found to be seropositive through ANC testing and eligible for ART were referred to the ART clinic, located on the same premises as ANC, but physically separated and separately staffed.	Proportion of treatment eligible pregnant women enrolling into HIV care within 60 d of HIV diagnosis; Proportion of women initiating ART during pregnancy.
Van der Merwe [40]	ART	South Africa (164)	Historically controlled trial	HIV testing, ART adherence counselling and treatment preparation took place within ANC. Thereafter, women were referred to hospital for initiation and follow–up of ARV treatment, which, whenever possible, was provided by the same staff members who began treatment preparation.	Pregnant women with indications for ARV treatment were referred to a hospital located approximately 1 km away, for preparation and initiation of treatment and long–term follow–up. These women were "fast–tracked" into treatment.	Pregnancy outcomes; Time–to–treatment initiation; Gestational age at ARV treatment initiation; Time from ARV treatment initiation to childbirth; Time between HIV diagnosis and receiving CD4 cell count results.
Ong'ech 2012 [38]	PMTCT	Kenya (363)	Prospective cohort study	Early infant HIV testing and prophylaxis were provided in the Maternal and Child Health clinic.	Infants were escorted to the Comprehensive Care Clinic, within the same health facility, for all HIV–related services.	Rates of attendance at each study visit (9 and 12 mo) and receipt of services for: infant HIV testing and prophylaxis at 6–8 weeks, receipt of immunizations at 14 weeks, continuation of prophylaxis at 6 mo, measles immunization at 9 mo, and HIV antibody testing at 12 mo.

Table 1. Continued

Study	Services integrated	Setting (participants)	Study design	Intervention description	Control description	Outcome measures
Pfeiffer 2010 [45]	ART	Mozambique (unknown)	Retrospective cohort study	At integrated sites, HIV–positive women were referred to the ART clinic from ANC services within the same health unit.	At vertical sites, HIV–positive women were referred to the ART clinic from ANC services at other health units.	Loss to follow–up from referrals of HIV–positive women from PMTCT services to ART services.
Stitson 2010 [42]	ART	South Africa (14987)	Retrospective cohort study	Site 1: ART initiated within the antenatal clinic when obstetricians with an HIV specialisation were on site. Site 2: women were referred by letter to a separate ART service located within 100 m of the maternity unit on the same premises.	Eligible women at the ANC clinic were referred to another site for HIV counselling and opt–in testing. ART was delivered at a separate primary health care facility approximately three kilometres from the antenatal service, using a referral letter.	Proportion of women who received more than 8 weeks of HAART; initiation of HAART in pregnancy.
Stinson 2013 [41]	ART	South Africa (14617)	Retrospective cohort study.	See Stinson 2010.	See Stinson 2010.	Proportion of women who initiated ART before delivery; Time to treatment initiation.
Turan 2012 [36]	ART, PMTCT	Kenya (1123)	Cluster–RCT	At the fully integrated sites, HIV positive women were provided all ANC, PMTCT, and HIV services in the ANC clinic, including HAART for women who were eligible.	In the control (non–integrated) clinics ANC and basic PMTCT services were provided in one visit, with referral to a separate clinic in the same health facility for HIV care and treatment (including HAART if indicated, opportunistic infection prophylaxis, education, and adherence counselling).	Baseline data only (aims to report HIV–free infant survival at 6 mo; rates of maternal enrolment in HIV care and treatment; infant HIV testing uptake at 3 mo).
Vo 2012 [35] (substudy of Turan 2012 [36])	ART	Kenya (326)	Nested cross–sectional study	See Turan 2012	See Turan 2012	Satisfaction; Preferred service model; average wait times.
Winestone 2012 [37] (substudy to Turan 2012 [36]	ART, PMTCT	Kenya (36 providers)	Qualitative study	See Turan 2012	See Turan 2012	Provider perceptions of quality of care.
Munkhuu 2009 [39]	Congenital syphilis testing	Mongolia (7700)	Cluster–RCT	The one–stop service included: (i) on–site screening for syphilis using rapid syphilis tests at the first antenatal visit and at the third trimester of gestation; (ii) immediate on–site treatment for seropositive women and their sexual partners; and (iii) pre– and post–test counselling.	After being admitted to the antenatal clinic, a pregnant woman could visit any District General Hospital or the National Center of Infectious Diseases for free initial and confirmatory syphilis testing. Women testing positive would be sent to a venereologist for appropriate case management and follow–up control, including contact tracing and counselling.	Uptake of syphilis testing at the first visit and third trimester; Receipt of adequate treatment (ie, completion of 3 doses of treatment before delivery); Treatment rates for sexual partners.
Bronzan 2007 [43]	Congenital syphilis testing	South Africa (1250)	Non–randomised controlled trial	On–site antenatal syphilis screening	Off–site syphilis screening	Percentage of eligible women who received 1, 2, or 3 appropriately timed weekly doses of penicillin; Acceptability of onsite testing to nurse clinicians.
Rahman 2011 [48]	Various	Bangladesh (20766)	Controlled before–and–after study	Set of maternal and neonatal interventions, following the continuum of care approach from pregnancy to delivery to the postnatal period, with improved links between community– and facility–based service delivery modes.	In the control areas, women receive pregnancy, delivery, and post–natal care from various government health facilities.	Perinatal mortality; Rates of facility deliveries and caesarean section.

ANC – Antenatal care; ART – Antiretroviral therapy; ARV – Antiretroviral; HAART – Highly active antiretroviral therapy; PMTCT – Prevention of mother–to–child transmission

Only two of the included studies involved randomised controlled trials, in both cases with cluster randomisation at the level of the health care facility [36,49]. We furthermore included one non–randomised controlled trial [43], one stepped–wedge cluster non–randomised trial [46], two controlled before–and–after studies [44,48], one prospective [38] (1) and two retrospective [41,42,45] cohort studies, and three historically controlled trials [39,40,47]. For one of the included cluster–RCTs only baseline data were available at the time of the review [36], however, additional data on patient satisfaction with and provider's perception of the intervention were published separately in a cross–sectional study [35] and as a purely qualitative study [37].

Description of interventions

Nine of the 12 included studies focused on integration of HIV–related services with ANC. Of these, four studies focused exclusively on integration of antiretroviral therapy (ART) for HIV–infected pregnant women with ANC services [40,41,45,46], four on measures for PMTCT of HIV infection [38,39,44,49], and one on HIV care and treatment services for both mother and child [35–37]. Additionally, two studies discussed the integration of syphilis screening and treatment services with ANC [43,49]. Only one study described the integration of services during the postnatal care period with ANC services [48]. All of the included studies described integration primarily from the perspective of delivery of services. While the necessity for integration of other health system functions was briefly touched upon in the study by Pfeiffer and others [45], this was not described as part of the intervention.

In the included studies, integrated delivery of services generally entailed delivery of multiple services by the same health care provider or by an integrated care team, with all services provided either within the ANC clinic or otherwise within the same premises as the ANC clinic. However, in one study [40], only HIV testing and counselling were fully integrated within the ANC service, whereas initiation and follow–up of treatment for HIV–infected women were performed at a separate facility. In the comparison groups, similar services were usually provided as stand–alone services either within the same facility as the ANC clinic or at a nearby health facility. These services could be accessed by referral from the ANC clinic.

Quality assessment and risk of bias

Of the included studies, only two met the 'gold standard' of evidence offered by the RCT design. All other studies used designs that are generally considered more prone to bias and confounding. The risk of bias for six studies (including the two RCTs, two CBA studies, one NRCT and one stepped–wedge trial) was assessed using the EPOC criteria. Only one of the RCTs described a random method of allocation and reported blinding of the study investigators [36]. The other RCT provided scant methodological detail and the study protocol was not available [49]. Similarly, the two CBA studies [44,48], as well as the NRCT [43] also did not report sufficient methodological information to assess risk of bias. **Table 2** provides a summary of the risk of bias assessment (using EPOC criteria) of included RCTs, SWTs, CBAs and NRCTs.

For the remaining six studies, the risk of bias was rated against the three categories of the Newcastle–Ottawa scale. **Table 3** shows the risk of bias assessment for included NRS based on the Newcastle–Ottawa Scale. For the study group category, one cohort study [35] and two historically controlled trials [40,47] scored the maximum of four stars; two studies (presented in three papers) received three stars [39,41,42] and one study failed to provide information on all but one of these criteria, receiving 1 star [45]. For the group comparability category, no studies received two stars. For the outcome category, five studies [35–39,44] reported the use of routine clinic and programme records to collect data, which may be assumed secure; one study did not report its data source at all [45]. As all included NRS used uptake and utilisation of services during pregnancy as their primary outcome, the period of follow–up until delivery was considered sufficient for all seven studies. This also meant

Table 2. Risk of bias assessment (EPOC criteria) of included RCTs, SWTs, CBAs and NRCTs

	Munkhuu 2009 [49] (cRCT)	Turan 2012 [36] (cRCT)	Killam 2010 [46] (SWT)	Geelhoed 2013 [44] (CBA)	Rahman 2011 [48] (CBA)	Bronzan 2007 [43] (NRCT)
Sequence generation	U	L	H	L	N/A	N/A
Allocation concealment	U	U	H	U	N/A	N/A
Blinding	U	L	L	U	U	U
Complete outcome data	L	L	L	N/A	N/A	N/A
No selective outcome reporting	U	N/A	U	U	U	U
Group comparability	L	L	L	U	L	U
Protection against contamination	U	L	U	U	U	H
Free from other sources of bias	L	U	L	U	L	H

cRCT – cluster–randomized controlled trial, CBA – controlled before–and–after trial, NRCT – non–randomised controlled trial, SWT – stepped wedge trial, H – High risk, L – Low risk, U – Unclear, N/A – Not applicable

Table 3. Risk of bias assessment for included NRS based on the Newcastle–Ottawa scale

	Ong'ech 2012 [38]	Pfeiffer 2010 [45]	Stinson 2010, 2013 [41,42]	Kasenga 2009 [47] (HCT)	Van der Merwe 2006 [40] (HCT)	van't Hoog 2005 [39] (HCT)
Study group:						
Representativeness	★	–	★	★	★	★
Selection of control	★	★	–	★	★	★
Exposure	★	–	★	★	★	–
Baseline	★	–	★	★	★	★
Cohort comparability:	★	–	★	–	★	–
Outcome:						
Assessment methods	★	–	★	★	★	★
Follow–up	★	★	★	★	★	★
Loss–to–follow–up*	–	–	–	–	–	–
Total	**7 stars**	**2 stars**	**6 stars**	**6 stars**	**7 stars**	**5 stars**

*As all included NRS used uptake and utilisation of services during pregnancy as their primary outcome, no follow–up beyond the point of recorded uptake of services was reported. We therefore did not award any stars in this category.

that "loss to follow–up" was not applicable in most cases, as no follow–up beyond the point of recorded uptake of services was required. Hence, we did not award any stars in this category. Overall, one cohort study [35] and one HCT [40] scored seven or eight stars; three studies scored five or six stars [39,41,42,47] and one scored just two stars [45].

Uptake and utilisation of health services

Utilisation outcomes for studies that examined integration of HIV services were grouped into four main themes: uptake of counselling & testing, enrolment, treatment initiation and follow–up & attendance. **Figure 2** shows a Forest plot of uptake and utilisation of HIV services (integrated care vs controls) for the included studies.

Three studies reported outcomes related to uptake of testing and counselling [38,39,47], suggesting higher uptake of HIV testing in integrated clinics [39,47]. Treatment initiation was higher in integrated clinics: one of the studies which did not find an effect had a very small sample size [47], and more recent outcomes from the same study as Stinson 2010 reported positive effects [41]. Effect on uptake of services and treatment initiation could not be estimated in Ong'ech and others, as all PCR testing and co–trimaxazole initiation was complete in both intervention and control groups. In the CBA study [44,48], there was an improvement in follow–up of HIV–exposed infants (registration, follow–up visits, serological testing) in both groups, but the progress could not be attributed to integrated MCH services and difference–in–difference estimates were not provided. Only one study reported on uptake of other services (immunisations for HIV infected infants) and follow–up care (attendance at PNC appointments, and continuation with prophylaxis), and suggested that integrated HIV services improved continuity of care for HIV infected infants [38].

For HIV–services, three studies reported on timeliness of treatment initiation or treatment duration at delivery [40,41,46]. Time to receiving test results and time to treatment initiation were shorter in integrated delivery models than in control groups in all three studies. Duration of ART before delivery and gestational age at ART initiation were comparable across integrated and control service delivery models. **Table 4** summarises the findings from the included studies on the timeliness of treatment initiation.

Two studies reported uptake and utilisation of services after integration of syphilis screening to ANC services [43,49]. Syphilis screening coverage was universal in the integrated model at the first antenatal visit, and was still significantly higher during the third trimester as compared with the control group; therefore, case detection was also higher in the intervention group. Appropriate treatment for patients with syphilis and their partners also improved in the integrated care delivery models. **Figure 3** shows a Forest plot of the results of uptake and utilization of syphilis screening services (integrated care vs controls) for the included studies.

Only one study reported outcomes relevant to integrating ANC to PNC; however, the study examined a multifaceted service delivery intervention involving strengthening both community and facility based care, as well as implementing evidence–based care [48]. While ANC coverage, facility delivery, and caesarean section rates were significantly higher in the post intervention period, the progress may not be attributable to the intervention.

Health outcomes

Three studies reported health outcomes (**Figure 4** shows a Forest plot of health outcomes, as measured by odds of adverse health outcomes in integrated care vs controls) [40,48,49]. The results were not pooled due to heterogeneity in type of service integration. One study found that both stillbirths and neonatal deaths were lower in regions where an integrated package of strengthened ANC and

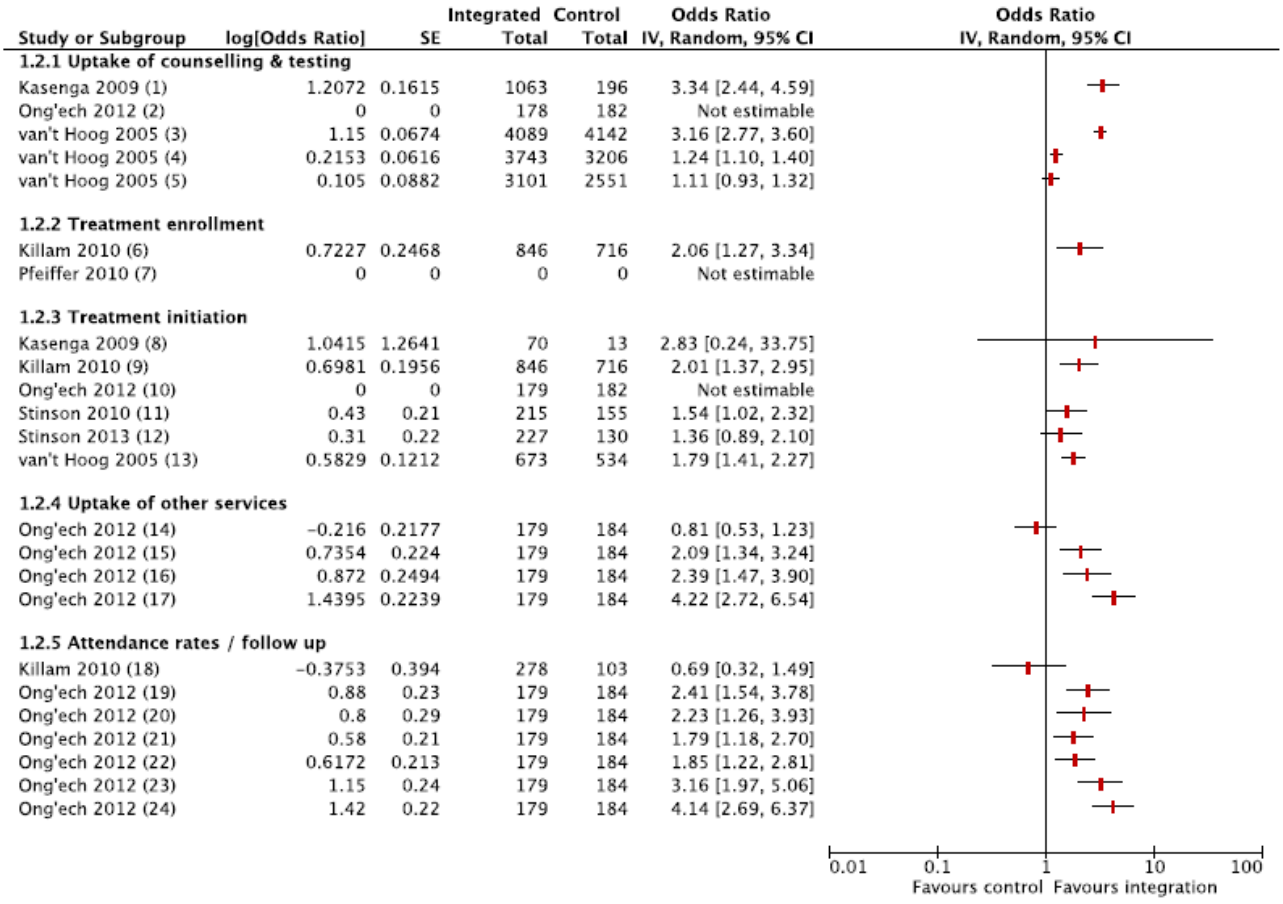


Figure 2. Uptake and utilisation of HIV services (integrated care vs controls). (1) HIV testing within ANC; (2) Infant DBS–PCR testing at 6–8 weeks; (3) Pre–test counselling; (4) Post–test counselling; (5) HIV testing within ANC; (6) Enrollment to HIV–care within 60 days of diagnosis; (7) Women registered for HIV care <30 days post–test (missing data, contact); (8) Nevirapine at delivery; (9) ART initiation during pregnancy; (10) Infant CTX initiation at 6–8 weeks (100% success in intervention group); (11) ART; (12) HAART; (13) Nevirapine uptake; (14) Measles immunization at 9 months; (15) Oral polio vaccine at 14 weeks nths; (16) Complete vaccination by 12 months; (17) DPT vaccine at 14 weeks; (18) 90–day retention among patients initiating ART; (19) 9–month postnatal visit; (20) 6–month postnatal visit; (21) Continuation of CTX prophylaxis at 6 months; (22) 14–week postnatal visit, (23) 12–month postnatal visit; (24) HIV antibody test at 12 months.

PNC services was delivered by community health workers, as compared with usual government care, and the adjusted odds ratio (OR) for perinatal deaths in intervention settings was 0.74 (95% confidence interval (CI) 0.62–0.88) [48]. The numbers of HIV–infected infants born to HIV+ mothers and those with congenital syphilis also were lower where testing and counselling were integrated to ANC services [40,49].

User experience

Data on user experience with and preferences regarding integrated care were collected in one sub–study of a cluster randomised trial of HIV–integrated services [35,36]. In adjusted models, overall user satisfaction with care was associated with a preference for integrated services (odds ratio, OR=2.03, 95% CI 1.07–3.85), and attending an integrated clinic (OR 10.34, 95% CI 2.08–51.3). Interactions between HIV status and integration suggest that integration improved HIV–infected women's satisfaction with their overall clinic experience, while it did not have an effect on HIV–uninfected women [35]. One study reported on the satisfaction of caretakers for HIV–infected infants in intervention and control groups, but did not provide any data [38]. At the end of one year of follow–up, there was no difference in satisfaction with the integrated vs usual care models.

Two studies reported on user satisfaction for the intervention groups only [44,49]. For one–stop integrated MCH services for HIV–infected infants, health care providers reported high satisfaction and "a subjective feeling of increased effectiveness" [44]. Over 86% of women attending two antenatal clinics in Ulaanbaatar, Mongolia, strongly agreed or agreed that they preferred receiving syphilis testing in the same place as ANC, allowing them to get same–

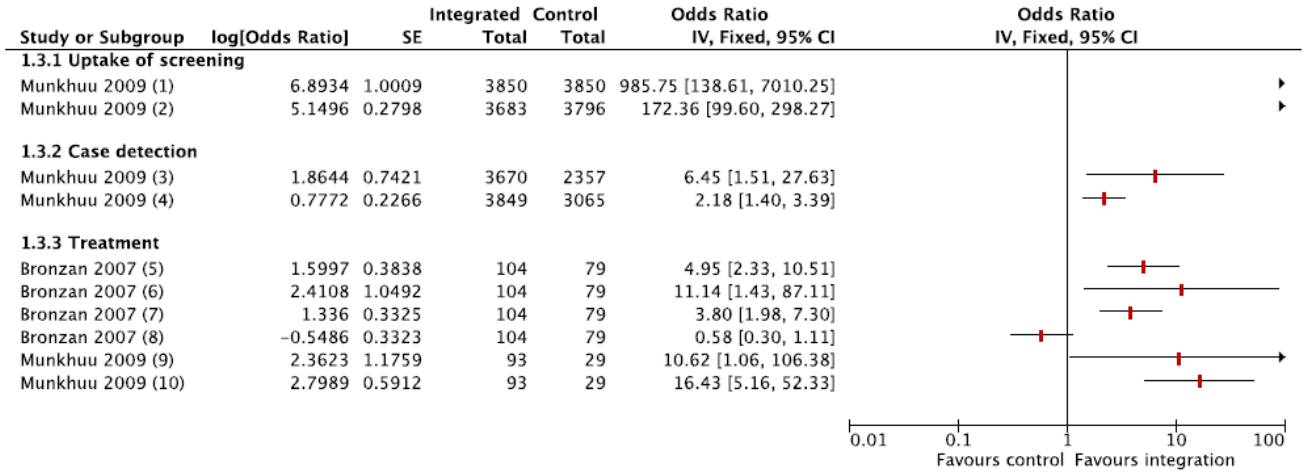


Figure 3. Uptake and utilization of syphilis screening services (integrated care vs controls). (1) Coverage at 1st antenatal visit; (2) Coverage at 3rd trimester; (3) Cases at 1st antenatal visit; (4) Cases at 3rd trimester; (5) At least one appropriately timed penicillin dose/week; (6) One appropriately timed penicillin dose/week; (7) Two appropriately timed penicillin doses/week; (8) Three appropriately timed penicillin doses/week; (9) Adequate treatment; (10) Partner treatment.

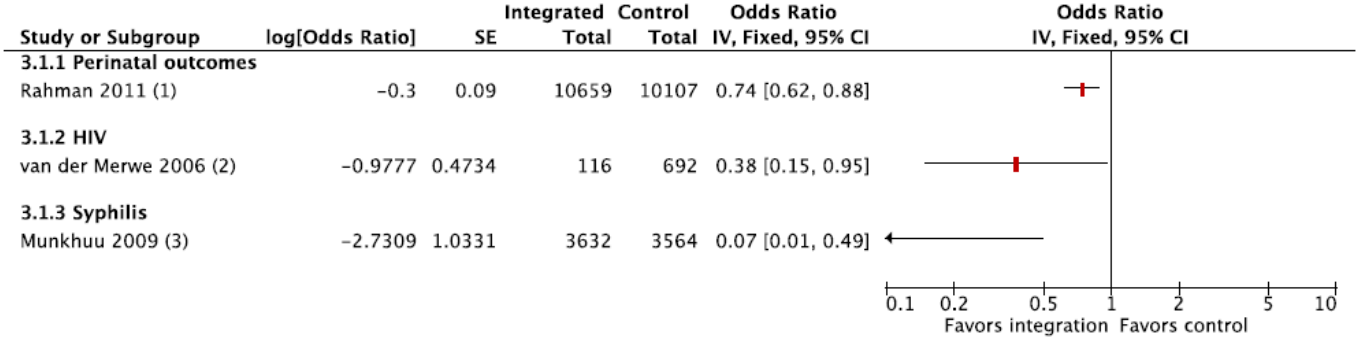


Figure 4. Health outcomes (odds of adverse health outcomes in integrated care vs controls). (1) Perinatal mortality, adjusted; (2) Number of HIV infections among infants born to HIV+ mothers; (3) Number of congenital syphilis cases.

day results and receive counselling and treatment from ANC providers. 80% were satisfied with the one–stop service, but 38% found the rapid testing stressful and less confidential. Most providers were also satisfied with integrated services, not reporting any significant problems or that syphilis counselling and treatment interfered with routine antenatal care [49]. Providers report, however, that integrated services lead to high staff workloads [44,49].

DISCUSSION

We found 12 studies that compared delivery of health services integrated into ANC with other, non–integrated, models of delivery of the same set of services. Our review finds some, albeit limited, evidence that integrated delivery results in improved uptake and utilisation of these services. Increased uptake of testing (HIV and syphilis) and PMTCT services, and earlier initiation of ART for HIV–infected mothers were, in turn, associated with lower rates of congenital infection with HIV and syphilis. In general, women also reported improved satisfaction with integrated services. These findings support the view that integrating additional health services into ANC can result in improved access to and uptake of essential health services for pregnant women. However, the reported evidence is largely based on non–randomised studies with moderate– to high–risk of bias, and therefore should be interpreted with caution.

Overall completeness and applicability of evidence

This review adds to a growing body of literature on integration of specific services into antenatal care settings, such

Table 4. Timeliness of treatment initiation

	Measure	Integrated	Control	P–value
Duration of ART before delivery (weeks):				
Killam 2010 [46]	Mean (SD)	10 (N/A)	11 (N/A)	NS
van der Merwe 2006 [40]	Median (IQR)	7 (3.9–11.2)	5 (2–10)	NS
Gestational age at ART initiation (weeks):				
Killam 2010 [46]	Mean (SD)	22 (N/A)	22 (N/A)	NS
van der Merwe 2006 [40]	Median (IQR)	32 (28–35)	33.5 (31–36)	0.042
Stinson 2013 [41]	Median (IQR)	31 (28–34)	30 (27–34)	NS
Time to receiving CD4 cell count (days):				
van der Merwe 2006 [40]	Median (IQR)	29 (11.5–45)	50 (22–92)	0.047
Time to treatment initiation (days):				
Stinson 2013 [41]	Median (IQR)	36 (N/A)	59 (N/A)	<0.001
van der Merwe 2006 [40]	Median (IQR)	37 (22–63)	56 (30–103)	0.041

SD – standard deviation, IQR – interquartile range, N/A – not applicable, NS – not significant

as PMTCT [11,19,26] and HIV services [25]. Of special interest is the review by Tudor Car [24], which looks at the effect of integration of perinatal PMTCT interventions aimed at reducing MTCT of HIV. It bases its findings on five studies, of which four were included in this review. It found that "there is very limited, non–generalisable evidence of improved PMTCT intervention uptake in integrated PMTCT programmes." A separate review by Lindegren and others looked at the impact of integrating HIV services with Maternal, Neonatal and Child Health (MNCH) services [25]. The focus of the review by Lindgren and others different from ours in that it looked at integration of HIV services into ANC, but also considered the reverse (ie, integration of ANC services into HIV services), or integration of both types of services into a pre–existing set of services. Across these different forms of integration, Lindegren and others found that for most studies integration had an apparent positive impact on reported outcomes. Several studies included in the review by Lindegren and others reported mixed or no effects, and one study reported negative outcomes due to providing integrated services [25]. These findings are generally consistent with those reported in our review.

Strikingly, the large majority of studies (nine out of 12) we retrieved concerned the integration of HIV–related services, in particular PMTCT and ART, into ANC. Two other studies dealt with integration of syphilis screening into ANC. However, we found no studies on integration of, for example, screening and treatment for other STIs, tuberculosis, malaria, non–communicable diseases or mental health issues into ANC that met the inclusion criteria. Whilst this emphasis on HIV is perhaps understandable in the context of countries with a high burden of HIV, this review reveals that there are few studies that have explored the potential of using ANC contacts as an entry point for health care services for women. This apparent deficiency was previously also addressed by Kerber and others, who noted that even in countries with good coverage of ANC services, coverage of effective interventions such as PMTCT remains low [5]. Since ANC often represents the most important, if not the only, point of contact a woman in LMIC has with formal health care services, our findings demonstrate lost opportunities for providing essential preventive and curative services.

Furthermore, the almost complete absence of studies looking at the potential benefits of integrating PNC services with ANC underscores the insufficient attention given to PNC in general, and suggests continued fragmentation of the continuum of maternal and child health care, particularly in the crucial post–partum period. As Kerber and others remarked, this fragmentation of the continuum suggests a "consensus has not been reached on a minimum package of postnatal interventions, with the strategies and mix of skills that are necessary for delivery." [5] This is a critical shortcoming that urgently needs to be addressed.

Only two of the included studies explicitly addressed the potential drawbacks of service integration and its impact on service quality, noting that integrated delivery of services could theoretically lead to inadvertent disclosure of HIV status as HIV–infected women would require longer appointments than non–infected women [37], and could result in unnecessary treatment if the new service model requires easier–to–use but less accurate testing techniques [49]. One study found that nurses considered the impact of integration on their workload acceptable [44]; no other impacts on the health system or other health services were discussed. This limited attention to the impact of integration on service quality and on the wider health system is cause for some concern. Decisions on whether or not to integrate specific services should be based on system–wide consideration of *all* potential costs and consequences, including unintended ones. However, the studies included in the review did not estimate costs and economic consequences of integration.

Potential biases and limitations

This review has four main limitations. First, although we used a robust and tested search strategy, it is nonetheless possible that we missed relevant studies. However, comparison with other reviews with a similar scope (ie, integration of services into maternal and child health care) [11,24,25,50], validates our strategy as we retrieved all relevant titles cited there.

Second, we were unable to retrieve the full text for 23 publications that we considered potentially eligible based on their titles and, where available, abstracts. Many of these were published in national or regional journals, often in languages other than English. Whilst this may have skewed our findings towards studies set in Anglophone countries and those published by European and North American researchers, it should be noted that out of the 23 missing studies only three were published from 2000 onwards. By comparison, all included studies were published in 2005 or later. We therefore consider it unlikely that many of the missing studies would have been eligible for inclusion, or that this could have had a significant effect on our overall findings.

Third, a potentially more important source of information not reported here is formed by programme evaluations that have not been published in the peer–reviewed literature, but have been prepared by funding institutions and implementing organisations. These additional data are included in a separate publication [51], which more generally discusses barriers and enablers to integration of services into ANC.

Fourth, as our review focused specifically on the impact of a service delivery model in which services were integrated into ANC, we required studies to compare findings to a service model in which the same, or a similar, set of services was provided in a non–integrated fashion. Without such a comparison it would not have been possible to distinguish between outcomes due to the availability of the services themselves, and those related to their mode of delivery. As a result, we excluded studies in which services that had not been previously available were directly introduced into the ANC setting. This applied in particular to PMTCT services. Also studies that did not clearly describe whether services had been previously available or, if so, how these were delivered, had to be excluded. This limited our evidence base to studies that very explicitly compared service delivery models, despite the fact that others also discussed similar integrated services.

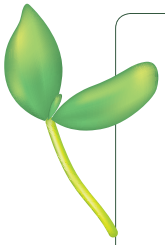
CONCLUSIONS

Implications for policy and practice

This review highlights the potential for improving maternal and child health care by integrating additional services with antenatal care, capitalising on the opportunities presented by relatively high rates of ANC coverage in many LMICs to develop integrated, evidence–based and cost–effective interventions with common delivery strategies for target populations [5]. The content and complexity of such a service package should be informed by the local health system capacity and epidemiological context and can evolve over time. However, care should also be taken to minimise the risks involved, such as potential deterioration of service quality and patient satisfaction, or overburdening frontline health workers.

Implications for research

There is a large evidence gap on the possible impacts for uptake and utilisation of essential services and health outcomes from integration of services with ANC. What little evidence is available is of insufficient quality to allow formulation of policy recommendations for other LMICs that may benefit from integration of health services. There is a clear need for more rigorously conducted studies, ideally involving comparison between different service delivery models with random allocation. However, additional quasi–experimental studies, and demonstration projects complemented by modelling studies, could also provide valuable insights in this area and in particular should help in understanding the role of contextual factors in achieving specific outcomes.



Funding: This research was funded by the Bill & Melinda Gates Foundation and the Harvard School of Public Health. The funders had no role in the study design, or in the interpretation of the results.

Authorship declaration: TdJ and IGU were responsible for development of the methodology, all stages of data collection and analysis, and reporting. EA and JZ contributed to data collection, screening of papers and data extraction. RA was responsible for the conception of the study and contributed to drafting and finalisation of the manuscript. All authors take responsibility for this study and its findings.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare no conflict of interest.

REFERENCES

1 Global Health Observatory Data Repository WHO. Maternal and reproductive health [Internet]. Geneva, Switzerland: WHO, 2013. Available: http://apps.who.int/gho/data/node.main.REPWOMEN39?lang=en. Accessed: 11 June 2014.

2 WHO Maternal Mortality Fact Sheet N°348. Geneva, Switzerland: WHO, 2012. Available: http://www.who.int/mediacentre/factsheets/fs348/en/. Accessed: 11 June 2014.

3 Trends WHO. in Maternal Mortality: 1990 to 2013 [Internet]. Geneva, Switzerland; 2014. Available: http://www.who.int/reproductivehealth/publications/monitoring/maternal-mortality-2013/en/. Accessed: 11 June 2014.

4 Khan KS, Wojdyla D, Say L, Gülmezoglu AM, Van Look PFA. WHO analysis of causes of maternal death: a systematic review. Lancet. 2006;367:1066-74. Medline:16581405 doi:10.1016/S0140-6736(06)68397-9

5 Kerber KJ, de Graft-Johnson JE. Bhutta Z a, Okong P, Starrs A, Lawn JE. Continuum of care for maternal, newborn, and child health: from slogan to service delivery. Lancet. 2007;370:1358-69. Medline:17933651 doi:10.1016/S0140-6736(07)61578-5

6 Children WHO. reducing mortality. Fact Sheet N°178. Geneva, Switzerland: WHO, 2013. Available at: http://www.who.int/mediacentre/factsheets/fs178/en/#. Accessed: 11 June 2014.

7 Tinker A, ten Hoope-Bender P, Azfar S, Bustreo F, Bell R. A continuum of care to save newborn lives. Lancet. 2005;365:822-5. Medline:15752509 doi:10.1016/S0140-6736(05)71016-3

8 Warren C, Daly P, Toure L, Mongi P. Postnatal care. In: Lawn J, Kerber K, editors. Opportunities for Africa's newborns: practical data, policy and programmatic support for newborn care in Africa. Geneva, Switzerland: WHO, 2006. p.79–90. Available: http://www.who.int/pmnch/media/publications/africanewborns/en/. Accessed: 11 June 2014.

9 Lincetto O, Mothebesoane-Anoh S, Gomez P, Munjanja S. Antenatal Care. In: Lawn J, Kerber K, editors. Opportunities for Africa's newborns: practical data, policy, and programmatic support for newborn care in Africa. Geneva, Switzerland: WHO, 2006. p. 51–62. Available: http://www.who.int/pmnch/media/publications/africanewborns/en/. Accessed: 11 June 2014.

10 USAID. ACCESS Program update: focused antenatal care–achieving results in antenatal care: improving maternal and newborn outcomes through integration of Services. Washington, DC, 2008. Available: http://www.jhpiego.org/files/ACCESS_resbriefANC_ENjul2008.pdf. Accessed: 11 June 2014.

11 Tudor Car L, van Velthoven M, Brusamento S, Elmoniry H, Car J, Majeed A, et al. Integrating prevention of mother-to-child HIV transmission (PMTCT) programmes with other health services for preventing HIV infection and improving HIV outcomes in developing countries [Review]. Cochrane Database Syst Rev. 2011;6:CD008741. Medline:21678382

12 Atun R, de Jongh T, Secci F, Ohiri K, Adeyi O. Integration of targeted health interventions into health systems: a conceptual framework for analysis. Health Policy Plan. 2010;25:104-11. Medline:19917651 doi:10.1093/heapol/czp055

13 Shigayeva A, Atun R, McKee M, Coker R. Health systems, communicable diseases and integration. Health Policy Plan. 2010;25 Suppl 1:i4-20. Medline:20966108 doi:10.1093/heapol/czq060

14 Mills A. Vertical vs horizontal health programmes in Africa: idealism, pragmatism, resources and efficiency. Soc Sci Med. 1983;17:1971-81. Medline:6670002 doi:10.1016/0277-9536(83)90137-5

15 Criel B, De Brouwere V, Dugas S. Integration of vertical programmes in multi-function health services. Stud Heal Serv Organ Policy. 1973;3. Antwerp, Belgium: ITGPress. Available: http://www.itg.be/itg/generalsite/infservices/downloads/shsop03.pdf. Accessed: 11 June 2014.

16 Atun R, Lazarus JV, Van Damme W, Coker R. Interactions between critical health system functions and HIV/AIDS, tuberculosis and malaria programmes. Health Policy Plan. 2010;25 Suppl 1:i1-3. Medline:20966104 doi:10.1093/heapol/czq062

17 Gröne O, Garcia–Barbero M. Integrated care: A position paper of the WHO European office for integrated health care services. Int J Integr Care. 2001;1:e21. Medline:16896400

18 Yartey J, Kumoji K. Technical consultation on the integration of HIV interventions into maternal, newborn and child health services. Geneva, Switzerland: WHO, 2006. Available: http://www.who.int/maternal_child_adolescent/documents/hiv_interventions/en/. Accessed: 11 June 2014.

19 Tudor Car L, Brusamento S, Elmoniry H, van Velthoven MHMMT, Pape UJ, Welch V, et al. The uptake of integrated perinatal prevention of mother–to–child HIV transmission programs in low– and middle–income countries: a systematic review. PLoS ONE. 2013;8:e56550. Medline:23483887 doi:10.1371/journal.pone.0056550

20 WHO. Making health systems work: integrated health services–what and why? Technical Brief No: 1. Geneva, Switzerland: WHO, 2008. Available: http://www.who.int/healthsystems/technical_brief_final.pdf. Accessed: 11 June 2014.

21 Samb B, Evans T, Dybul M, Atun R, Moatti J-P, Nishtar S, et al. An assessment of interactions between global health initiatives and country health systems. Lancet. 2009;373:2137-69. Medline:19541040 doi:10.1016/S0140-6736(09)60919-3

22 Atun R, Jaffar S, Nishtar S, Knaul FM, Barreto ML, Nyirenda M, et al. Improving responsiveness of health systems to non-communicable diseases. Lancet. 2013;381:690-7. Medline:23410609 doi:10.1016/S0140-6736(13)60063-X

23 Manzi F, Schellenberg JA, Hutton G, Wyss K, Mbuya C, Shirima K, et al. Human resources for health care delivery in Tanzania: a multifaceted problem. Hum Resour Health. 2012;10:3. Medline:22357353

REFERENCES

24 Tudor Car L, Van Velthoven MHMMT, Brusamento S, Elmoniry H, Car J, Majeed A, et al. Integrating prevention of mother-to-child HIV transmission programs to improve uptake: a systematic review. PLoS ONE. 2012;7:e35268. Medline:22558134 doi:10.1371/journal.pone.0035268

25 Lindegren ML, Kennedy C, Bain-Brickley D, Azman H, Creanga A, Butler L, et al. Integration of HIV/AIDS services with maternal, neonatal and child health, nutrition, and family planning services. Cochrane Database Syst Rev. 2012;9:CD010119. Medline:22972150

26 Both JMC, van Roosmalen J. The impact of Prevention of Mother to Child Transmission (PMTCT) programmes on maternal health care in resource-poor settings: looking beyond the PMTCT programme–a systematic review. BJOG. 2010;117:1444-50. Medline:20937071 doi:10.1111/j.1471-0528.2010.02692.x

27 Legido-Quigley H, Montgomery CM, Khan P, Atun R, Fakoya A, Getahun H, et al. Integrating tuberculosis and HIV services in low- and middle-income countries: a systematic review. Trop Med Int Health. 2013;18:199-211. Medline:23217030 doi:10.1111/tmi.12029

28 WHO. Sexual and reproductive health and HIV linkages: evidence review and recommendations. Geneva, Switzerland: WHO; 2009. Available: http://www.who.int/reproductivehealth/publications/linkages/hiv_2009/en/. Accessed: 11 June 2014.

29 Atun R, de Jongh TE, Secci FV, Ohiri K, Adeyi O, Car J. Integration of priority population, health and nutrition interventions into health systems: systematic review. BMC Public Health. 2011;11:780. Medline:21985434 doi:10.1186/1471-2458-11-780

30 The Cochrane Collaboration. Cochrane handbook for systematic reviews of interventions Version 5.1.0. Higgins J, Green S, editors. London: The Cochrane Collaboration, 2011.

31 Dudley L, Garner P. Strategies for integrating primary health services in low– and middle–income countries at the point of delivery. Cochrane Database Syst Rev. 2011;7:CD003318. Medline:21735392

32 Effective Practice and Organisation of Care (EPOC). Suggested risk of bias criteria for EPOC reviews. EPOC Resources for review authors. Oslo: Norwegian Knowledge Centre for the Health Services; 2013. Available: http://epocoslo.cochrane.org/epoc-specific-resources-review-authors. Accessed: 11 June 2014.

33 Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses [Internet]. Ottawa Hospital Research Institute, 2014. Available: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed: 11 June 2014.

34 Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol. 2005;5:13. Medline:15840177 doi:10.1186/1471-2288-5-13

35 Vo BN, Cohen CR, Smith RM, Bukusi EA, Onono MA, Schwartz K, et al. Patient satisfaction with integrated HIV and antenatal care services in rural Kenya. AIDS Care. 2012;24:1442-7. Medline:22296261 doi:10.1080/09540121.2011.652357

36 Turan JM, Steinfeld RL, Onono M, Bukusi EA, Woods M, Shade SB, et al. The study of HIV and antenatal care integration in pregnancy in Kenya: design, methods, and baseline results of a cluster–randomized controlled trial. PLoS ONE. 2012;7:e44181. Medline:22970177 doi:10.1371/journal.pone.0044181

37 Winestone LE, Bukusi E, Cohen C. Acceptability and feasibility of integration of HIV care services into antenatal clinics in rural Kenya: a qualitative provider interview study. Glob Public Health. 2012;7:149-63. Medline:22043837 doi:10.1080/17441692.2011.621964

38 Ong'ech JO, Hoffman H, Kose J. Provision of services and ccare for HIV-exposed infants: a comparison of maternal and child health clinic and HIV comprehensive care clinic models. J Acquir Immune Defic Syndr. 2012;61:83-9. Medline:22592589 doi:10.1097/QAI.0b013e31825bd842

39 van't Hoog AH, Mbori-ngacha DA, Marum LH, Otieno JA, Misore AO, Nganga LW, et al. Preventing mother-to-child transmission of HIV in Western Kenya: operational issues. J Acquir Immune Defic Syndr. 2005;40:344-9. Medline:16249710 doi:10.1097/01.qai.0000160712.86580.ff

40 van der Merwe K, Chersich MF, Technau K, Umurungi Y, Conradie F, Coovadia A. Integration of antiretroviral treatment within antenatal care in Gauteng Province, South Africa. J Acquir Immune Defic Syndr. 2006;43:577-81. Medline:17031321

41 Stinson K, Jennings K, Myer L. Integration of antiretroviral therapy services into antenatal care increases treatment initiation during pregnancy: a cohort study. PLoS ONE. 2013;8:e63328. Medline:23696814 doi:10.1371/journal.pone.0063328

42 Stinson K, Boulle A, Coetzee D, Abrams E, Myer L. Initiation of highly active antiretroviral therapy among pregnant women in Cape Town, South Africa. Trop Med Int Health. 2010;15:825-32. Medline:20497405 doi:10.1111/j.1365-3156.2010.02538.x

43 Bronzan RN, Mwesigwa-Kayongo DC, Narkunas D, Schmid GP, Neilsen G, Ballard RC, et al. On-site rapid antenatal syphilis screening with an immunochromatographic strip improves case detection and treatment in rural South African clinics. Sex Transm Dis. 2007;34:S55-60. Medline:17139234 doi:10.1097/01.olq.0000245987.78067.0c

44 Geelhoed D, Lafort Y, Chissale É, Candrinho B, Degomme O. Integrated maternal and child health services in Mozambique: structural health system limitations overshadow its effect on follow-up of HIV–exposed infants. BMC Health Serv Res. 2013;13:207. Medline:23758816 doi:10.1186/1472-6963-13-207

45 Pfeiffer J, Montoya P. Integration of HIV/AIDS services into African primary health care: lessons learned for health system strengthening in Mozambique-a case study. J Int AIDS Soc. 2010;13:3. Medline:20180975 doi:10.1186/1758-2652-13-3

REFERENCES

46 Killam WP, Tambatamba BC, Chintu N, Rouse D, Stringer E, Bweupe M, et al. Antiretroviral therapy in antenatal care to increase treatment initiation in HIV–infected pregnant women: a stepped-wedge evaluation. AIDS. 2010;24:85-91. Medline:19809271 doi:10.1097/QAD.0b013e32833298be

47 Kasenga F, Byass P, Emmelin M, Hurtig A-K. The implications of policy changes on the uptake of a PMTCT programme in rural Malawi: first three years of experience. Glob Health Action. 2009;2:1-7. Medline:20027274

48 Rahman A, Moran A, Pervin J, Rahman A, Rahman M, Yeasmin S, et al. Effectiveness of an integrated approach to reduce perinatal mortality: recent experiences from Matlab, Bangladesh. BMC Public Health. 2011;11:914. Medline:22151276 doi:10.1186/1471-2458-11-914

49 Munkhuu B, Liabsuetrakul T, Chongsuvivatwong V, McNeil E, Janchiv R. One-stop service for antenatal syphilis screening and prevention of congenital syphilis in Ulaanbaatar, Mongolia: a cluster randomized trial. Sex Transm Dis. 2009;36:714-20. Medline:19773681 doi:10.1097/OLQ.0b013e3181bc0960

50 Suthar AB, Hoos D, Beqiri A, Lorenz-Dehne K, McClure C, Duncombe C. Integrating antiretroviral therapy into antenatal care and maternal and child health settings: a systematic review and meta-analysis. Bull World Health Organ. 2013;91:46-56. Medline:23397350 doi:10.2471/BLT.12.107003

51 de Jongh TE, Gurol-Urganci I, Allen E, Jiayue Zhu N, Atun R. Barriers and enablers to integrating maternal and child health services to antenatal care in low and middle income countries. BJOG. 2016;123:549-57. Medline:26861695. doi:10.1111/1471-0528.13898

Electronic supplementary material:
The online version of this article contains supplementary material.

journal of global health

PAPERS

Patterns and determinants of antenatal care utilization: analysis of national survey data in seven countdown countries

Ghada Saad–Haddad[1], Jocelyn DeJong[1], Nancy Terreri[2], María Clara Restrepo–Méndez[3], Jamie Perin[4], Lara Vaz[5], Holly Newby[6], Agbessi Amouzou[6], Aluísio JD Barros[3], Jennifer Bryce[7]

[1] Faculty of Health Sciences, American University of Beirut, Beirut, Lebanon
[2] Independent consultant
[3] Federal University of Pelotas, Pelotas, Brazil
[4] Institute for International Programs, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
[5] Save the Children, Washington D.C., USA
[6] Division of Data, Research and Policy, United Nations Children's Fund, New York City, NY, USA
[7] Institute for International Programs, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Background Antenatal care (ANC) is critical for improving maternal and newborn health. WHO recommends that pregnant women complete at least four ANC visits. Countdown and other global monitoring efforts track the proportions of women who receive one or more visits by a skilled provider (ANC1+) and four or more visits by any provider (ANC4+). This study investigates patterns of drop–off in use between ANC1+ and ANC4+, and explores inequalities in women's use of ANC services. It also identifies determinants of utilization and describes countries' ANC–related policies, and programs.

Methods We performed secondary analyses using Demographic Health Survey (DHS) data from seven Countdown countries: Bangladesh, Cambodia, Cameroon, Nepal, Peru, Senegal and Uganda. The descriptive analysis illustrates country variations in the frequency of visits by provider type, content, and by household wealth, women's education and type of residence. We conducted a multivariable analysis using a conceptual framework to identify determinants of ANC utilization. We collected contextual information from countries through a standard questionnaire completed by country–based informants.

Results Each country had a unique pattern of ANC utilization in terms of coverage, inequality and the extent to which predictors affected the frequency of visits. Nevertheless, common patterns arise. Women having four or more visits usually saw a skilled provider at least once, and received more evidence–based content interventions than women reporting fewer than four visits. A considerable proportion of women reporting four or more visits did not report receiving the essential interventions. Large disparities exist in ANC use by household wealth, women's education and residence area; and are wider for a larger number of visits. The multivariable analyses of two models in each country showed that determinants had different effects on the dependent variable in each model. Overall, strong predictors of ANC initiation and having a higher frequency (4+) of visits were woman's education and household wealth. Gestational age at first visit, birth rank and preceding birth interval were generally negatively associated with initiating visits and with having four or more visits. Information on country policies and programs were somewhat informative in understanding the utilization patterns across the countries, although timing of adoption and actual implementation make direct linkages impossible to verify.

Conclusion Secondary analyses provided a more detailed picture of ANC utilization patterns in the seven countries. While coverage levels differ by country and sub–groups, all countries can benefit from specific in–country assessments to properly identify the underserved women and the reasons behind low coverage and missed interventions. Overall, emphasis needs to be put on assessing the quality of care offered and identifying women's perception to the care as well as the barriers hindering utilization. Country policies and programs need to be reviewed, evaluated and/or implemented properly to ensure that women receive the recommended number of ANC visits with appropriate content, especially, poor and less educated women residing in rural areas.

Correspondence to:
Ghada Saad–Haddad
Faculty of Health Sciences
American University of Beirut
Beirut, Lebanon
gs29@aub.edu.lb

Antenatal care (ANC), defined as the care provided to a woman during her pregnancy, is an essential component of reproductive health care. ANC can serve as a platform for the delivery of highly–effective health interventions that can reduce preventable maternal and newborn deaths [1,2]. ANC services offer pregnant women an entry point to the health care system, providing appropriate screening, intervention and treatment throughout pregnancy, and encouraging women to seek a skilled birth attendant for their delivery [3]. Furthermore, using ANC allows women to receive information about improving maternal health through proper nutrition and self–care during pregnancy; and throughout the postpartum period, such as the benefits of exclusive breastfeeding and counseling on family planning methods [4].

The current World Health Organization (WHO) recommendation is that each woman receives a minimum of four goal–oriented or focused ANC visits for low–risk deliveries, to be supervised or attended by a skilled ANC attendant [4]. The timing of the first visit should be before 16 weeks of pregnancy, the second visit should be between 24 and 26 weeks, the third visit between 30 and 32 weeks, and the fourth visit between 36 and 38 weeks [5]. WHO defines a thorough set of essential elements for each visit (**Box 1**).

Coverage of ANC has been used globally as one of the indicators to track progress towards Target 5.B (achieving universal access to reproductive health by 2015) under Millennium Development Goal 5 (MDG 5) to improve maternal health [12]. The official ANC indicators for global tracking are: (1) the proportion of women with a recent live birth who report at least one ANC visit with skilled health personnel (ANC 1+); and (2) the proportion of women with a recent live birth who report at least four ANC visits with any provider (ANC 4+) [12].The Countdown to 2015 for Maternal, Newborn and Child Survival, a global movement that tracks coverage for evidence–based interventions in 75 countries that account for more than 95% of maternal and child deaths [13], also reports on the ANC 1+ and ANC 4+ indicators.

There have been numerous studies of the determinants of ANC use in low– and middle–income countries. Fewer studies have examined the determinants of use by frequency of antenatal care visits, comparatively, and through inferential analyses [14-21]. There have also been several analyses of equity in utilization of ANC services. Relevant articles stratify utilization by urban/rural place of residence [16,22–24], and less frequently, by mother's education [15,25], wealth [15,26], income [25], and ethnicity [25]. However, little is known about the frequency of ANC visits in general, especially as a comparative presentation across countries. No previous study, to our knowledge, has examined utilization in terms of what the globally measured ANC indicators might be missing with respect to associations between women's characteristics and their patterns of visits. Moreover, qualitative studies, or studies that use both qualitative and quantitative methods, are fewer in number [27]. These studies focus on contextual aspects such as the presence of health care workers in the community, availability of affordable care, household characteristics and per-

Box 1. The evolution of World Health Organization guidelines for antenatal care visits

The concept of antenatal care originated in Europe in the early decades of the 20th century. It is believed that the ANC model and the recommendations set at that time formed the foundation for ANC programs worldwide. The model indicated that visits should begin around 16 weeks of gestation, followed by visits at 24 and 28 weeks, then fortnightly visits until 36 weeks, and finally, weekly visits until delivery [6].

This 'Western model' was implemented for developing countries without taking into consideration contextual factors, which are especially important in low–resource settings [7]. WHO therefore developed a new model of ANC, consisting of a reduced number of visits and specifying the evidence–based interventions to be provided at each visit, including: assessment of the pregnant woman; screening for pre–eclampsia, anemia, syphilis, and HIV; provision of preventive measures such as checking of iron and folate dosage, tetanus toxoid immunization, anti–malarial precautions, and advice on labor or danger signs; advise on proper self–care, nutrition, and substance abuse; and counseling on the importance of family planning [5]. These recommendations are referred to as "focused" or "goal–oriented" ANC. Clinical evidence at the time the recommendations were released indicated that health outcomes were similar for women who received the four focused visits and women who received standard ANC with more visits [7,8].

Dowswell and colleagues [6] in an updated Cochrane systematic review using new methods of assessment, showed a statistically significant increase in perinatal mortality in low– and middle–income settings among women who received focused ANC compared to women who received standard ANC. In a 2011 statement, WHO acknowledged this and planned to provide updated guidelines for ANC based on their findings to be generated from additional secondary analyses [9]. The results of a secondary analysis looking at the WHO ANC trial were published in 2013, again showing a substantial increase in perinatal mortality among women receiving the focused ANC compared to those receiving the standard package, especially between 32 and 36 weeks of gestation. However, the findings also showed high levels of heterogeneity between the populations in the trials, and suggested that differences in perinatal mortality between the control and intervention groups could be attributed to different settings, populations or even quality of care received [10]. The WHO is re–evaluating its ANC guidelines, an exercise which is expected to be completed by the end of 2015 [11].

ceived distance from the health care facility, waiting time at the facility [27], women's perceptions about ANC, and their experiences, attitudes, beliefs and perceived need for services [27,28].

For this paper, we purposely selected a limited number of Countdown countries to examine and understand the underlying patterns of ANC utilization that are not revealed when relying solely on the globally measured ANC indicators. We identify whether a significant drop–off in utilization occurs after a certain number of visits. We also describe the number of ANC visits by the type of provider, and the content received overall during ANC. We examine the coverage of ANC by three measures of inequality. Finally, we use several environmental, population and individual characteristics to analyze utilization patterns in the selected countries. In addition, contextual information on policy and program structure was collected from the selected countries for the purpose of improving understanding of ANC coverage levels and drop off.

DATA AND METHODOLOGY

Data

The selection of countries was based on several criteria with the desire to have six to seven Countdown countries from different world regions, each with a Demographic Health Survey (DHS) in 2010 or later. We chose countries with extreme coverage levels (high or low) of ANC 4+, ANC 1+ and skilled birth attendance. We also selected a couple of countries identified as priority countries for eliminating mother–to–child transmission of HIV (list of the Countdown countries in Table S1 in **Online Supplementary Document**). The seven selected Countdown countries are Bangladesh, Cambodia, Cameroon, Nepal, Peru, Senegal and Uganda.

Data for our analyses were obtained from nationally–representative household surveys conducted under the DHS program [29]. Information on the number of ANC visits for the most recent live birth in the five years preceding the survey for each woman in the sample was found in the women's individual questionnaire. Women who responded "don't know" or had a missing response were excluded from the analysis. Missing variables found in responses to the other variables we chose for our descriptive and inferential statistics were handled similarly and observations were dropped from the analysis (proportions of missing varied between variables in each country and across countries but never exceed a proportion of 0.7% of the total sample size per country). The following DHS surveys were used for the analyses presented in this work: Bangladesh 2011, Cambodia 2010, Cameroon 2011, Nepal 2011, Peru 2012, Senegal 2010, Uganda 2011.

Methodology

The bulk of this study consisted of a thorough descriptive analysis to unpack the ANC indicators. We analyzed ANC visit frequency by type of provider reported and by content interventions received. In each survey, women were asked to list the providers they saw during any ANC visit. We categorized women who reported having ANC into those who saw a skilled provider for at least one visit and those who saw unskilled providers only (the list of providers in Table S2 in **Online Supplementary Document**). We selected a limited number of evidence–based content interventions that should be routinely administered during ANC visits and data usually available through DHS: blood sample taken, blood pressure taken, urine sample taken and being told about pregnancy problems. In Bangladesh, only being told about pregnancy problems was available, as the other questions were not asked. We examined differences in content received among women who saw a skilled provider vs those who saw only an unskilled provider, and the pattern of content received by wealth quintile. We charted the frequency of visits by gestational age at first ANC visit (by trimester). In Bangladesh, no data were collected on this variable. We also investigated inequalities in utilization of ANC visits by three dimensions of inequality, specifically, women's education (none, primary, secondary, and higher), household wealth quintile (five wealth quintiles from poorest to richest as defined by DHS) and the area of residence (rural or urban). We described the differences in proportions of women's reported frequency of visits by each of the three dimensions.

To identify the determinants that affected women's choices in initiating ANC and the seeking patterns among women who reported receiving ANC, we adopted Anderson's Behavioral Model for Healthcare Use [30], specifically four components of the model and a selection of 15 determinants (**Figure 1**). Anderson's Behavioral Model has been used extensively to understand utilization in different health care settings [32]. Numerous studies have made use of this conceptual model to study the determinants of antenatal care utilization [17,19,32,33]. We assessed the factors that influence the frequency of ANC visits for two comparisons: those reporting no visits vs those reporting one or more visits; and those reporting one to three visits vs those reporting four or more visits. After examining the bivariate relationships between each determinant and the dependent outcomes, we performed multivariable logistic regression analyses. Using the strategy of hierarchical entry of variables [34], we first included the external environment factors into the models to assess their association with the outcome variable. Using backward elimination we exclude factors not significant ($P<0.05$) at the level being entered, one at a time, starting with the variable with the

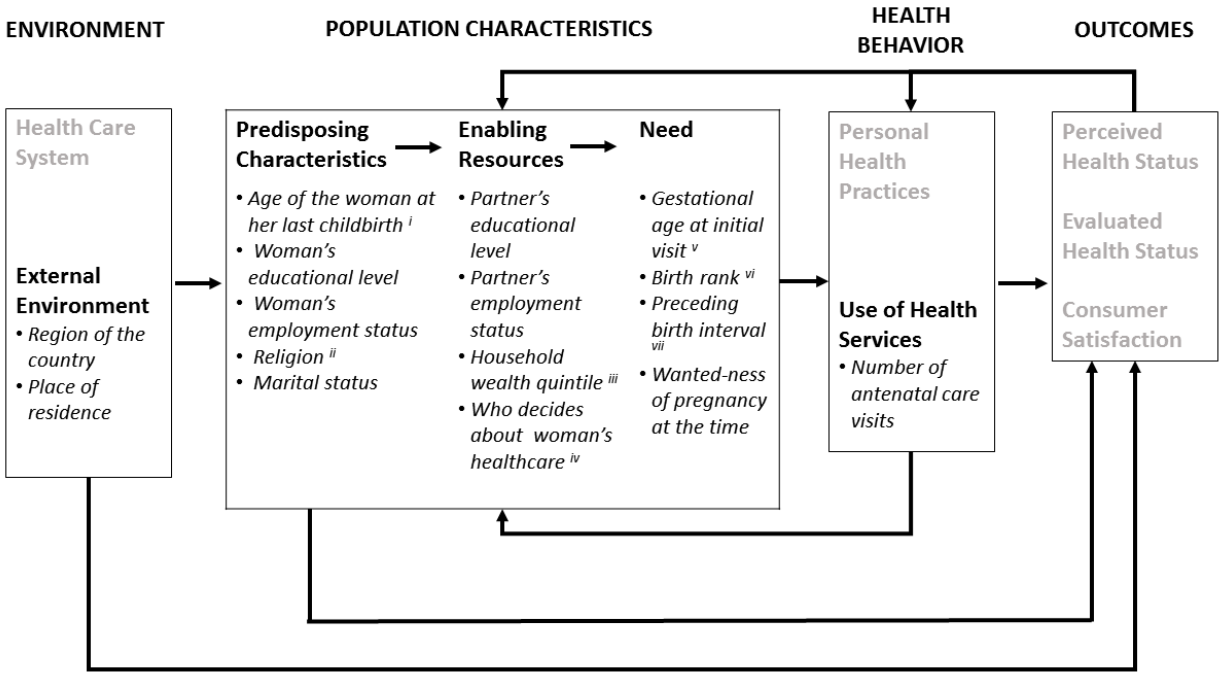


Figure 1. The conceptual framework based on Anderson's Behavioral Model of Healthcare Use and the corresponding determinants used in our secondary analysis [31]. Source: Anderson 1995 [31]. [i] Age at woman's most recent birth was calculated by subtracting the last child's date of birth from the woman's date of birth divided by 12. [ii]Religion was categorized as dominant religion and other religions. [iii]Household wealth quintile is made up of five wealth quintiles from poorest to richest as constructed by DHS where each quintile represents 20% of the households in the study sample. [iv]The variable 'who decides about woman's health care' is categorized as: woman alone, woman & partner, partner alone, someone else.[v]Gestational age at first ANC visit was grouped into trimesters. [vi]Birth rank was categorized as: 1st – 2nd birth, 3rd – 4th birth, 5th birth or more. [vii]Preceding birth interval was grouped into: first birth (no interval), less than 2 years interval, 2–3 years interval, more than 3 years interval.

highest *P*–value. We re–ran the models until all the variables at the level being entered were significant. After looking at the model with environmental variables only, we added the predisposing characteristics to the models followed by a reapplication of the backward elimination procedure. The enabling factors and the need factors were then added to the models using similar procedures.

We analyzed the data with Stata (StataCorp. 2013. Stata Statistical Software: Release 13; College Station, TX: StataCorp LP) using the 'svy' prefix to take into account the complex sample design, including sampling weights and clustering.

Collection of information on ANC–related national policies and programs

We collected information on policy and programming for ANC in each of the seven countries by identifying a government official, researcher and/or non–governmental organization (NGO) staff knowledgeable about current and past ANC policy and programs and willing to assist. Each key informant was encouraged to contact additional resource persons as necessary and to provide the research team with copies of relevant documents. We developed and provided informants with a self–administered standard. Topics covered in the questionnaire included information on country policies and guidelines, with details on recommended timing, number and content of ANC visits. Additional information requested included the locations of ANC service provision in each country, the types of providers, incentives for women to seek care, user fees, incentives for providers, communication or social marketing around ANC, and how the ANC service is organized. Most questions were open–ended, and a final question asked for any additional comments from the informant on how uptake of ANC services might be improved. To assist the informant in filling out the questionnaire and for verification purposes, each country questionnaire included pre–completed descriptive information from the latest country DHS or other publication, when available.

One member of the research team reviewed global policies relevant to ANC, compiled results from the country questionnaires, reviewed documents provided by key informants and others found through online searches. Follow–up with key informants was made to provide missing information or to resolve discrepancies. Descriptive summaries and tables were completed for each country along

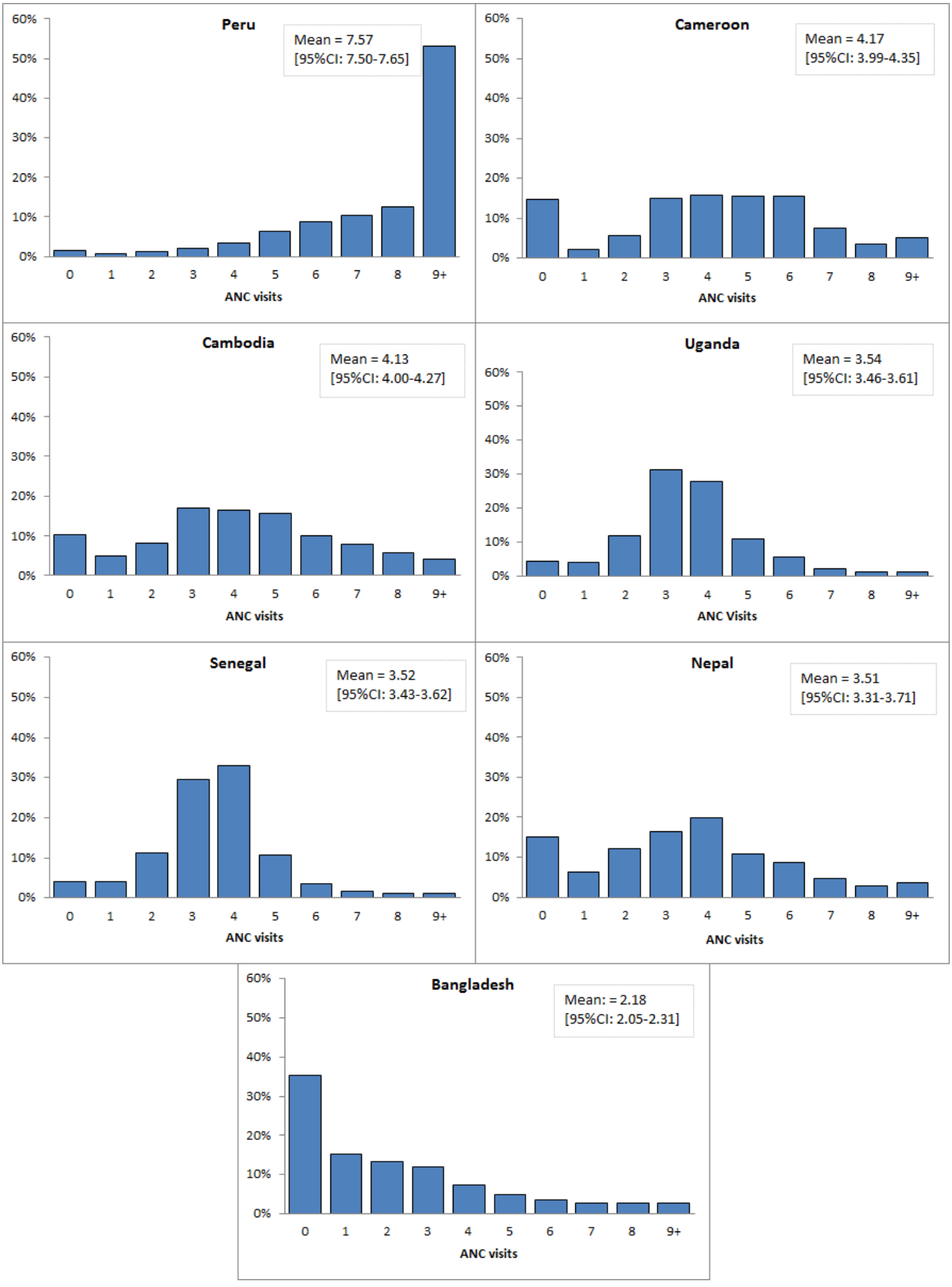


Figure 2. Percentage of women who had a live birth in the five years preceding the DHS surveys reporting zero to more than nine ANC visits for their most recent live birth, and mean of ANC visits among all these women (95% confidence intervals), in seven Countdown countries.

with an overall summary of findings. Results were shared with the research team and used in analyzing the country results.

RESULTS

The seven countries we selected for this secondary analysis are Bangladesh, Cambodia, Cameroon, Nepal, Peru, Senegal and Uganda. The un–weighted sample sizes of women aged 15–49 years in each country were: 7319 (ever–married women only) in Bangladesh; 6421 women in Cambodia; 7576 women in Cameroon; 4079 women in Nepal; 7991 women in Peru; 8008 women in Senegal and 4818 women in Uganda.

Descriptive analysis

Table 1 describes the proportions of women reporting one or more and four or more ANC visits with a skilled provider or any provider. With the exception of Bangladesh and Nepal, 85–96% of the women reported at least one ANC visit with a skilled provider. In Bangladesh, around one–third of the women reported not receiving any ANC. The proportion of women who reported four or more visits ranged from 48% to 63% in five of the seven countries. Bangladesh was at the low end, with 24% and Peru stood out with 94%. Despite the fact that the globally measured ANC indicators are not fully comparable, because the ANC 1+ indicator refers to visits with a skilled provider and ANC 4+ refers to visits with any provider, it is important to note that in five out of the seven countries, over 90% of the women who reported receiving four or more visits with any provider also reported receiving at least one visit with a skilled provider (**Table 1**).

The distribution of number of ANC visits varies from country to country, as shown in **Figure 2**. Peru has the highest mean (7.6); Cameroon and Cambodia have a mean of just over four reported ANC visits. The distributions in Bangladesh and Peru represent two extremes, with a right skewed distribution in Bangladesh (35% of women with no visits) and a left skewed distribution in Peru (nearly no women reporting zero ANC visits). Most women who reported no ANC visits reside in rural areas, are in the two poorest quintiles of their national populations, and have less than a primary school education (data not shown).

We present the cumulative distribution of ANC visits by provider type and the relative decline in proportions of women across the visits in **Figure 3**. In Bangladesh, Cambodia, Cameroon and Peru, the proportions of women who reported receiving ANC appears to gradually decrease as the number of visits increases. In Senegal and Uganda, there seems to be a pronounced drop off between three and four or more visits; in Nepal this noticeable drop off occurs between four and five visits. Similar to the results in **Table 1**, most women reported receiving care from a skilled provider during one or more ANC visits. The relative decline in the proportion of women who reported receiving care from unskilled providers decreased faster than the relative decline of the proportion of women who reported receiving care from skilled providers, as the reported number of visits increased.

We show the percentage of women receiving selected content interventions during any ANC visit among women reporting one to three or four or more visits in **Figure 4**. Women who reported four or more visits reported receiving at least one content intervention more than women reporting one to three visits, even though the increase in proportions varied across countries and among interventions; this pattern is also visible after stratifying the percentage of women receiving content by type of provider reported (Table S3 in **Online Supplementary Document**) However, a considerable proportion of women who reported the recommended four or more ANC visits did not receive any of the essential interventions at least once. Women who reported receiving ANC services and seeing a skilled provider at least once, seemed to report receiving more content

Table 1. Percentage of women who had a live birth in the five years preceding the DHS surveys who reported one ANC visit with a skilled provider and four or more visits with any provider or skilled provider for their most recent live birth, in seven Countdown countries

	One or more ANC visits with any provider (%)	ANC 1+ (with a skilled provider) (%)	ANC 4+ (with any provider) (%)	Four or more ANC visits with a skilled provider* (%)	Women reporting ANC 4+ with any provider and present as a subset among women reporting ANC 1+ with a skilled provider (%)
Bangladesh 2011	64.6	51.7	23.9	19.9	83.4
Cambodia 2010	89.6	89.1	59.6	59.4	99.7
Cameroon 2011	85.4	84.9	62.9	62.7	99.7
Nepal 2011	84.9	58.2	50.1	40.0	79.8
Peru 2012	98.4	96.0	94.4	92.2	97.6
Senegal 2010	95.8	93.2	51.2	50.2	98.1
Uganda 2011	95.7	94.8	48.5	48.1	99.3

DHS – Demographic Health survey, ANC – antenatal care

*At least one visit of the four or more visits is with a skilled provider.

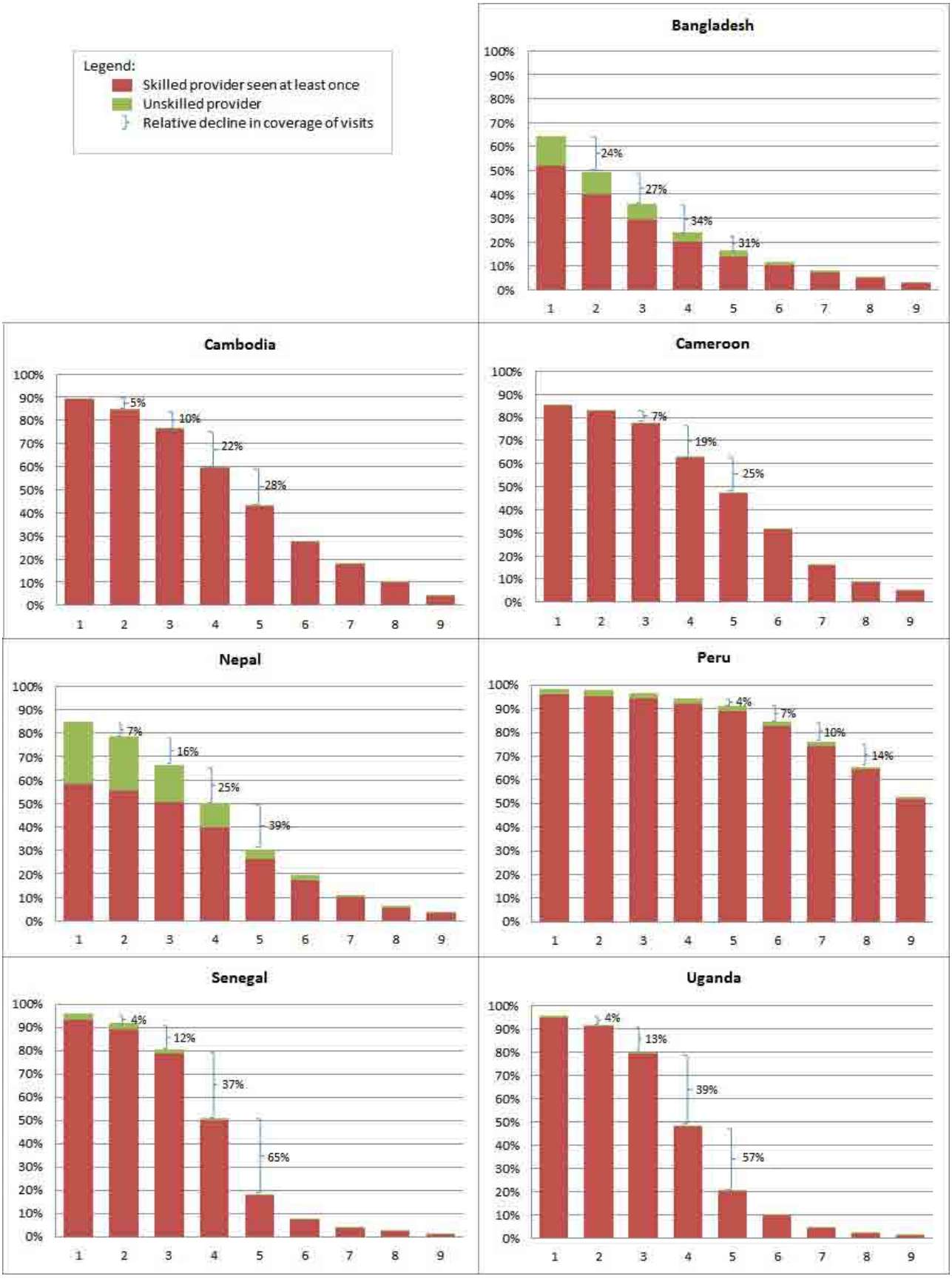


Figure 3. Cumulative percentage of women who had a birth in the five years preceding the DHS surveys by number of ANC visits and type of provider for their most recent live birth, in seven Countdown countries.

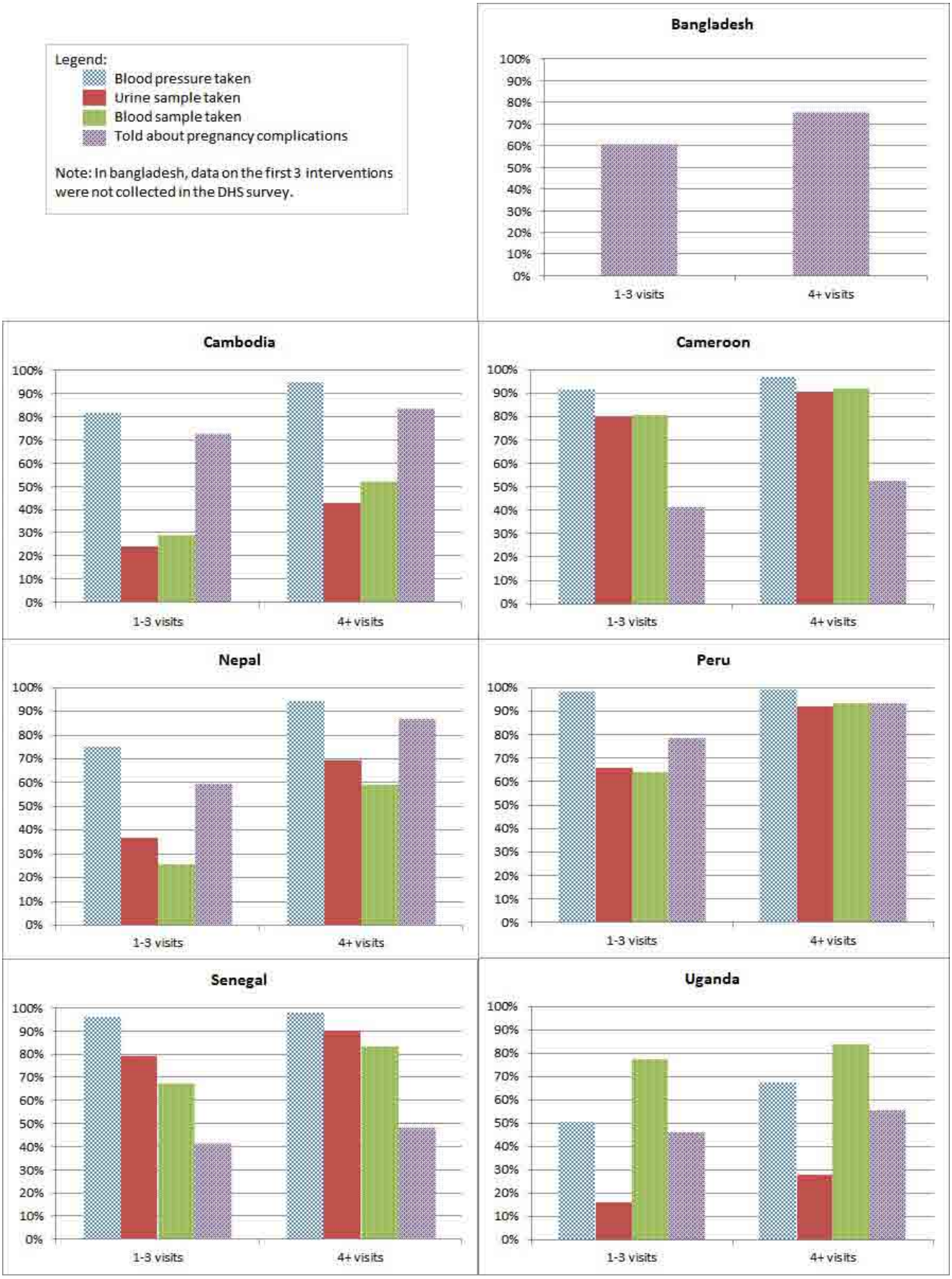


Figure 4. Percentage of women receiving content interventions during any ANC visit among women reporting one to three ANC visits or four or more ANC visits for their last live birth in the five years preceding the DHS survey, in seven Countdown countries.

interventions than women who received care from unskilled providers (Table S3 in **Online Supplementary Document**). Stratification of women's reported content by household wealth quintile (Table S4 in **Online Supplementary Document**), showed that as women's wealth status increased, so did their proportions of reporting receipt of content interventions at least once during any visit.

We looked at gestational age at first ANC visit among women who reported receiving ANC visit for their most recent birth in the five years preceding the DHS survey. The results show that women who had their first ANC visit during the first trimester reported a higher number of visits overall (Figure S5 in **Online Supplementary Document**). In Cameroon and Uganda, the proportion of women who made their ANC debut during the second trimester is high relative to other countries. Nearly 40% of women in Cameroon and around half of the women in Uganda started ANC during their second trimester. In all seven countries, the proportion of women who report starting ANC in the third trimester is around 5% with the exception of Uganda, where 13% of the women report receiving ANC for the first time in the third trimester; these women generally report three or less visits.

We present the distribution of women who reported receiving ANC visits by household wealth quintile in **Figure 5**. Wide disparities in the proportions of women reporting utilization exist across the wealth quintiles, except in Peru and Uganda. Top inequality exists where women in the richest wealth quintile are much better off than the rest; bottom inequality exists where women in the poorest wealth quintile are worse off than the rest of the women [35]. Inequality patterns differ by country. In Bangladesh, ANC utilization patterns clearly show top inequality, whereas in Cambodia and Nepal a pattern of top inequality begins to emerge only as the number of visits increases. The greatest disparities among countries are found in the proportions of women's reported visits by educational level (Figure S6.A in **Online Supplementary Document**). Women with the highest level of education report the highest proportions of visits. A pattern of top inequality emerges as the number of visits increases in all countries except Peru, where the inequality by woman's education is minimal and linear. Inequalities also exist in ANC utilization by place of residence (Figure S6.B in **Online Supplementary Document**). In all seven countries, women living in urban areas reported higher proportions of visits compared to their counterparts residing in rural areas. As a result of the drop off in utilization in Senegal and Uganda (**Figure 2**), the proportions of women who reported four or more visits show a considerable decline across all the wealth quintiles, educational levels and by urban–rural residence, in addition to a noticeable widening of the gap across categories of the three stratifiers as the number of visits reaches four or more ANC visits.

Model–based results

In the multivariable analysis, we sought to identify determinants of ANC initiation in Model A (zero visits vs one or more visits), and of frequency of visits in Model B (one to three visits vs four or more visits). Within each country, determinants predicting initiation of care and frequency of visits differed somewhat, except in Bangladesh where similar determinants predicted the outcome measures in both models at relatively similar odds ratios (OR). We present the results of Model B for the seven countries in **Table 2**; the results of model A and B for each country are found in in Tables S7A–G in **Online Supplementary Document**. When we entered the external environment factors, initially, place of residence had a significant effect on ANC in both models in all countries, except Uganda; however, as the subsequent levels were added to the models place of residence became insignificant. The exceptions are Bangladesh and Senegal where women residing in rural areas were less likely to report at least one visit, in Model A (OR: 0.67 and 0.62, respectively) and less likely to report four or more visits in Model B (OR: 0.52 and 0.74, respectively), compared to women residing in urban areas. Generally, among women's predisposing characteristics, educational level was the strongest predictor of the outcome measures. In Peru, Senegal and Uganda, educational level was significantly positively associated with initiation of care (in Model A) only. In Bangladesh, Cambodia, Cameroon and Nepal, having any level of education significantly increased the odds of initiating ANC and having a higher frequency of visits, compared to having no level of education. Woman's age at last birth became less significant as the determinants from other levels were added into the Models. The only age group which recorded a significant effect on the outcome measures was the 20–34 years age group, where women in this age group in Cambodia, Nepal and Uganda were more likely to report having at least one ANC visit compared to women who were less than 20 years old (odds ratios ORs: 1.58, 1.54 & 1.62, respectively). While in Cambodia, Cameroon and Peru, women aged 20–34 were more likely to have four or more visits compared to women who were less than 20 years old (ORs: 1.43, 1.57 & 2.48, respectively). The effect of the other 'predisposing' characteristics (woman's occupational status, religion and marital status) were generally not significantly related to women's reported frequency of visits. Within the enabling resources, current partners' education affected the outcome measures positively in several countries. In Bangladesh, having a secondary or higher education and in Cambodia and Nepal having a secondary education increased the odds of initiat-

Table 2. Multivariable analysis results for Model B (1–3 ANC visits vs 4 or more ANC visits) for seven Countdown countries after hierarchical enter of the determinants into the analysis

Characteristics		Bangladesh			Cambodia			Cameroon			Nepal			Peru			Senegal			Uganda		
		OR	95% CI	P–value	OR	95% CI	P–value	OR	95% CI	P–value	OR	95% CI	P–value	OR	95% CI	P–value	OR	95% CI	P–value	OR	95% CI	P–value
External Environment:*																						
Place of residence	Urban	*Ref*			*Ref*			*Ref*			*Ref*			*Ref*			*Ref*			Exc†	†Exc	†Exc
	Rural	0.52	0.42–0.65	0.000	0.75	0.55–1.02	0.068	0.92	0.71–1.19	0.529	0.88	0.64–1.22	0.454	1.05	0.70–1.58	0.814	0.74	0.61–0.88	0.001	†Exc	†Exc	†Exc
Predisposing characteristics:																						
Woman's age at last birth	<20	Exc†	Exc†	Exc†	Ref			Ref			Ref			Ref			Exc†	Exc†	Exc†	Exc†	Exc†	Exc†
	20–34	Exc†	Exc†	Exc†	1.43	1.07–1.91	0.017	1.57	1.24–1.97	0.000	1.17	0.88–1.56	0.278	2.48	1.46–4.20	0.001	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†
	35–49	Exc†	Exc†	Exc†	1.48	0.98–2.24	0.064	1.46	1.06–2.01	0.021	1.21	0.70–2.10	0.490	1.67	0.82–3.37	0.155	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†
Woman's educational level	No education	Ref			Ref			Ref			Ref			Ref			Ref			Ref		
	Primary	1.58	1.19–2.10	0.002	1.27	1.00–1.61	0.046	1.15	0.88–1.50	0.312	1.75	1.32–2.32	0.000	0.46	0.21–1.03	0.058	1.00	0.83–1.21	0.971	1.05	0.82–1.34	0.721
	Secondary	2.19	1.67–2.87	0.000	1.62	1.21–2.16	0.001	1.56	1.13–2.16	0.007	2.28	1.72–3.01	0.000	0.74	0.31–1.79	0.508	0.84	0.62–1.14	0.265	1.04	0.77–1.40	0.794
	Higher	4.23	2.91–6.13	0.000	2.94	0.84–10.25	0.090	3.89	1.85–8.16	0.000	6.12	2.57–14.58	0.000	0.95	0.32–2.78	0.923	1.94	0.58–6.52	0.281	1.46	0.92–2.33	0.106
Woman's employment status	Did not work in the past 12 months	Exc†	Exc†	Exc†	Ref			†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc
	Worked in the past 12 months	†Exc	†Exc	Exc†	0.88	0.70–1.11	0.286	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc
Religion	Dominant religion	*Ref*			*Ref*			†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc
	Other religions	1.25	0.97–1.61	0.090	0.51	0.30–0.86	0.012	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc
Marital status	Currently in union	†Exc	†Exc	Exc†	†Exc	†Exc	†Exc	*Ref*			†Exc	†Exc	†Exc	*Ref*			†Exc	†Exc	†Exc	†Exc	†Exc	†Exc
	formerly/Never in union	†Exc	†Exc	Exc†	†Exc	†Exc	†Exc	0.84	0.60–1.18	0.318	†Exc	†Exc	†Exc	1.29	0.49–3.36	0.607	†Exc	†Exc	†Exc	†Exc	†Exc	†Exc
Enabling factors:																						
Current partner's educational level	No education	*Ref*			*Ref*			*Ref*			*Ref*			*Ref*			*Ref*			†Exc	†Exc	†Exc
	Primary	1.09	0.87–1.38	0.445	1.16	0.88–1.53	0.283	1.06	0.82–1.36	0.677	1.44	1.05–1.98	0.024	3.20	1.39–7.37	0.006	1.11	0.89–1.38	0.340	†Exc	†Exc	†Exc
	Secondary	1.37	1.09–1.73	0.008	1.59	1.16–2.18	0.004	1.18	0.87–1.59	0.288	1.56	1.12–2.18	0.009	2.48	1.04–5.92	0.041	1.23	0.91–1.67	0.176	†Exc	†Exc	†Exc
	Higher	1.95	1.47–2.59	0.000	2.13	1.17–3.87	0.014	1.20	0.73–1.96	0.466	1.28	0.81–2.03	0.292	2.10	0.67–6.62	0.205	1.85	1.16–2.94	0.010	†Exc	†Exc	†Exc
Current partner's most recent employment status	No recent occupation	†Exc	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	N/A‡	N/A‡	N/A‡	N/A‡	N/A‡	N/A‡				N/A‡	N/A‡	N/A‡
	Recently had an occupation	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	N/A‡	N/A‡	N/A‡	N/A‡	N/A‡	N/A‡				N/A‡	N/A‡	N/A‡
Household wealth quintile	Poorest	*Ref*						*Ref*			*Ref*			*Ref*			*Ref*			*Ref*		
	Poorer	1.16	0.87–1.54	0.325	1.23	0.97–1.55	0.089	1.17	0.93–1.46	0.176	1.04	0.77–1.41	0.795	1.36	0.92–2.02	0.122	1.17	0.99–1.38	0.061	1.16	0.90–1.49	0.244
	Middle	1.55	1.18–2.02	0.002	1.07	0.81–1.40	0.633	1.19	0.89–1.59	0.241	1.05	0.72–1.54	0.800	1.41	0.75–2.66	0.282	1.42	1.16–1.73	0.001	1.23	0.94–1.61	0.125
	Richer	2.15	1.63–2.83	0.000	1.59	1.19–2.12	0.002	1.60	1.13–2.26	0.008	1.55	1.04–2.33	0.033	1.29	0.58–2.85	0.529	1.57	1.20–2.06	0.001	1.66	1.21–2.29	0.002
	Richest	4.08	2.97–5.60	0.000	1.71	1.20–2.44	0.003	1.84	1.24–2.73	0.002	2.36	1.44–3.85	0.001	3.07	0.82–11.49	0.095	1.95	1.42–2.68	0.000	1.98	1.42–2.77	0.000
Decision regarding woman's health care	Woman alone	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	*Ref*			*Ref*			Exc†	Exc†	Exc†	Exc†	Exc†	Exc†
	Woman & partner	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	0.93	0.72–1.19	0.554	0.81	0.53–1.25	0.347	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†
	Partner alone	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	0.72	0.53–0.99	0.041	0.65	0.42–0.99	0.047	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†
	Someone else/other	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	0.69	0.52–0.91	0.010	0.33	0.06–1.82	0.200	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†
Need factors:																						
Gestational age at initial visit	First trimester	**N/A‡**	**N/A‡**	**N/A‡**	*Ref*			*Ref*			*Ref*			*Ref*			*Ref*					
	Second trimester	**N/A‡**	**N/A‡**	**N/A‡**	0.23	0.20–0.27	0.000	0.24	0.20–0.28	0.000	0.36	0.29–0.46	0.000	0.18	0.12–027	0.000	0.29	0.25–0.35	0.000	0.26	0.21–0.32	0.000
	Third trimester	**N/A‡**	**N/A‡**	**N/A‡**	0.02	0.01–0.04	0.000	0.03	0.02–0.04	0.000	0.02	0.01–0.06	0.000	0.01	0.01–0.02	0.000	0.08	0.04–0.15	0.000	0.05	0.03–0.08	0.000
Birth rank	1st – 2nd births	*Ref*			*Ref*			Exc†	Exc†	Exc†	*Ref*			*Ref*			*Ref*			Exc†	Exc†	Exc†
	3rd – 4th births	0.87	0.72–1.04	0.128	0.85	0.69–1.05	0.124	Exc†	Exc†	Exc†	0.85	0.65–1.11	0.223	0.72	0.45–1.14	0.156	0.87	0.73–1.04	0.121	Exc†	Exc†	Exc†
	5th births or more	0.43	0.29–0.64	0.000	0.55	0.40–0.76	0.000	Exc†	Exc†	Exc†	0.58	0.37–0.91	0.018	0.53	0.30–0.93	0.026	0.77	0.65–0.91	0.003	Exc†	Exc†	Exc†
Preceding birth interval	First birth	*Ref*			*Ref*			*Ref*			*Ref*			Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†
	Less than 2 years	0.71	0.51–0.97	0.033	0.75	0.55–1.03	0.076	0.64	0.47–0.86	0.004	0.60	0.42–0.86	0.005	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†
	2–3 years	0.62	0.48–0.81	0.000	0.82	0.63–1.08	0.156	0.72	0.55–0.95	0.022	0.69	0.51–0.95	0.021	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†
	More than 3 years	0.80	0.68–0.95	0.010	0.71	0.56–0.91	0.005	0.76	0.59–0.98	0.032	0.84	0.65–1.10	0.215	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†
Pregnancy wanted at the time	Then	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†
	Later	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†
	No more	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†	Exc†

ANC – antenatal care, OR – odds ratio, 95% CI – 95% confidence interval, Ref – reference category, Exc – excluded, N/A – not available

*Results related to the determinant "region of the country" can be found in Tables 7A–G in **Online Supplementary Document.**

†Excluded by backward elimination during hierarchical entry at the level or due to insignificance at the bivariate level.

‡Country data was not available for determinant.

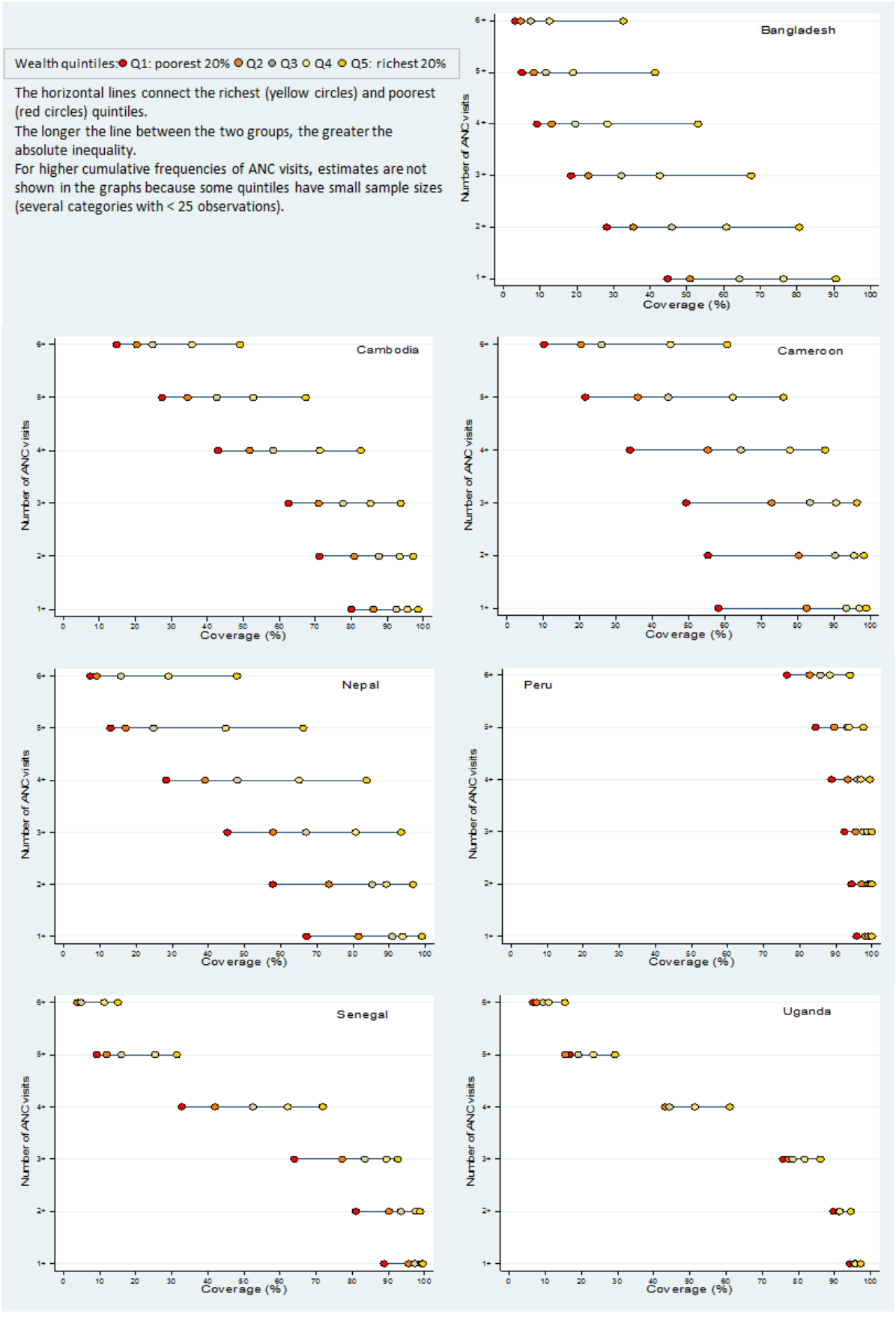


Figure 5. Percentage of women who had a live birth in the five years preceding the DHS surveys by number of ANC visits and household wealth quintiles, in seven Countdown countries.

ing care (Model A) and the odds of having four or more visits compared to women whose current husbands have no education. Household wealth was the strongest predictor of the outcome measures in this level. Household wealth was positively associated with reporting four or more visits in all countries. The richest quintiles presented odds ratios of 2.4 to 6.1 and 1.7 to 3.1 compared to the poorest in Models A and B, respectively; with Senegal having the highest effect–odds ratio of 7.7 in Model A and Bangladesh having the highest effect–an odds ratio of 4.1 in Model B. Decision regarding the woman's health care was only significant in Cameroon (Model A), Nepal (both models) and Peru (Model B) and had a negative effect on the outcome measure. In the fourth and final hierarchical level, the "need" factors, women whose gestational age at initial visit was in the second or third trimester were significantly less likely to have four or more visits compared to women who had their initial visit in the first trimester. Birth rank had a significant negative association with initiation of ANC in Cambodia and Nepal; in Bangladesh, Peru and Senegal the negative effect of birth rank is significant at the 5th birth or higher. Birth rank is also negatively associated with the frequency of ANC visits among women's 5th birth or more in Bangladesh, Cambodia, Nepal, Peru and Senegal. Preceding birth interval was a strong predictor of the outcome measure in both models in Bangladesh, Cambodia, Cameroon and Nepal showing a negative association.

Descriptive review of ANC–related national policies and programs

Each of these seven countries has its own set of national policies, strategies and guidelines around health–related issues and ANC, specifically. **Tables 3** and **4** summarize the information obtained through the questionnaire. Countries vary widely in terms of their ANC–related policies, programs, standards, and guidelines. Here we use selected examples to explore how these variations may have affected the ANC utilization patterns presented in the descriptive and multivariable analysis above. We describe this link with caution, because our data sets are not sufficiently complete or quantitative to determine directional causality.

Bangladesh and Peru, at the two extremes of ANC 1+ and ANC 4+ coverages among these seven countries, have comprehensive guidelines and policies related to ANC. Although, ANC utilization is relatively low in Bangladesh, trends in coverage of ANC 1+ and ANC 4+ have been increasing steadily since the early 1990s [13]. Unlike Bangladesh and the other five countries, the Peruvian government goes beyond the WHO guidelines of four ANC visits and recommends a minimum of six scheduled visits.

Both Senegal and Uganda show a distinct drop–off in ANC utilization between the third and fourth visits. In Senegal, several reproductive health–related policies were either updated or developed between 2002 and 2005, and one of the changes included moving from a standard of three to four ANC visits. In Uganda, the government has adopted a four–visit, focused ANC approach, and recently introduced guidelines addressing HIV/AIDS and prevention of mother–to–child transmission that refers to ANC as a platform for care and treatment. However, the reported number of visits by gestational age during the first ANC visit (Table S5 in **Online Supplementary Document**) showed that 66% and 13% of Ugandan women report initiating care during their second and third trimester, respectively, which inevitably means there is less time to complete the recommended number of visits prior to childbirth. The results of Uganda's multivariable analysis also show that as the gestational age at first ANC visit increases, women are significantly less likely to report receiving four or more ANC visits compared to three or fewer.

In Nepal, the proportion of women receiving care only from an unskilled provider was the highest among the seven countries, followed by Bangladesh. Our contextual information showed that both these countries had clear guidelines permitting unskilled providers to offer certain ANC services. The *National Medical Standards for Reproductive Health* guideline, adopted by the Nepali government in 2009, states that in the absence of a skilled birth attendant in the facilities serving rural areas, a maternal and child health worker or a health assistant (categorized as unskilled providers in our study) can provide ANC services [36]. With 90% of our sample of Nepali women residing in rural areas, high proportions of women may have only had access to unskilled providers offering ANC services. Furthermore, some reports from Nepal refer to unskilled providers such as health assistants, auxiliary health workers, maternal and child health workers, and village health workers as trained professionals [37]. In Bangladesh, similar to Nepal, the majority of women reside in rural areas (around 75% of our sample), which are served by a complex network of public health facilities offering ANC services by skilled and unskilled providers. At community level, providers now considered unskilled for ANC, historically provided services at primary facilities and household level through both government and non–government agencies [38].

DISCUSSION

The globally–measured ANC indicators, ANC 1+ and ANC 4+, need to be accompanied by more detailed analysis of ANC utilization patterns in each country in order to unpack the underlying factors and inequalities that play a role in women's uptake of ANC services.

We intentionally selected countries for analysis from different world regions and with varying levels of ANC 1+, ANC

Table 3. Summary of policies and programs related to ANC, in seven Countdown countries

	Selection of national policies and their reference to ANC	Policies & efforts to tackle inequities	ANC–related campaigns or communication efforts
Bangladesh	*National Maternal Health Strategy*–2001: • Specifies the supply of equipment for ANC, the delivery of care and a monitoring checklist. • Outlines interaction with pregnant women and their families to ensure ANC uptake and popularization of service delivery mechanisms as well as the use of ANC for birth preparedness. *Health, Population and Nutrition Sector Development Plan (HPNSDP), 2011–2016:* • ANC mentioned as a key service in emergency obstetric care needs and management. • Specifically mentions distribution of folic acid/and iron supplementation. *National Plan of Action for Adolescent Sexual and Reproductive Health*, 2013: • Specifically published to tackle the major concern of teenage pregnancies which make up around 30% of adolescents aged 15–19 y old.	The *HPNSDP*–2011–2016 prioritizes the improvement of ANC by: • Emphasizing maternal, newborn, child and adolescent health interventions/services in urban slums, hard to reach and low performing areas. • Prioritizing areas of high maternal mortalities and geographically & socially disadvantaged population.	Available through various types: • Television and radio programs. • Mass communication during the Safe Motherhood days when ANC is promoted along with other services. • Posters and pamphlets available at health facilities.
Cambodia	*Health Sector Strategic Plan for 2008–2015:* • Plans to scale up access to and coverage of health services, especially comprehensive reproductive, maternal, newborn and child health services. *Safe Motherhood Clinical Management Protocols for Referral Hospital*–June 1013 & *Safe Motherhood Clinical Management Protocols for Health Centre*–July 2010: • Provide technical updates regarding frequency of ANC visits (from 2+ to 4+), timing (as early as knowing the missing period), and additional services (screening).	The *Health Sector Strategic Plan* – 2008–2015 has pro–poor policies including: • Pro–poor health financing systems • Exemptions for the poor • Expansion of the health equity funds • Health insurance	To promote early ANC visits an ANC campaign took place in 2009 using both mass media and interpersonal communication in selected geographical areas.
Cameroon	Prenatal care centres (*Soins Prenataux Recentres*)–2006 • Includes change from the focus on the number of visits to the quality of the visit. *Prevention of Mother to Child Transmission of HIV*–November 2008	A project initiated by the World Bank in 14 districts to test performance–based financing addresses some aspects of inequity.	No specific efforts noted
Nepal	*National Health Policy* – 1991: • Adopted the safe motherhood approach with *the Safe Motherhood Program* being a priority. *Safe Motherhood Policy*–1996: • Focuses on improving maternal health in a holistic way *National Policy on Skilled Birth Attendants NPSBA)*–2006 *National Medical Standard for Reproductive Health*–2009*:* • An updated version of the *National Maternity Care Guidelines (NMCG)* which was released in 2006. • A standard reference document for essential clinical materials and tools in support of patient care using the latest evidence in maternal and neonatal care. • Uses the concepts of focused antenatal care. • Emphasis is on every pregnant woman being at risk, birth preparedness and complication readiness, providing quality rather than quantity of antenatal care.	No specific efforts noted.	• The Government implemented a communication strategy. • ANC–related messages are broadcast through radio.
Peru	*Comprehensive Health Insurance* (*The Seguro Integral de Salud (SIS)) – 2001*: • Aims to protect the health of Peruvians who do not have health insurance, prioritizing those vulnerable populations who are in poverty or extreme poverty. *Budgeting for results* (*Presupuesto por Resultados)* – 2008: • It proposes action based on critical problem solving and includes Strategic Programs such as the one for Mother and Newborn (which was established for women in extreme poverty & no health insurance). *Technical document: National Strategic Plan for Reduction of Maternal and Neonatal Mortality (RN No. 207–2009)* *Technical Guides: Intervention Model to improve Access, Quality and Use of facilities that provide obstetric and neonatal functions (RM No. 223–2009/MINSA)* • provides strategies to improve availability, accessibility and use of facilities. *Technical Standard for the comprehensive care of maternal health (RM No. 827–2013/MINSA)* • Establishes the technical requirements and administrative procedures, based on scientific evidence, that allow to deliver quality care in preparation for pregnancy, refocused prenatal care, institutional and skilled delivery care and postpartum care.	• The *Seguro Integral de Salud (SIS)*–2001 • The *Mother and Newborn Strategic Program* • The *Technical Standard for Vertical Delivery with Attention to Intercultural Adaption*: intended to improve access for Andean and Amazonian women of childbearing age. • The establishment of Maternal Waiting homes to increase access to delivery care in health facilities.	Different media used to communicate importance of ANC

Table 3. Continued

	Selection of national policies and their reference to ANC	Policies & efforts to tackle inequities	ANC–related campaigns or communication efforts
Senegal	*National Program for the Prevention of Maternal Mortality (Programme Nacional de Prevention de la Mortalite Maternelle)*–1990 *Politique de Santé et d'Action Sociale* (Health Policy and Social Action)–1995 • Placed reproductive health as one of the cornerstones. *Population Policy Statement (Déclaration de Politique de Population)* • Established in 1998 & updated in 2002 to be consistent with the ICPD. *National Program of Reproductive Health (Programme Nacional de santé de la Reproduction)*–2002 *A multi–sectoral roadmap* • Developed to accelerate the reduction of maternal and neonatal mortality in order to achieve MDGs 4 & 5.	The national strategy for all women of reproductive age has elements for making services available to all–geographically, financially, socio–culturally, and to all religious groups through: • Increasing points of service delivery. • Provision of minimum package of reproductive health services at health facilities. • Adjusting the fees/costs according to people's abilities to pay.	• Government conducted nationwide scale up campaign with radio and television spots on malaria prevention with SP and use of ITNs. • NGOs support this campaign by broadcast general messages on antenatal care through local–community radio.
Uganda	*The National Policy Guidelines and Service Standards for Sexual and Reproductive Health and Rights*–2012 (3rd update) • Sets rules and regulations governing reproductive health services including antenatal services • Outlines tasks that guide service provision and describe aspects of ANC services • Emphasizes integration of services such as access to services for sexually transmitted infections and HIV/AIDS services at the ANC clinic *Integrated National Guidelines on Antiretroviral Therapy, Prevention of Mother to Child Transmission and on Infant & young Child feeding*–2011 • Facilitates integration of services and to promote a family–centered approach for HIV and AIDS care and treatment. • ANC is recognized as a platform for this care and treatment.	A voucher scheme for pregnant women is being piloted in a few areas.	• Radio messaging on particular aspects of ANC, eg, malaria prevention and prevention of mother to child transmission of HIV • Some projects in limited geographic areas have used phone text messages to ANC clients

ANC – antenatal care, SP – sulfadoxine–pyrimethamine, ITN – Insecticide–treated bednet, NGO – non–governmental organization, ICPD – International Conference on Population and Development, MDG – Millennium Development Goals, HIV – human immunodeficiency virus, AIDS – Acquired Immune Deficiency Syndrome.

4+ and skilled birth attendance coverage. Skilled birth attendance is lowest, at 32% and 36%, in Bangladesh and Nepal, respectively [13]; in addition to having the lowest proportions of one or more and four or more ANC visits, our results also showed that these two Asian countries reported the highest prevalence of ANC provision by an unskilled provider. However, local definitions of what constitutes a skilled provider seem to vary in these two countries as described in the descriptive review of national policies and programs.

Nevertheless, the majority of women in the seven Countdown countries reported receiving care from a skilled provider at least once. The reported content interventions, on the other hand, require more attention. Even in Peru, where over 90% of women reported receiving four or more visits, evidence–based content was highest relative to other countries yet not universal. Regardless of countries' diverse settings, women who reported four or more ANC visits, who received care from a skilled provider at least once and were better off in terms of household wealth, reported receiving a higher proportion of each of the four content interventions. Similar results were also presented by Hodgins et al (2014), who looked at DHS data on content interventions of ANC visits from countries [31]. In their analysis, the proportion of content interventions (out of eight) among women who reported four or more visits ranged from 32% to 85% in the 41 countries and the overall average was 60% [31]. These relationships need to be explored further at the country level to understand whether content interventions are not being provided during ANC visits or are being postponed to later during the pregnancy, resulting in missed opportunities for women who report a low number of visits. Or, on the contrary, the perceived usefulness and quality of the interventions offered at health care facilities may play a role in women's decisions about whether to return for subsequent visits. We did not consider the health care facilities providing the ANC services in our quantitative analysis, yet this may play an important role in women's ANC utilization patterns especially if the quality of care is perceived as poor. Powell–Jackson and colleagues looked at the quality of ANC services in the private commercial sector, private not–for–profit sector, public sector and home from DHS data in 46 low– and middle–income countries [39]. The content of care score was worst in home–based care, where women received the least number of ANC services, followed by both the public and private commercial sectors with similar scores and the private not–for–profit had the highest ANC content score. The researchers conclude that the private commercial and

Table 4. Summary of national standards and guidelines for ANC, in seven Countdown countries

	Recommended number of ANC visits	Presence of guidelines for visit content	Where ANC services are provided	Who provides the ANC services	Presence of user fees	Incentives for women's utilization
Bangladesh	Follows the WHO recommendation: • 1st visit: before 16 weeks. • 2nd visit: 24–28 weeks. • 3rd visit: 30–32 weeks. • 4th visit: 36–38 weeks.	Present	Provided at both private sector and public sector (primary, secondary and tertiary facilities) and through NGOs. Home–based ANC may be provided in rural areas.	In urban areas and the private sector, doctors usually provide ANC. In rural areas there is a wider array of skilled and unskilled providers who offer ANC services.	No public sector fees. Private facilities charge fees for service.	In some Upazila Health Centers (public facilities) patients receive transportation cost.
Cambodia	Follows the WHO recommendation: • 1st visit: before 16 weeks (or as soon as possible after a missed menstrual period). • 2nd visit: 24–28 weeks. • 3rd visit: 30–32 weeks. • 4th visit: 36–38 weeks.	Present	Provided at health centers (primary facilities) or hospitals (tertiary level).	Types of providers & services are the same in urban and rural public facilities. Services in private facilities depend on ability to pay. ANC services are generally provided by midwives.	Public facilities have user fee schemes. Private sector facilities have a fee–for–service.	Some schemes offer indirect incentives through: • Health equity fund • Voucher scheme linking ANC services to other MCH services.
Cameroon	Recommends four visits: • 1st visit at 1–16 weeks amenorrhea. • 2nd visit at 28 weeks. • 3rd visit at 32 weeks. • 4th visit at 36 weeks.	None provided	Present at all health facilities.	Providers do not vary according to public/private sector or to rural/urban areas. ANC services are provided by various skilled & unskilled providers and at various workstations in one facility.	Both the public and private sector charge fees at different rates.	No incentives available
Nepal	Recommends four visits: • 1st visit at 4 months • 2nd visit at 6 months • 3rd visit at 8 months • 4th visit at 9 months	Present	In rural areas, ANC is provided at sub–health posts, health posts and district hospitals. In urban areas, ANC is provided at private clinics and maternity hospitals.	All service providers should be skilled birth attendants (these include nurses and doctors). If these skilled providers are not available at Sub–health posts and out–reach clinics, then MCH Workers can provide ANC services.	No public sector fees. Private sector charges vary.	Incentives provided to women who complete 4 ANC visits and have an institutional delivery.
Peru	Recommends a minimum of 6 visits: • 1st visit: at less than 14 weeks • 2nd visit: 14–21 weeks • 3rd visit: 22–24 weeks • 4th visit: 25–32 weeks • 5th visit: 33–36 weeks • 6th visit: 37–40 weeks	Present	Most ANC services are provided through the network of 8000 public facilities. Home visits are made when women miss their scheduled visit.	Service providers are mainly skilled. Unskilled providers are usually involved in the health team particularly at the first level of the health system.	Fees depend on different funding sources.	Specific program created in 2005, provides program grants for direct transfers to benefit the poorest families, rural and urban.
Senegal	Recommends at least 4 visits: • 1st visit at 3 months • 2nd visit at 6 months • 3rd visit at 8 months • 4th visit at 9 months	Present	Provided through health huts, health posts, maternity centers or hospitals and private clinics.	Standards and protocols stipulate that only skilled providers can provide ANC services at both public and private facilities	Both public and private sector facilities charge fees but at different rates.	Insecticide–treated bednets are provided to pregnant women.
Uganda	Recommends four focused ANC visits: • 1st visit: 0– 16 weeks (after two missed periods). • 2nd visit: 16–28 weeks. • 3rd visit: 28–36 weeks. • 4th visit: after 36 weeks. • 4th visit: after 36 weeks.	Present	Provided at hospitals, health centers, and sometimes at outreach clinics.	Skilled providers provide ANC services in all facilities. Unskilled providers such as community health workers & village health team members can provide information. Nursing assistants & nurse aids are being phased out.	No public sector fees. In Private not–for–profit facilities fees are subsidized. Private for profit sector generally does not subsidize ANC except for immunization.	Mama Kit of essential supplies to use during delivery (gloves, protective sheets, baby receiving sheet, soap) are provided to pregnant women.

ANC – antenatal care, WHO – World Health Organization, NGO – non–governmental organization, MCH – maternal and child health.

public sectors are both very diverse and show lots of variation in quality of care, possibly depending on the economic status of women seeking ANC care [39].

In order to ensure high coverage of accessible and equitable ANC services, programs and policies need to focus on women with low levels of education, living in poor and rural households. Our findings indicate that household wealth is an important determinant of ANC initiation in all seven countries, and of the overall frequency of visits in all countries except Peru. This is consistent with a systematic review of the relevant literature [27]. Our multivariable analysis results showed that women who come from poorer households are less likely than richer women to initiate care, and among those who do seek ANC, less likely to have four or more visits. These multivariable results complement the results of the equity analysis, and highlight the need to address financial barriers to accessing ANC services. ANC services are offered by the public sector free of charge in Bangladesh, Nepal and Uganda, and yet utilization is relatively low. Unexpected fees for prescribed medications or tests, and indirect costs related to transportation to the facility, have been associated with women's choices of health care services, and need to be considered [28]. The exceptional case of Peru may reflect its unique combination of political will, economic growth, broad societal participation, pro–poor strategies and increased spending in health and related sectors in the last two decades, which led to reduction in socioeconomic inequalities in health and significant progress in coverage of RMNCH interventions especially among the most deprived groups and areas of the country [40].

Education allows women to be more autonomous, more knowledgeable about health care services, and therefore to exert greater control over health–related decisions. We would therefore expect women's education to have a positive influence on the initiation and frequency of ANC visits [19,20,27], and this is supported by our results. The wide disparities in ANC utilization by levels of women's education may also be due to the uneven distribution of women within each category; women who have attained higher levels of education are fewer in number and generally better off than those who report having low or no education. The results of the multivariable analysis showed that education was a significant determinant for initiating ANC, and to a lesser extent, for reporting four or more ANC visits. Similar to the results of Guliani and colleagues [19], who looked at the use of ANC services and their frequency across 32 low income countries, the association of women's education was stronger with the initiation of ANC visits than with the overall number of visits. This may be because women with no education are not included in the second model, which looks at the frequency pattern of utilization [19].

The equity analysis showed that women residing in rural areas have lower proportions of ANC utilization than women residing in urban areas, and our policy data suggested important differences in services and providers in urban and rural areas in most country settings. We were therefore surprised that urban–rural residence emerged as a significant determinant of ANC visits only in Bangladesh and Senegal. A systematic review looking at early use of ANC services and type of residence concluded that the association was not consistent [41], hence further country analysis is required to identify the contextual factors that affect ANC use. A study looking at contextual influences of 13 sub–regions in Nepal on women's ANC patterns identified important sub–regional variations in ANC use, which need to be taken into consideration at the policy–making level [42]. Our findings reinforce the importance of regional differences in ANC utilization within countries, and suggest that further analysis of this relationship is likely to generate information useful for ANC program planning.

We collected some information on barriers to ANC utilization from national surveys or ethnographic studies through our key informant interviews, but were often unable to obtain full and relevant information. Qualitative studies exploring barriers to antenatal care are available in the literature, and can contribute to the interpretation of our findings. In Bangladesh, despite multiple maternal health policies and an extensive array of public sector ANC facilities, women are not utilizing ANC services as recommended. This has been attributed by some to women's perception that pregnancy is a normal event that does not need medical care and interventions [43]. In Cambodia, five types of barriers to maternal health care use have been identified as needing to be addressed to increase ANC utilization: financial, physical, cognitive, organizational and psychological/socio–cultural [44]. One study reported that the use of ANC services by pregnant women in Nepal was greatly influenced by mothers–in–law [45]; this is corroborated by our multivariable results showing a negative association between women's reports that they are not responsible for health care decisions and the frequency of visits. Another study reports that mothers in Uganda viewed ANC services as deficient, and are dissatisfied with the perceived quality of the interventions offered during visits [46]. WHO has identified several barriers to the provision of quality ANC, including perceptions of poor quality of care, distance, cost, stigma, social and traditional influences, perceptions that pregnancy is a healthy state that does not need specific care, and disrespect for and abuse of women in health service settings [47]. Furthermore, a systematic scoping review performed to understand what women seek during pregnancy, found that across diverse settings, having a positive pregnancy experience was what

mattered to pregnant women and this was characterized by four themes: preserving physical and sociocultural normality; maintaining a healthy pregnancy for mother and baby; effective transition to progressive labour and birth; and ensuring positive motherhood [48].

Our study has some limitations that need to be kept in mind when interpreting the results. DHS data are collected on ANC–related questions, only, for women's most recent live birth in the five years preceding the survey. Hence, pregnancies resulting in a miscarriage or a stillbirth are excluded and no data on the ANC utilization patterns of these pregnancies is collected; data which may provide important insights to the patterns and quality of care in these cases specifically. Furthermore, the DHS data we use is from interviews with women and their responses to the frequency of ANC visits, types of providers seen, and content interventions received for their most recent live birth. As a result, there may be potential recall bias, an issue which is receiving increased awareness in mothers' reports of services received [49]. Women may also be biased in their reports of the type of provider from whom they received care, especially in settings where several types of providers offer ANC care. Other limitations related to the type of provider are the fact that the DHS data cannot tell us the type of provider visited during each of the woman's ANC visit and that the choice of providers may be restricted to who is available at the health care facility in the community. We report on content received during any ANC visit, and while it does imply a minimum level of quality of care, we are unable to confirm it, because the DHS does not assess whether the content interventions were offered in a proper and timely way. Women are also asked to report on their own and their partner's employment status at the time of the survey interview, and this may have changed since the time of their most recent pregnancy. The DHS does not include questions that are directly related to the barriers to seeking ANC or accessing ANC facilities. However, we used the question about who makes decisions about the woman's health care as a measure of one potential barrier to access of services. Our findings do not reflect other potential barriers such as distance to a facility or associated financial costs. Another limitation is that the data we were able to obtain on policies and programs in each country varied greatly, and because most of our key informants volunteered their time, it was not always possible to confirm all details or to seek additional information and clarification within the time frame of this study.

Nevertheless, this analysis has several strengths. We examine ANC utilization from a new perspective, focusing specifically on information missed by global tracking of only two indicators. The study brings together several types of analysis–descriptive, equity, and inferential analysis–to generate new and detailed results of specific characteristics of women and their households that are associated with ANC utilization patterns in seven diverse countries.

CONCLUSION

The results of this study indicate that reporting the globally measured ANC indicators, ANC 1+ and ANC 4+, is useful to provide an overall idea of the proportions of ANC utilization in countries. However, descriptive and multivariable analyses generated a much better understanding of each country's unique pattern of ANC utilization, as well as the characteristics of women not currently receiving adequate care. The presence of variations across countries suggests the need for specific in–country assessments, national panels, or advisory groups to look more closely at national data, commission specific studies and perhaps try different models of ANC to find ways to achieve universal ANC coverage.

A number of predominant aspects of ANC utilization patterns emerged across the seven Countdown countries. Our results highlight the need to focus on evidence–based content interventions offered to women during their ANC visits. Further quantitative assessments of the frequency and timeliness of content interventions by different types of providers and in different settings are needed to ensure proper administration of the WHO–recommended interventions. Moreover, qualitative studies looking into the barriers of ANC use and women's perceptions of ANC services in each country are essential. There is a growing body of literature that focuses on women's perceptions of pregnancy and quality of ANC services and how important this aspect is on ANC uptake. The current DHS protocol asks about barriers to seeking health care in general; it may provide important insights to include a question specifically about the numerous barriers that may affect women's initiation of ANC and completion of the four recommended visits.

Inequality in ANC utilization patterns among women of different wealth statuses, educational backgrounds and places of residence need to be considered at the policy–making level across most of the countries we studied. These dimensions of inequality were strong predictors of ANC utilization and higher frequency of visits, except for place of residence. The influence of place of residence on ANC utilization in Bangladesh and Senegal suggests the need to assess the health care services offered in rural areas. And the lack of significance of this factor in the five other countries suggests that there are variations within each place of residence that need to be identified and used to provide effective interventions. While we found that policies and guidelines related to ANC as well as 'Safe Motherhood' strategies were incorporated into the national policies,

across the seven countries, yet, there is a need to ensure evaluation and proper implementation of these policies and strategies. Peru is an example of successful implementation of political and structural reform, which took place in the 1990s, and led to the enhancement of health systems and infrastructure, reduction in poverty, and the introduction of insurance schemes, one of which is for pregnant mothers, among other groups [50].

With the end of the MDG era, few countries have achieved the MDG5 goal of reducing maternal mortality by three quarters, and most have a long way to go before achieving universal access to reproductive health services [13]. Most maternal deaths are preventable, and the causes of these deaths are known [2,47,51]. By increasing attention and investment to providing quality maternity, antenatal, and post–partum care, life–saving interventions may be administered properly and in a timely manner by skilled health providers to help improve maternal and neonatal health and their survival [2]. In the post–2015 agenda, as the Sustainable Development Goals and their measurable indicators are being set, it is essential to include targets for ending preventable maternal deaths and to ensure that the momentum focusing on maternal and reproductive health–with ANC as a vital component–continues [51].

Acknowledgements: We would like to thank Liliana Carvajal (UNICEF), Allisyn Moran (USAID) and Cesar Victora (Federal University of Pelotas) for their valuable insights during the conceptualization of this research. We are also grateful for the assistance of the key informants in each country: From Bangladesh–Dr Sushmita Roy (Center for Child and Adolescent Health, ICDDR,B) and D.M. Emdadul Hoque (Center for Child and Adolescent Health, ICDDR,B), in addition to numerous resource persons. From Cambodia–Rathmony Hong (UNICEF Cambodia) and Dr Lam Phirun (National Reproductive Health Program, Ministry of Health Cambodia). From Cameroon–Dr Marquise Kouo Ngamby (Reproductive Health Unit, Ministry of Health). From Nepal–Radha Paudel (Action Works Nepal). From Peru–Luisa Sacieto Carbajo (La Salud del Recién Nacido, Peru) and Maria Gabriela Villlalobos (Latin America and Caribbean Region, Save the Children USA). From Senegal–Medoune Diop (Counterpart International). And from Uganda–Miriam Sentongo (Ministry of Health, Uganda).

Funding: This work was funded through a sub–grant from the U.S. Fund for UNICEF under the Countdown to 2015 for Maternal, Newborn and Child Survival grant from the Bill & Melinda Gates Foundation.

Disclaimer: The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the U.S. Fund for UNICEF.

Authorship declarations: All authors were involved in the conceptualization and designed of this manuscript. GSH performed the secondary data analyses and wrote the first draft of the manuscript. All other authors provided critical input to the data analyses and manuscript. NT conducted the contextual data collection process and presented the data in a thorough report to be used in the manuscript. All authors read and approved the final manuscript.

Competing Interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare no conflict of interest.

REFERENCES

1 United Nations. Framework of actions for the follow–up to the Programme of Action of the International Conference on Population and Development Beyond 2014. United Nations, 2014.
2 United States Agency for International Development. Ending preventable maternal mortality: USAID maternal health vision for action – Evidence for strategic approaches. Washington DC: USAID, 2015.
3 Lincetto O, Mothebesoane–Anoh S, Gomez P, Munjanja S, Munjanja S. Chapter 2: Antenatal care In: Lawn J, Kerber K, editors. Opportunities for Africa's Newborns: practical data, policy and programmatic support for newborn care in Africa. Geneva: World Health Organization; 2006. p. 51–62.
4 World Health Organization. Standards for maternal and neonatal care. Geneva: World Health Organization; 2007.
5 World Health Organization. Pregnancy, childbirth, postpartum and newborn care: a guide for essential practice. Report No.: 3rd edition. Geneva: WHO, 2015. 1–184 p.
6 Dowswell T, Carroli G, Duley L, Gates S, Gülmezoglu AM, Khan–Neelofur D, Piaggio GGP. Alternative versus standard packages of antenatal care for low–risk pregnancy. Cochrane Database Syst Rev. 2010;(10):CD000934. Medline:20927721
7 World Health Organization. WHO Antenatal care randomized trial: manual for the implementation of the new model. Geneva: World Health Organization, 2002.

REFERENCES

8 Villar J, Carroli G, Khan–Neelofur D, Piaggio G, Gülmezoglu M. Patterns for routine antenatal care for low–risk pregnancies care. Cochrane Database Syst Rev. 2001;(4):CD000934. Medline:11687086

9 World Health Organization. WHO statement on antenatal care. Geneva: WHO, 2011.

10 Vogel JP, Abu Habib N, Souza JP, Gulmezoglu AM, Dowswell T, Carroli G, et al. Antenatal care packages with reduced visits and a perinatal mortality: A secondary analysis of the WHO antenatal care trial. Reprod Health. 2013;10:19-26. Medline:23577700 doi:10.1186/1742-4755-10-19

11 Kearns A, Millar K. Continuum of care series: Providing for mom and baby through antenatal and postnatal care. Maternal Health Task Force Blog 2014 Aug 1. Available: http://www.mhtf.org/2014/08/01/call-for-posts-continuum-of-care-series-providing-for-mom-and-baby-through-antenatal-and-postnatal-care/. Accessed: 1 February 2015.

12 United Nations. Indicators for monitoring the Millennium Development Goals: definitions, rational, concepts and sources 2012. Available: http://mdgs.un.org/unsd/mi/wiki/MainPage.ashx. Accessed: 9 April 2015.

13 World Health Organization. Countdown to 2015. Fulfilling the health agenda for women and children: the 2014 report. Geneva: WHO, 2014.

14 Magadi MA, Madise NJ, Rodrigues RN. Frequency of timing of antenatal care in Kenya: Explaining the variations between women of different communities. Soc Sci Med. 2000;51:551-61. Medline:10868670 doi:10.1016/S0277-9536(99)00495-5

15 Sepehri A, Sarma S, Simpson W, Moshiri S. How important are individual, household and commune characteristics in explaining utilization of maternal health services in Vietnam? Soc Sci Med. 2008;67:1009-17. Medline:18635302 doi:10.1016/j.socscimed.2008.06.005

16 Alexandre PK, Saint–Jean G, Crandell L, Fevrin E. Prenatal care utilization in rural areas and urban areas of Haiti. Rev Panam Salud Publica. 2005;18:84-92. Medline:16156958 doi:10.1590/S1020-49892005000700002

17 Beeckman K, Louchx F, Putman K. Determinants of the number of antenatal visits in a metropolitan region. BMC Public Health. 2010;10:527-36. Medline:20809951 doi:10.1186/1471-2458-10-527

18 Gayawan E. A poisson regression model to examine spatial patterns in antenatal care utilization in Nigeria. Popul Space Place. 2014;20:485-97. doi:10.1002/psp.1775

19 Guliani H, Sepehri A, Serieux J. Determinants of prenatal care use: evidence from 32 low–income countries across Asia, sub–Saharan Africa and Latin America. Health Policy Plan. 2014;29:589-602. Medline:23894068 doi:10.1093/heapol/czt045

20 Kishowar Hossain AH. Utilization of antenatal care services in Bangladesh: an analysis of levels, patterns, and trends from 1993 to 2007. Asia Pac J Public Health. 2010;22:395-406. Medline:20498123 doi:10.1177/1010539510366177

21 Shrestha G. Factors related to utilization of antenatal care in Nepal: a generalized linear approach. Journal of Kathmandu Medical College. 2013;2:69-74.

22 Celik Y, Hotchkiss DR. The socio–economic determinants of maternal health care utilization in Turkey. Soc Sci Med. 2000;50:1797-806. Medline:10798333 doi:10.1016/S0277-9536(99)00418-9

23 Magadi MA, Zulu EM, Brockerhoff M. The inequality of maternal health care in urban sub–Saharan Africa in the 1990s. Pop Stud (Camb). 2003;57:347-66. Medline:14602534 doi:10.1080/0032472032000137853

24 Mekonnen Y, Mekonnen A. Factors influencing the use of maternal healthcare services in Ethiopia. J Health Popul Nutr. 2003;21:374-82. Medline:15038593

25 Saxena D, Vangani R, Mavalankar DV, Thomsen S. Inequity in maternal health care service utilization in Gujarat: Analyses of district–level health survey data. Glob Health Action. 2013;6:1-9. Medline:23469890

26 Houweling TAJ, Ronsmans C, Campbell OMR, Kunst AE. Huge–poor–rich inequalities in maternity care: an international comparative study of maternity and child care in developing countries. Bull World Health Organ. 2007;85:745-54. Medline:18038055 doi:10.2471/BLT.06.038588

27 Simkhada B, van Teijlingen ER, Porter M, Simkhada P. Factors affecting the utilization of antenatal care in developing countries: systematic review of the literature. J Adv Nurs. 2008;61:244-60. Medline:18197860 doi:10.1111/j.1365-2648.2007.04532.x

28 Finlayson K, Downe S. Why do women not use antenatal services in low– and middle–income countries? A meta–synthesis of qualitative studies. PLoS Med. 2013;10:e1001373. Medline:23349622 doi:10.1371/journal.pmed.1001373

29 USAID. The DHS Program: Demographic and Health Surveys. Maryland: ICF International. Available: http://dhsprogram.com/. Accessed: 1 November 2012.

30 Andersen RM. Revisiting the behavioral model and access to medical care: Does it matter? J Health Soc Behav. 1995;36:1-10. Medline:7738325 doi:10.2307/2137284

31 Hodgins S, D'Agostino A. The quality–coverage gap in antenatal care: toward better measurement of effective coverage. Glob Health Sci Pract. 2014;2:173-81. Medline:25276575 doi:10.9745/GHSP-D-13-00176

32 Babitsch B, Gohl D, Lengerke TV. Re–visiting Anderson's behavioral model of health services use: a systematic review of studies from 1998–2011. Psychosoc Med. 2012;9:Doc11. Medline:23133505

33 Chakraborty N, Islam AM, Chowdhury RI, Bari W, Akhter HH. Determinants of the use of maternal health services in rural Bangladesh. Health Promot Int. 2003;18:327-37. Medline:14695364 doi:10.1093/heapro/dag414

34 Victora CG, Huttly SR, Fuchs SC, Olinto MT. The role of conceptual frameworks in epidemiological analysis: a hierarchical approach. Int J Epidemiol. 1997;26:224-7. Medline:9126524 doi:10.1093/ije/26.1.224

REFERENCES

35 Boerma JT, Bryce J, Kinfu Y, Axelson H, Victora CG. Mind the gap: equity and trends in coverage of maternal, newborn, and child health services in 54 countdown countries. Lancet. 2008;371:1259-67. Medline:18406860 doi:10.1016/S0140-6736(08)60560-7

36 Mo HP. [Nepal]. National Medical Standard for Reproductive Health. Report No.: Volume III: Maternal and Newborn Care. Kathmandu: Family Health Division, Ministry of Health and Population, Government of Nepal, 2009.

37 Dahal RK. Utilization of antenatal care services in rural area of Nepal. Int J Collab Res Intern Med Public Health. 2013;5:120-31.

38 Mridha MK, Anwar I, Koblinski M. Public–sector maternal health programmes and services for rural Bangladesh. J Health Popul Nutr. 2009;27:124-38. Medline:19489411 doi:10.3329/jhpn.v27i2.3326

39 Powell–Jackson T, Macleod D, Benova L, Lynch C, Campbell O. The role of private sector in the provision of antenatal care: a study of demographic and health surveys from 64 low– and middle–income countries. Trop Med Int Health. 2015;20:230-9. Medline:25358532 doi:10.1111/tmi.12414

40 Huicho L, Segura E, Huayanay C, Niño de Guzman J, Restrepo-Mendez MC, et al. Placing child health and nutrition at the centre of the antipoverty political agenda: the Peruvian experience, 2000-2013. Lancet Glob Health. 2015. submitted

41 Say L, Raine R. A systematic review of inequalities in the use of maternal health care in developing countries: examining the scale of the problem and the importance of context. Bull World Health Organ. 2007;85:812-9. Medline:18038064 doi:10.2471/BLT.06.035659

42 Mathews SA, Gubhaju B. Contextual influences on the use of antenatal care in Nepal. Report No.: DHS Geographic Studies 2. Calverton, Maryland: ORC Macro, 2004.

43 Government of Bangladesh. Bangladesh Maternal Mortality and Health care Survey (BMMS). Bangladesh: Government of Bangladesh, 2010.

44 Matsuoka S, Aiga H, Rasmey LC, Rathavy T, Okitsu A. Perceived barriers to utilization of maternal health services in rural Cambodia. Health Policy. 2010;95:255-63. Medline:20060193 doi:10.1016/j.healthpol.2009.12.011

45 Simkhada B, Porter MA, van Teijlingen ER. The role of mothers–in–law in antenatal care decision–making in Nepal: A qualitative study. BMC Pregnancy Childbirth. 2010;10:34. Medline:20594340 doi:10.1186/1471-2393-10-34

46 Ndyomugyenyi R, Neema S, Magnussen P. The use of formal and informal services for antenatal care and malaria treatment in rural Uganda. Health Policy Plan. 1998;13:94-102. Medline:10178189 doi:10.1093/heapol/13.1.94

47 The independent Expert Review Committee. Every woman, every child: a post–2015 vision: the third report of the independent Expert Review Group on Information and Accountability for Women's and Children's health. Geneva: WHO, 2014.

48 Downe S, Finayson K, Tuncalp O, Gulmezoglu AM. What matters to women: A systematic scoping review to identify the processes and outcomes of antenatal care provision that are important to healthy pregnant women. BJOG. 2015. doi:10.1111/1471-0528.13819. Epub ahead of print. Medline:26701735

49 Stanton CK, Rawlins B, Drake M, dos Anjos M, Cantor D, Chongo L, et al. Measuring coverage in MNCH: testing the validity of women's self–report of key maternal and newborn health interventions during the peripartum period in Mozambique. PLoS ONE. 2013;8:e60694. Medline:23667427 doi:10.1371/journal.pone.0060694

50 PAHO/WHO, Partnership for Maternal, Newborn & Child Health, World Bank and Alliance for Health Policy and Systems Research. Success factors for women's and children's health: Peru. Geneva: World Health Organization; 2015.

51 Requejo JH, Bhutta ZA. The post–2015 agenda: staying the course in maternal and child survival. Arch Dis Child. 2015;100 Suppl 1:S76-81. Medline:25613979 doi:10.1136/archdischild-2013-305737

journal of global health

Assessing the validity of indicators of the quality of maternal and newborn health care in Kenya

Ann K Blanc[1], Charlotte Warren[2], Katharine J McCarthy[1], James Kimani[3], Charity Ndwiga[3], Saumya RamaRao[1]

[1] Population Council, New York, NY, USA
[2] Population Council, Washington D.C., USA
[3] Population Council, Nairobi, Kenya

Background The measurement of progress in maternal and newborn health often relies on data provided by women in surveys on the quality of care they received. The majority of these indicators, however, including the widely tracked "skilled attendance at birth" indicator, have not been validated. We assess the validity of a large set of maternal and newborn health indicators that are included or have the potential to be included in population–based surveys.

Methods We compare women's reports of care received during labor and delivery in two Kenyan hospitals prior to discharge against a reference standard of direct observations by a trained third party (n = 662). We assessed individual–level reporting accuracy by quantifying the area under the receiver operating curve (AUC) and estimated population–level accuracy using the inflation factor (IF) for each indicator with sufficient numbers for analysis.

Findings Four of 41 indicators performed well on both validation criteria (AUC>0.70 and 0.75<IF<1.25). These were: main provider during delivery was a nurse/midwife, a support companion was present at birth, cesarean operation, and low birthweight infant (<2500 g). Twenty–one indicators met acceptable levels for one criterion only (11 for AUC; 9 for IF). The skilled birth attendance indicator met the IF criterion only.

Interpretation Few indicators met both validation criteria, partly because many routine care interventions almost always occurred, and there was insufficient variation for robust analysis. Validity is influenced by whether the woman had a cesarean section, and by question wording. Low validity is associated with indicators related to the timing or sequence of events. The validity of maternal and newborn quality of care indicators should be assessed in a range of settings to refine these findings.

Correspondence to:
Ann K. Blanc, PhD
Population Council
New York, NY, USA
ablanc@popcouncil.org

Nearly 275 000 maternal deaths occurred globally in 2011, nearly all of which took place in low– and middle–income countries (LMIC) [1]. Most of these countries did not reduce maternal mortality to levels targeted in the Millennium Development Goals (MDG5) [1]. Progress has been hindered, in part, by a lack of reliable maternal health data, especially on maternal deaths [2]. Measurement challenges are particularly significant in LMIC with irregular and incomplete health system reporting.

To measure progress in maternal health, monitoring agencies have relied on tracking indicators proposed as measures of quality of care, such as the proportion of births attended by a skilled birth attendant, that are assumed to be strongly correlated with maternal mortality [3]. Such indicators are routinely assessed in population—based household survey programs, such as the Demographic and Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS), in which female respondents report on events surrounding recent births [4]. Despite their widespread use, the majority of proposed quality of care indicators, including skilled birth attendance, have not been validated [1,5,6]. In fact, numerous researchers have noted the lack of correlation between these indicators and maternal mortality levels [5,7–9]. These researchers argue that information on the category of provider at birth is deficient as a measure of quality of care as it relies on assumptions about provider training and competence as well as access to essential supplies and equipment. It is important therefore to identify alternate indicators that describe the actual content of care, can be reported with accuracy, and have the potential to be included in routine data collection programs.

A growing, but still limited, body of research has examined the validity of indicators of the quality of care in the intrapartum and early postpartum period. To our knowledge, however, no study has yet reported on how accurately women can recall the skill level of their provider at birth, although there have been some attempts to look at data quality issues [10]. Furthermore, the few validation studies that have taken place have generally compared maternal self–reports with hospital records, which may be incomplete or inaccurate, or have been conducted in high–income settings, where maternal mortality rates are generally low [11–15].

To address this gap, this study assessed women's ability to report on a set of quality of maternal and newborn health care indicators that are either currently in use or have the potential to be included in routine survey–based data collection. In spite of its limitations, it seems likely that the "skilled birth attendance" indicator will continue to be used and so we assess how accurately women report on the skill level of their provider during delivery. We compare women's self–reports of maternal and newborn care received against third party observations during labor and delivery. Finally, we provide suggestions for modifications to data collection procedures that could improve the measurement of maternal and newborn health care.

METHODS

Study sites

Validation exercises were conducted in two high volume public hospitals located in Kisumu District and Kiambu District in western and central Kenya, respectively. According to the 2014 Kenya DHS, nationally, 61% of births in the five years preceding the survey were delivered in a health facility; in Kisumu and Kiambu districts the prevalence was 70% and 93%, respectively [16]. Facility–based delivery is less likely among older women, those who have lower education, are poorer, or reside in rural areas [16]. Fertility levels among women in the two districts are lower than the national rate, with the total fertility rate in Kisumu at 3.6 births per woman and in Kiambu at 2.7, compared with 3.9 nationally [16].

Data collection

Data collection took place from July to September 2013. All pregnant women aged 15 to 44 who were admitted to a study facility maternity unit and in early labor were invited to participate. Participants included eligible women who underwent labor and delivery and were able to provide consent.

Our reference standard for validity analysis is data collected by trained researchers who observed providers in the maternity admission room and labor and delivery rooms using a structured checklist–type form. Observers were registered Kenyan nurse/midwives with at least three years of experience in a maternal and newborn health unit and previous research experience. Observations were used as the reference standard as they reflected all facets of caregiving including events related to the birth itself as well as interactions between the women and provider, before, during and up to one hour after delivery. In the few cases in which clarification was needed (eg, in the event the mother and infant were taken into separate rooms, the observer remained with the mother) observations were supplemented by checking facility records and by asking providers.

Exit interviews with women took place prior to hospital discharge. Data collectors who were degree holders in a social science interviewed women using a structured questionnaire. Interview questionnaires were translated into Kiswahili, Dholuo and Kikuyu and were administered in the woman's language of preference.

All data collectors received four days of intensive training on the study procedures, the rationale behind each element of the client questionnaire and observation checklist to ensure full understanding of the instrument components, and how to record responses and observations.

Ethical review

Written informed consent was obtained from all participants and their attending providers prior to participation. All women and providers were provided with a description of the study and procedures, including their right to refuse participation at any time. In Kenya, pregnant adolescents

between ages 15–17 years are considered "emancipated minors" and their written informed consent was also obtained [17–19]. Staff who provide labor and delivery care were identified by the hospitals' obstetrics and gynaecology director, and approached for recruitment and consent. No providers refused participation.

Prior to participant enrollment, the study protocol was approved by the ethical review committees of the Population Council and the Kenya Medical Research Institute (KEMRI).

Indicator selection

To identify indicators to be validated, a landscaping scan of published and grey literature was conducted from April to July 2012. The scan focused on indicators of the quality and content of care received during labor and delivery and the health outcomes related to this period [20]. Indicators were included if they were currently in use or proposed for use in household survey programs such as the DHS and MICS or reflected standard practices of maternal and newborn labor and delivery care. The scan yielded a list of 285 indicators. This list was assessed by a group of public health experts specializing in maternal health to select a set of 80 indicators for validity testing. Indicators were selected on the basis of their wide use and/or potential to assess the critical elements of maternal and newborn care during the initial assessment of the woman, the first, second and third stages of labor, and immediate postnatal period.

Analysis

Sample size was calculated assuming 50% prevalence for all indicators, given that some harmful practices would rarely occur, and some beneficial practices would almost always occur, at 60% sensitivity ±6% precision, 70% specificity ±6% precision, with type 1 error set at $\alpha = 0.05$ assuming a normal approximation to a binomial distribution. These specifications imply a minimal sample size of 500, which was increased to 600 women to allow for 20% attrition in a separate study to re–interview women approximately one year following delivery.

Statistical analysis was performed using Stata Version 12 [21] to assess indicator accuracy at the individual and population level. For individual–level reporting accuracy, we calculated the sensitivity (ie, true positive rate) and specificity (ie, true negative rate) of indicators by constructing two–by–two tables for each indicator that had at least five counts per cell [22].

Missing pairwise data were excluded. To summarize the accuracy of each indicator, we quantified the area under the receiver operating curve (AUC), which plots the sensitivity (ie, true positive rate) of each indicator against its false positive rate (1–specificity). To measure uncertainty associated with validity, we estimated 95% confidence intervals (CI), assuming a binomial distribution. In practice, the AUC represents the "average accuracy of a diagnostic test" [23,24]. AUC values range from 1.0 (perfect classification accuracy) to 0 (zero accuracy). An AUC value of 0.5 is the equivalent of a random guess.

To assess the population–based validity of indicators, we estimated the inflation factor (IF), also known as the Test to Actual Positives (TAP) ratio [25]. The IF reflects the prevalence of the indicator as it would be reported by women in a survey after accounting for sensitivity and specificity (Pr) divided by the true prevalence (ie, observer report) (P). By comparing the ratio of the estimated survey–based prevalence to its true prevalence, we calculated the degree to which each indicator would be over or under–estimated by women's self–report (IF = Pr/P) [25,26].

The prevalence of women's self–report in a survey (Pr) is calculated by applying each indicator's estimated sensitivity (SE) and specificity (SP) to its true prevalence (P), using the following equation: $Pr = P \times (SE+SP-1)+(1-SP)$ [26]. We caution that the estimated survey–based prevalence is dependent on the observed prevalence of the indicator. Therefore, IF estimates reflect the magnitude of over or under–estimation in the study setting. To illustrate the implications of the IF estimates for other contexts in which the true prevalence is different from our study setting (eg, outside of a hospital facility), we model the estimated survey prevalence for select indicators across all possible coverage levels (ie, true prevalence ranging from 0 to 100%) using the above equation [27].

We categorized individual–level reporting accuracy as high ($AUC>0.70$), moderate ($0.60<AUC<0.70$), and low ($AUC<0.60$) [22] and the degree of bias reflected by the IF as low ($0.75<IF<1.25$), moderate ($0.50<IF<1.5$) and large ($IF<0.50$ or $IF>1.5$) [11]. In order to summarize indicator validity in terms of both individual and population–level accuracy, we considered indicators with high AUC and low IF to have high overall performance [22].

Role of the funding source

The funders of the study had no role in data collection, analysis, interpretation or writing of the study results, or decision to submit for publication.

RESULTS

Sample description

1039 women admitted to the maternity unit at participating study facilities were recruited to participate. Of those, 676 women were observed (Kiambu = 395, Kisumu = 281). Approximately one–third of women were not observed because they were not in labor but required monitoring on

the antenatal ward, did not progress into labor, or they progressed rapidly into labor and full observation was not possible (**Figure 1**). Fourteen women who were observed did not participate in the exit interview.

Participants' background characteristics and differences by facility location are presented in **Table 1**. The majority of women were under age 25 with fewer than two prior births, married, and with no or primary education. A greater percentage of Kisumu participants were never married, while fewer were married/living together or separated/widowed.

Validation results

The full list of indicators selected for validity testing is presented in **Table 2**. The table provides the prevalence for each indicator as reported by women and observers, which, for some indicators, varied substantially. For example, 73% of women reported that the provider(s) washed his or her hands or used antiseptic before any initial examination, while 27% of observers recorded that this took place. "Don't know" responses were minimal for most indicators. However, four indicators for which the proportion of women who responded "Don't know" exceeded 5% are reported in **Table 3**. Two of these indicators refer to the immediate postnatal period: whether the newborn was immediately dried after birth and whether the newborn was immediately wrapped in a towel. Having a cesarean section as opposed to a vaginal delivery was significantly associated with responding "Don't

Table 1. Sample background characteristics by facility location

	% total sample (N = 662)	% Kiambu (N = 388)	% Kisumu (N = 274)	*P*–value*
Age in years:				0.504
15–19	14.7	12.4	17.9	
20–24	40.8	41.8	39.4	
25–29	29.9	30.2	29.6	
30–34	8.6	9.3	7.7	
35–39	5.6	5.9	5.1	
40–44	0.5	0.5	0.4	
Prior parity (total number of live births):				0.435
0	50.2	49.7	51.3	
1	26.5	28.8	22.9	
2	14.0	13.2	14.9	
3	6.0	5.2	7.3	
4 or more	3.3	3.1	3.6	
Educational attainment:				0.001
None	10.3	10.3	10.2	
Primary	44.0	45.9	41.2	
Secondary	29.5	33.2†	24.1†	
Higher	16.3	10.6†	24.5†	
Marital status:				0.001
Single, never married	14.7	9.8†	21.5†	
Married/living together	83.4	87.6†	77.4†	
Separated/widowed	2.0	2.6†	1.1†	
Type of delivery:				0.679
Vaginal	86.6	87.0	85.9	
Cesarean section	13.4	13.0	14.1	

*Based on χ^2 test comparing facility locations, statistically significant at $P<0.05$.

†Statistically significant pairwise comparisons using the Holm–Bonferroni correction to adjust for multiple comparisons.

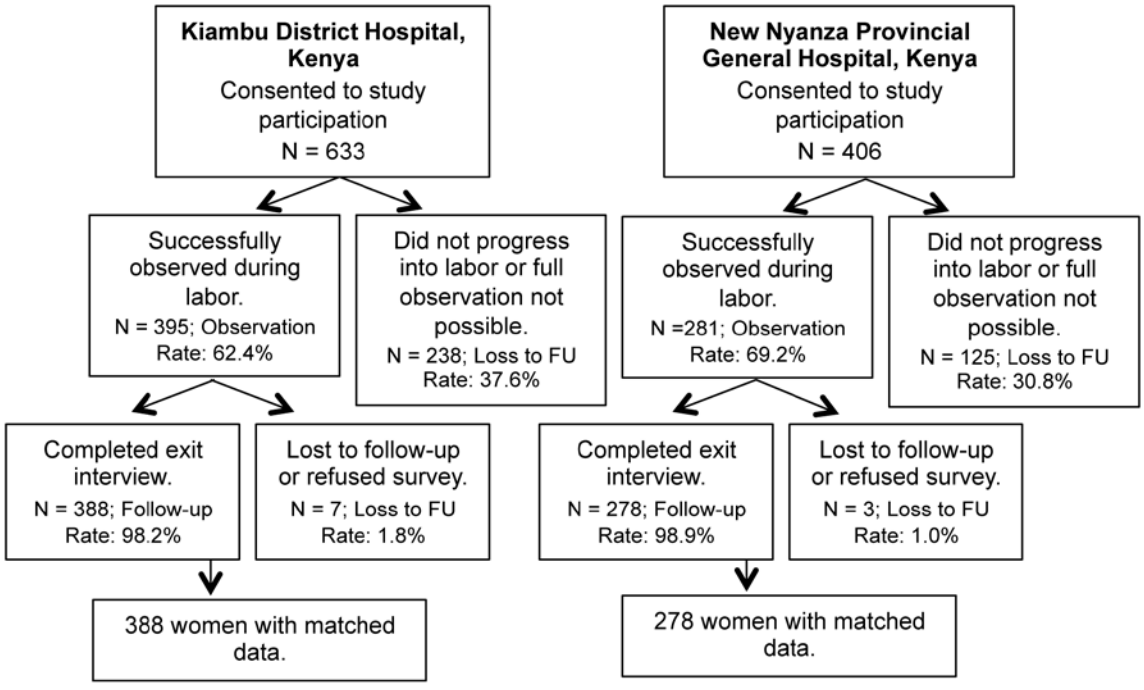


Figure 1. Participant response rates.

Table 2. List of indicators assessed and reported prevalence*

Indicator	N	Women's self–report, prevalence (%)	Observer report, prevalence (%)	At least 5 counts/ cell
Initial client assessment:				
Type of facility where gave birth (public hospital)	651	98.0	100.0	No
Referred to facility because of a problem	655	8.2	7.9	Yes
HIV status checked	659	25.2	94.7	No
Offered HIV test	660	8.3	1.7	No
Receives HIV test	654	8.6	9.3	No
Provider washes hands with soap and water or uses antiseptic before any initial examination	467	73.1	26.6	Yes
Takes blood pressure	654	93.4	87.0	Yes
Takes urine sample	654	5.7	1.4	No
Checks fetal heart rate with fetoscope/ultrasound	659	95.8	99.7	No
Wears high–level disinfected or sterile gloves for vaginal examination	658	99.9	99.9	No
Provider respectful care:				
Woman allowed to drink liquids/eat	624	66.8	42.0	Yes
Encourages/assists woman to ambulate during labor	644	87.0	77.5	Yes
Encourages/assists woman to assume different positions in labor	649	14.3	58.2	Yes
Woman allowed to have a support person present during labor and delivery	648	8.8	9.1	Yes
A support person is present at birth	644	3.7	4.8	Yes
First stage of labor:				
Induces labor by uterotonic (IV, IM, tablet)	630	10.8	4.6	Yes
Augments labor with uterotonic (by IV line, IM injection, or tablet)	625	39.2	22.4	Yes
Uterotonic received (to induce or augment labor)	619	43.8	27.1	Yes
Membranes ruptured (to induce or augment labor)	650	3.1	42.3	Yes
Skilled birth attendance–main provider:				
Main provider labor				
Skilled main provider labor†	649	89.9	92.6	Yes
Main provider labor–doctor or medical resident	649	9.6	0.5	No
Main provider labor–doctor (ob–gyn)	649	9.6	0.3	No
Main provider labor–medical resident	649	0.0	0.2	No
Main provider labor–medical intern	649	0.2	1.9	No
Main provider labor–nurse/midwife	649	80.2	92.1	Yes
Main provider labor–clinical officer	649	2.2	0.6	No
Main provider labor–facility support/ staff aide	649	0.2	0.3	No
Main provider labor–student nurse	649	2.3	2.8	No
Main provider labor–support companion	649	0.6	1.7	No
Main provider labor–no one or other	649	4.6	0.0	No
Main provider delivery				
Skilled main provider delivery†	644	94.3	92.9	Yes
Main provider delivery– doctor (ob–gyn) or medical resident	644	19.3	11.8	Yes
Main provider delivery– doctor (ob–gyn)	644	19.1	3.0	No
Main provider delivery–medical resident	644	0.2	8.9	No
Main provider delivery–medical intern	644	0.3	1.1	No
Main provider delivery–nurse/midwife	644	75.0	81.1	Yes
Main provider delivery–clinical officer	644	2.2	0.8	No
Main provider delivery–student nurse	644	2.2	4.8	Yes
Main provider delivery–no one or other	644	0.9	0.5	No
Second and third stage of labor:				
Episiotomy performed	545	22.9	18.2	Yes
Uterotonic administered within few minutes of delivery (via injection, IV medication, or oral/rectal tablets)	562	96.8	98.8	No
Uterotonic received 1–3 min after birth	552	96.9	81.5	No
Uterotonic received after delivery of placenta	552	59.1	2.4	Yes
Applies controlled cord traction	561	97.5	98.9	No
Performs uterine massage after delivery of placenta	558	88.4	98.6	No
Position of mother at birth–on back	645	94.7	99.8	No
Health provider wore gloves during delivery of baby	563	100.0	99.8	No
Immediate postnatal newborn care:‡				
Newborn given to mother immediately after birth	611	59.9	57.6	Yes

Table 2. Continued

Indicator	N	Women's self–report, prevalence (%)	Observer report, prevalence (%)	At least 5 counts/ cell
Newborn immediately dried with towel/cloth	594	96.1	99.5	No
Newborn placed immediately skin to skin on mother's chest§	602	78.9	16.3	Yes
Newborn immediately skin to skin on mother (2 item indicator)#	596	29.2	16.2	Yes
(Of newborns not on skin) Newborn wrapped with towel	86	90.7	91.9	No
Breastfeeding initiated within first hour of birth	551	76.4	53.0	Yes
Something other than breastmilk given to baby within first hour of delivery	572	1.9	1.1	No
Baby bathed within the first hour after birth	604	2.8	0.1	No
Baby weighed	635	99.8	100.0	No
Low birth–weight baby (<2500g)	579	6.7	7.8	Yes
High birth–weight baby (<4500g)	579	1.0	1.0	No
3 elements of newborn care (immed. · dried + on skin + breastfed in first hour)	506	71.5	9.3	Yes
3 elements of newborn care (immed. · dried, 2 item skin–to–skin#, breastfed in first hour)	501	25.1	9.2	Yes
Immediate postnatal care for mother:				
Palpates uterus 15 min after delivery of placenta	557	88.3	70.2	Yes
Provider did at least one post–delivery health check	649	96.0	94.9	No
In first post–delivery exam, provider checks for bleeding	627	62.0	90.6	Yes
In first post–delivery exam, provider examines perineum	554	56.1	87.4	Yes
In first post–delivery exam, provider takes temperature	638	60.0	40.3	Yes
In first post–delivery exam, provider takes blood pressure	642	74.6	48.3	Yes
In first post–delivery exam, provider checks for involution	615	64.2	78.2	Yes
Woman asked for pain relief medication while at facility	638	32.1	10.5	Yes
Woman received pain relief medication	640	59.4	17.5	Yes
Maternal and infant outcomes:				
Cesarean section (C/S) performed	651	13.5	13.4	Yes
Decision for C/S taken before labor started	76	9.2	0.0	No
(Of women who had a C/S) C/S performed after labor started	76	90.8	100.0	No
(Of women who had a C/S) Provider decided C/S would be done	80	82.5	100.0	No
(Of women who had a C/S) Reason for C/S–prolonged/obstructed labor	76	32.9	67.1	Yes
Complications–any:	654	44.8	11.0	Yes
–Eclampsia	654	10.9	0.3	No
–Hemorrhage	654	11.2	4.6	Yes
–Prolonged labor (>12 h)	654	23.7	3.7	Yes
–None	654	51.5	89.0	Yes
(Of women who had complications) Blood products given	72	15.3	18.1	No
Stillborn delivery	651	0.9	1.4	No

*Table presents descriptive results. The sample size per indicator varied by women's responses to the question, and uses matched data, excluding 'Don't Know' and missing responses.

†Skilled provider is doctor (obstetrician/gynecologist, ob–gyn), nurse/midwife or medical resident.

‡Asked of mothers whose babies were breathing at birth.

§Newborn was placed against mother's chest after delivery.

#Indicator constructed from two skin–to–skin items: (1) newborn placed against mother's chest after delivery and (2) newborn was naked on skin (not wrapped in a towel).

Know" to both of these questions (OR = 15.3, 95% CI = 8.2–28.3, $P<0.001$; OR = 2.7, 95% CI = 1.2–5.9, $P=0.016$, respectively).

Of the 41 indicators with at least five cases in each cell of the two–by–two table, four had both high individual reporting accuracy and low population–level bias (**Table 4**). These were: the main provider during delivery was a nurse/midwife, a support companion was present during the birth, cesarean operation, and low birthweight infant (<2500 g). Receiving an episiotomy was close to meeting both criteria (AUC = 0.87, 95% CI = 0.83–0.89, IF = 1.26).

A total of 8 indicators had high individual reporting accuracy (AUC>0.70), 7 had moderate accuracy (AUC>0.60), and 26 had low accuracy (AUC<0.60). Indicators with high AUC results reflected events leading up to (eg, induction or augmentation of labor, episiotomy) and during the birth itself (eg, cesarean section, main provider during delivery was a doctor or medical resident, main provider during de-

Table 3. Indicators with greater than 5% "Don't know" responses

Respondent question	N	"Don't know" (%)
Did the health provider(s) wash his/her hands with soap and water or use antiseptic before examining you?	662	29.5
Was your baby dried off with a towel or cloth immediately after his/her birth, within a few minutes of delivery?	660	8.3
(Of women who reported newborn was not placed against her chest immediately after delivery) Was your baby wrapped in a towel or cloth immediately after birth?	170	20.6
In your first physical examination after delivery, did a health provider do a perineal exam?	662	9.8

Table 4. Validation results for indicators*

Indicator	Total number	Observer prevalence (%)	Sensitivity (95% CI)	Specificity (95% CI)	Self–report survey estimate (%), based on sensitivity and specificity	AUC (95% CI)	IF	High accuracy (AUC > 0.70) & low bias (IF 0.75–1.25)
Initial client assessment:								
Woman referred to facility because of a problem	655	7.9	25.0 (14.0–38.9)	93.2 (90.9–95.1)	8.2	0.59 (0.55–0.62)	1.04	IF
Provider washes hands with soap and water or uses antiseptic before initial examination	467	26.6	83.9 (76.2–89.9)	32.9 (28.0–38.2)	71.5	0.58 (0.54–0.63)	2.69	
Takes blood pressure	654	87.0	87.7 (84.9–90.2)	23.3 (11.8–38.6)	86.3	0.55 (0.52–0.59)	0.99	IF
Provider respectful care:								
Woman allowed to drink liquids or eat	624	42.0	72.5 (66.7–77.8)	37.3 (32.3–42.5)	66.8	0.55 (0.51–0.59)	1.59	
Encourages/assists woman to ambulate during labor	644	77.5	90.2 (87.2–92.6)	24.1 (17.4–31.9)	87.0	0.57 (0.53–0.61)	1.12	IF
Encourages/assists woman to assume different positions in labor	649	58.2	19.1 (15.2–23.4)	92.3 (88.4–95.1)	14.3	0.56 (0.52–0.59)	0.25	
Woman allowed to have a support person during labor and delivery	648	9.1	23.7 (13.6–36.6)	92.7 (90.3–94.7)	8.8	0.58 (0.54–0.62)	0.97	IF
Support companion present during birth	644	4.8	48.4 (30.2–66.9)	98.5 (97.2–99.3)	3.7	0.73 (0.70–0.77)	0.77	Both
First stage of labor:								
Induces labor with uterotonic	630	4.6	69.0 (49.2–84.7)	92.0 (89.6–94.1)	10.8	0.80 (0.77–0.84)	2.35	AUC
Augments labor with uterotonic	625	22.4	72.9 (64.7–80.0)	70.5 (66.2–74.5)	39.2	0.72 (0.68–0.75)	1.75	AUC
Uterotonic received (labor induction or augmentation)	619	27.1	78.0 (70.9–84.0)	69.0 (64.5–73.2)	43.8	0.73 (0.70–0.77)	1.61	AUC
Membranes ruptured (labor induction or augmentation)	650	42.3	4.0 (2.0–7.0)	97.6 (95.5–98.9)	3.1	0.51 (0.47–0.55)	0.07	
Skilled birth attendance:								
Skilled main provider† labor	649	92.6	90.5 (87.9–92.7)	16.7 (7.5–30.2)	90.0	0.54 (0.50–0.58)	0.97	IF
–Main provider labor nurse/midwife	649	92.1	81.1 (77.7–84.2)	27.5 (15.9–41.7)	80.4	0.54 (0.50–0.58)	0.87	IF
Skilled main provider† delivery	644	92.9	95.0 (92.9–96.6)	15.2 (6.3–28.9)	94.3	0.55 (0.51–0.59)	1.02	IF
–Main provider delivery doctor (ob–gyn)/ medical resident	644	11.8	82.9 (72.5–90.6)	89.3 (86.4–91.7)	19.3	0.86 (0.83–0.89)	1.63	AUC
–Main provider delivery nurse/midwife	644	81.1	86.2 (82.9–89.0)	73.0 (64.2–80.6)	75.0	0.80 (0.76–0.83)	0.93	Both
Main provider delivery student nurse	644	4.8	16.1 (5.5–33.7)	98.5 (97.2–99.3)	2.2	0.57 (0.53–0.61)	0.45	
Second and third stage labor:								
Episiotomy performed	545	18.2	82.8 (73.9–89.7)	90.4 (87.2–92.9)	22.9	0.87 (0.83–0.89)	1.26	AUC
Uterotonic received following delivery of placenta	552	2.4	53.9 (25.1–80.8)	40.8 (36.6–45.1)	59.1	0.47 (0.43–0.52)	25.1	
Immediate newborn postnatal care:								
Baby given to mother immediately after birth	611	57.6	66.5 (61.3–71.4)	49.0 (42.8–55.3)	59.9	0.58 (0.54–0.62)	1.04	IF
Baby placed immediately skin to skin on mother	602	16.3	78.6 (69.1–86.2)	21.0 (17.6–24.9)	78.9	0.50 (0.46–0.54)	4.85	
Baby placed immediately skin to skin on mother (2 item)†	596	16.2	26.8 (18.3–36.8)	70.3 (66.1–74.3)	29.2	0.49 (0.44–0.53)	1.80	
Breastfeeding within first hour of birth	551	53.0	88.4 (84.1–91.8)	37.1 (31.2–43.3)	76.4	0.63 (0.59–0.67)	1.44	
3 elements of essential newborn care (immediately dried, on mother's skin, breastfed within first hour)	506	9.3	70.2 (55.1–82.7)	28.3 (24.2–32.7)	71.5	0.49 (0.45–0.54)	7.70	

Table 4. Continued

Indicator	Total number	Observer prevalence (%)	Sensitivity (95% CI)	Specificity (95% CI)	Self–report survey estimate (%), based on sensitivity and specificity	AUC (95% CI)	IF	High accuracy (AUC > 0.70) & low bias (IF 0.75–1.25)
3 elements of essential newborn care (immediately dried, 2 item on mother's skin‡, breastfed within first hour)	501	9.2	19.6 (9.4–33.9)	74.3 (70.0–78.2)	25.2	0.47 (0.42–0.51)	2.7	
Low birthweight newborn (<2500g)	579	7.8	71.1 (55.7–83.6)	98.7 (97.3–99.5)	6.7	0.85 (0.82–0.88)	0.87	Both
Immediate postnatal care for mother:								
Palpates uterus 15 min after delivery of placenta	557	70.2	88.8 (85.2–91.7)	12.7 (8.0–18.7)	88.3	0.51 (0.46–0.55)	1.26	
First post–delivery exam, provider ask/ checks for bleeding	627	90.6	59.9 (55.7–63.9)	17.0 (8.4–29.0)	62.0	0.38 (0.35–0.42)	0.68	
First post–delivery exam, provider examines perineum	554	87.4	57.9 (53.3–62.3)	55.7 (43.3–67.6)	56.1	0.57 (0.53–0.61)	0.64	
First post–delivery exam, provider takes temperature	638	40.3	88.1 (69.3–80.3)	50.1 (45.0–55.3)	60.0	0.63 (0.59–0.66)	1.49	
First post–delivery exam, provider takes blood pressure	642	48.3	88.1 (85.1–92.5)	38.0 (32.3–43.5)	74.6	0.63 (0.59–0.67)	1.55	
First post–delivery exam, provider checks for involution	615	78.2	64.0 (59.6–68.3)	35.1 (27.0–43.8)	64.2	0.50 (0.46–0.54)	0.82	IF
Woman asked for pain relief medication during stay	638	10.5	35.8 (24.5–48.5)	68.3 (64.3–72.1)	32.1	0.52 (0.48–0.56)	3.06	
Woman received pain relief medication	640	17.5	85.7 (77.8–91.6)	46.2 (41.9–50.6)	59.4	0.66 (0.62–0.70)	3.39	
Maternal outcomes:								
Cesarean section (C/S) performed	651	1.4	93.1 (85.6–97.4)	98.8 (97.5–99.5)	13.5	0.96 (0.94–0.97)	1.01	Both
Reason for C/S– prolonged/obstructed labor	76	67.1	39.2 (25.8–53.9)	80.0 (59.3–93.2)	32.9	0.60 (0.47–0.70)	0.49	
Complications (any):	654	11.0	62.5 (50.3–73.6)	57.4 (53.3–61.4)	44.8	0.60 (0.56–0.64)	4.07	
– Hemorrhage	654	4.6	33.3 (17.3–52.8)	89.9 (87.3–92.2)	11.2	0.62 (0.58–0.65)	2.43	
– Prolonged labor	654	3.7	50.0 (29.1–70.9)	77.3 (73.8–80.5)	23.7	0.64 (0.60–0.67)	6.46	
– None	654	89.0	53.8 (49.6–57.9)	66.7 (54.6–77.3)	51.5	0.60 (0.56–0.64)	0.58	

CI – confidence interval; IF – inflation factor; AUC – receiver operating curve

Recommended indicators met both AUC and IF validation criteria.

*Validation analysis based on matched data, excluding 'Don't Know' responses. Sensitivity and specificity analysis was not performed for indicators that had fewer than 5 counts per cell.

†Skilled provider includes doctor (ob–gyn), medical resident or nurse/midwife.

‡Indicator constructed from two skin–to–skin items: (1) newborn placed against mother's chest after delivery and (2) newborn was lying naked against the mother's chest.

livery was a nurse/midwife, support person present during birth, low birthweight infant). Indicators with low value AUC results tended to be related to events immediately following the birth (eg, uterotonic received following delivery of the placenta) and postnatal health checks for the mother and newborn. For population–level bias, a total of 13 indicators had low bias (0.75<IF<1.25), 7 had moderate bias (0.5<IF<1.5), and 21 had large bias (IF<0.5 or IF>1.5). Indicators with large bias varied with respect to phase of labor and delivery, but those with the largest bias tended to have a low observed prevalence and be those that may require medical knowledge to report accurately (eg, experience of complications).

To assess women's ability to recall the type of provider who attended them, respondents were asked, "*Who was the main provider assisting you during delivery?*" There were sufficient cell counts to assess two provider categories with robust analysis: nurse/midwife and student nurse. The nurse/midwife indicator met both the high AUC and low IF criteria while the student nurse indicator had low individual accuracy (AUC = 0.57) and large bias (IF = 0.45). An indicator constructed in analysis that combines responses of "doctor", "medical resident" and "nurse/midwife" as "skilled attendants" had low individual accuracy (AUC = 0.55), primarily due to low specificity, and low population–level bias (IF = 1.02) [28]. Cross–tabulation results that compare women's reports to observers' reports on their main provider during delivery suggest a tendency for medical residents and nurse/midwives to be misclassified by women as doctors (**Table 5**).

Table 5. Cross–tabulation of main provider during delivery based on observer reports and women's responses

Self–report (number)	Observer report (number)							
	Doctor (obstetrician/gynecologist)	Medical resident	Medical intern	Nurse or midwife	Clinical officer	Student nurse	Other	Total
Doctor (obstetrician/gynecologist)	16	46	6	46	2	7	0	123
Medical resident	0	1	0	0	0	0	0	1
Medical intern	0	0	0	1	0	1	0	2
Nurse/midwife	2	7	1	450	3	17	3	483
Clinical officer	1	0	0	12	0	1	0	14
Student nurse	0	1	0	8	0	5	0	14
Support person	0	0	0	1	0	0	0	1
No one	0	0	0	2	0	0	0	2
Other	0	2	0	2	0	0	0	4
Don't know	1	3	0	0	0	0	0	4
Total	20	60	7	522	5	31	3	648

To illustrate the implications of indicator properties established in this study for other contexts, we plot the values of the predicted survey prevalence (Pr) of select indicators across all possible levels of intervention coverage (from 0 to 100%). **Figures 2** and **3** compare the predicted prevalence using the sensitivity and specificity calculated in this study (blue line), to perfect reporting accuracy assuming 100% sensitivity and specificity (black line) across all levels of coverage. Using the example of a high sensitivity and low specificity indicator such as "skilled birth attendance", these data demonstrate that in a high coverage setting the estimated survey–based prevalence from women's self–report more closely approximates the true prevalence while in low coverage settings, the estimated survey–based prevalence would greatly overestimate the true rate (**Figure 2**). For example, in a setting where the true prevalence of skilled attendance is 40%, rather than the 93% observed in this study (the red triangle), the estimated survey–based prevalence would exceed the true prevalence by 50 percentage points. In contrast, an indicator with both high sensitivity and high specificity, such as "cesarean operation", would generate a survey–based estimate that closely approximates the true prevalence across all coverage levels (**Figure 3**).

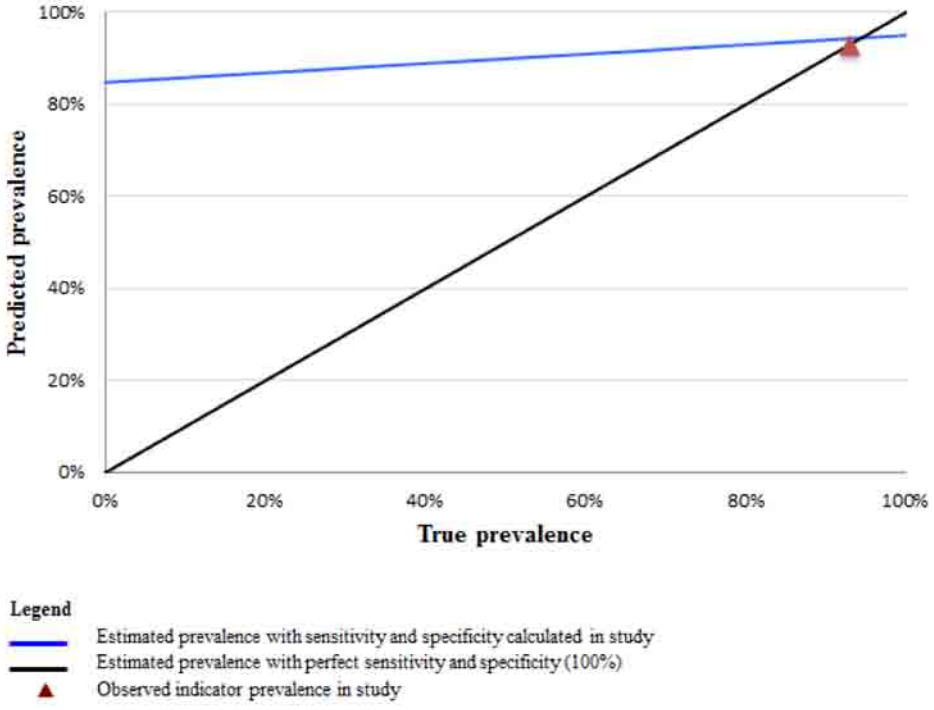


Figure 2. Predicted prevalence of skilled birth attendance based on sensitivity and specificity of women's reports by true prevalence.

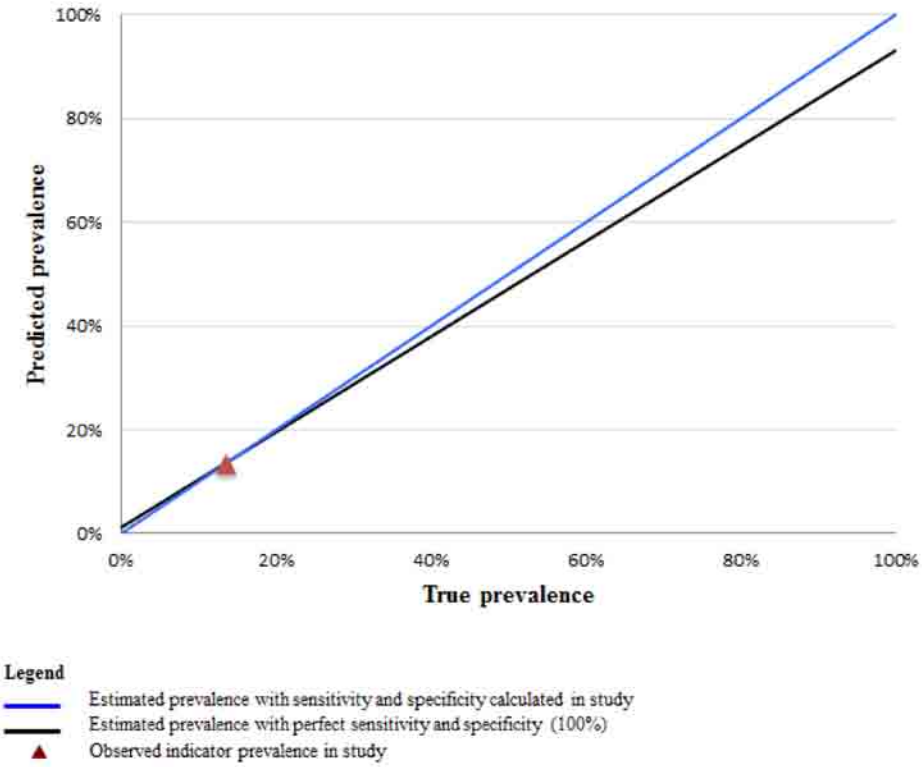


Figure 3. Predicted prevalence of caesarean operation based on sensitivity and specificity of women's reports by true prevalence.

DISCUSSION

This study provides validity results for 41 indicators of the quality of maternal and immediate newborn health care that are either currently in use or have the potential to be included in household surveys. Across phases of labor and delivery, we found indicators related to concrete, observable aspects of care or which reflected pain or concern were reported with higher accuracy. These results are consistent with previous studies which have found particularly memorable events, such as having a cesarean operation [11,22,29] and having a support person present [22], have high overall validity.

That a small number of the initial list of 82 indicators met both validation criteria is partly due to the fact that many preventative care interventions almost always occurred, and there was insufficient variation (ie, not enough cases in each cell) for robust analysis. For many preventative care indicators we found that most women accurately reported receiving the care (ie, high indicator sensitivity). For example, an indicator that is a proxy for receiving a uterotonic for the prevention of postpartum hemorrhage (ie, if an injection, IV medication or tablets were received in the first few minutes following birth), was accurately reported by nearly all women. Although the near universal implementation of this practice limited robust analysis, these results suggest some aspects of routine care can be accurately reported. However, given that there were few instances in which standard preventative interventions were not received, unless there was almost perfect negative classification by women, these indicators also tended to have low specificity. An alternate interpretation is that the observed pattern of high sensitivity and low specificity for many preventative practices may reflect "facility reporting bias" among women based on the expectation of receiving appropriate care. This finding has also been described in a study of women's reporting of maternal and child health care in China [11].

The potential for facility reporting bias may also be relevant for indicators on skilled birth attendance. Indicators measuring the assistance of a skilled provider had high sensitivity and low specificity for both labor and delivery. Women tended to underreport the presence of less skilled providers, such as student nurses, and over–report the presence of a doctor or obstetrician/gynecologist. The positive bias may also be due to differences in how women conceptualized who their "main" provider was. It is possible that women understood their "main" provider to be the attendant with the highest rank and who may have been deemed 'in–charge' of her care, while observers iden-

tified the primary provider as the attendant who administered the majority of the care to the woman.

Study findings also suggest the validity of some indicators may be dependent on context and question wording. Indicators that performed worse on the validity tests tend to be related to the timing or sequence of events, such as whether the newborn was placed skin–to–skin on the mother's chest immediately after delivery. A two–item question sequence that clarified the precise meaning of "skin–to–skin" greatly reduced women's overestimation of the practice compared to a one–item indicator (**Table 4**). These results are consistent with findings that women had difficulty reporting whether their newborn was placed skin–to–skin in a qualitative study of delivery and newborn care among women in Bangladesh and Malawi [30], but contrast with findings from a recent validation study in Mozambique [22]. The mixed results may be attributable to a longer recall period in the Mozambique study.

An influential aspect of the birth context was the type of delivery. Women who had a cesarean operation were less likely to be able to report on immediate newborn care than women with normal deliveries. This is reflected in high levels of "Don't Know" responses. This finding suggests that it may be worth excluding women with cesarean sections from questions about care immediately after birth in routine household surveys.

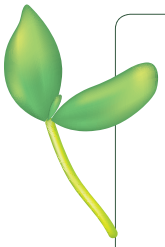
A number of indicators performed well on the IF test only; individual–level misclassification does not inherently signify that measurement at the aggregate level will be inaccurate [25]. In studies where the goal is to estimate the approximate population–based coverage of an indicator, false positives and false negatives may balance out to produce a close approximation of population level coverage (ie, indicators that meet the IF criteria alone). Knowing if an indicator's IF is large can inform when corrective methods may be needed to limit false positive reporting (eg, use of a two–item indicator).

Knowledge of whether an indicator is likely to be overestimated can also have significant programmatic implications. For example, where skilled birth attendance is over–reported, progress in scaling up the presence of higher cadre providers may not be as great as expected. It is important to recognize that the presence of a skilled provider is one aspect of receiving quality care, one which relies on the assumption that providers have received the necessary training to administer essential interventions and have acquired the competencies to address complications during childbirth. Additionally, even "skilled" providers may not be able to deliver adequate care if they do not have access to necessary equipment and supplies. Information on skilled attendance as reported by women should be corroborated with indicators on the content of care. When possible, we recommend that users also triangulate self–reported data on quality of care with other data sources such as information on stock–outs of essential medicines [4].

While a strength of this study is the use of direct observation as the reference standard, there are some limitations. Validation results are based on women seeking delivery in a large public hospital, and may not be generalizable to women who deliver in other types of facilities or at home. The lack of variation in hospital practices also limited the ability to analyze all of the indicators, which may have otherwise proven to be valid if we had collected data in a range of health institutions. Finally, our results inform a 'best case' scenario in terms of recall accuracy because women were interviewed shortly following delivery. To inform how recall changes over time, as well as to investigate women's understanding of concepts such as who their main provider was, a follow–up study is under way to re–interview women one year after delivery.

CONCLUSION

The measurement of the quality of maternal and newborn health care received in LMIC settings often relies on data from surveys of women. Little research has examined the validity of these indicators. The primary indicator of interest in this study–delivery by a skilled birth attendant–met validation criteria for reporting at the population level only and the results indicate that reporting accuracy may be particularly problematic where skilled birth attendance coverage is low. Indicator properties established here provide insight into contexts where indicator use is appropriate, and where modifications to data collection procedures or question construction may be warranted.



Acknowledgments: We acknowledge Hannah Taboada, who conducted a comprehensive literature review of indicators to be assessed in the study and developed the initial tools used in the study. In addition, we acknowledge the team who conducted the facility–based data collection including Jackie Kivunaga and Brian Mdawida. We appreciate the advice of the maternal health experts who advised on the study and are grateful to the women and providers who allowed us to share their experiences.

Funding: The Maternal Health Task Force through a grant from the Bill & Melinda Gates Foundation supported this study.

Authorship declaration: AKB, CW, and SR conceived of the study design and selected indicators for analysis: ICMJE criteria for authorship read and met. AKB, CW, KM, JK, CN and SR agree with manuscript and conclusions: CW, CN and JK collected the data. KM analyzed the data: KM. Wrote the paper: AKB CW KM.

Declaration of interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). The authors report no competing interests.

REFERENCES

1 Lozano R, Wang H, Foreman KJ, Rajaratnam JK, Naghavi M, Marcus JR, et al. Progress towards Millenium Development Goals 4 and 5 on maternal and child mortality: an updated systematic analysis. Lancet. 2011;378:1139-65. Medline:21937100 doi:10.1016/S0140-6736(11)61337-8

2 World Health Organization, United Nations Children's Fund, United Nations Population Fund, The World Bank. Trends in maternal mortality: 1990 to 2008. Geneva: World Health Organization, 2010. Available: https://www.unfpa.org/sites/default/files/pub-pdf/trends_matmortality90-08.pdf. Accessed: 8 December 2014.

3 Starrs A. The Safe Motherhood Action Agenda: Priorities for the next decade. New York: Inter-Agency Group for Safe Motherhood, Family Care International, 1977. Available: http://www.popline.org/node/526231. Accessed: 17 October 2013.

4 Hodgins S. Achieving better maternal and newborn outcomes: coherent strategy and pragmatic, tailored implementation. Glob Health Sci Pract. 2013;1:146-53. Medline:25276527 doi:10.9745/GHSP-D-13-00030

5 Graham WJ, Bell JS, Bullough CH. Can skilled attendance at delivery reduce maternal mortality in developing countries? In: Lerberghe W, Kegels G, De Brouwere, V, editors. Antwerp: Stud Health Serv Org Policy, 2001, pp. 97-130.

6 Koblinksy M, Conroy C, Kursehy N, Stanton ME. S. J. Issues in programming for safe motherhood. Arlington, VA: John Snow, Inc., 2000. Available: http://pdf.usaid.gov/pdf_docs/Pnack513.pdf. Accessed: 14 October 2013.

7 Scott S, Ronsmans C. The relationship between birth with a health professional and maternal mortality in observational studies: a review of the literature. Trop Med Int Health. 2009;14:1523-33. Medline:19793070 doi:10.1111/j.1365-3156.2009.02402.x

8 Lohela TJ, Campbell OM, Gabrysch S. Distance to care, facility delivery and early neonatal mortality in Malawi and Zambia. PLoS ONE. 2012;7:e52110. Medline:23300599 doi:10.1371/journal.pone.0052110

9 Hounton S, Menten J, Ouedraogo M, Dubourg D, Meda N, Ronsmans C, et al. Effects of a skilled care initiative on pregnancy-related mortality in rural Burkina Faso. Trop Med Int Health. 2008;13 Suppl 1:53-60. Medline:18578812 doi:10.1111/j.1365-3156.2008.02087.x

10 Stewart MK, Boerma JT, Ginneken JV. Assessment of the quality of data on maternity care in the Demographic and Health Surveys, 1986-1989. In: An assessment of the quality of health data in DHS-1 Surveys. Methodological Reports 2. Calverton, Maryland: Macro International Inc. 1994; 59-77. Available: http://www.dhsprogram.com/pubs/pdf/MR1/MR1.pdf. Accessed: 7 April 2015.

11 Liu L, Li M, Yang L, Ju L, Tan B, Walker N, et al. Measuring coverage in MNCH: A validation study linking population survey derived coverage to maternal, newborn and child health care records in rural China. PLoS ONE. 2013;8:e60762. Medline:23667429 doi:10.1371/journal.pone.0060762

12 Troude P, L'Hélias LF, Raison-Boulley AM, Castel C, Bouyer J, de La Rouchebrochard E. Perinatal factors reported by mothers: do they agree with medical records? Eur J Epidemiol. 2008;23:557-64. Medline:18560979 doi:10.1007/s10654-008-9268-9

13 Elkadry E, Kenton K, White P, Creech S, Brubaker L. Do mothers remember key events during labor? Am J Obstet Gynecol. 2003;189:195-200. Medline:12861162 doi:10.1067/mob.2003.371

14 Casey R, Rieckhoff M, Beebe SA, Pinto-Martin J. Obstetric and perinatal events: The accuracy of maternal report. Clin Pediatr (Phila). 1992;31:200-4. Medline:1563192 doi:10.1177/000992289203100402

15 Quigley MA, Hockley C, Davidson LL. Agreement between hospital records and maternal recall of mode of delivery: Evidence from 12,391 deliveries in the UK Millenium Cohort Study. BJOG. 2007;114:195-200. Medline:17166217 doi:10.1111/j.1471-0528.2006.01203.x

REFERENCES

16 Kenya National Bureau of Statistics (KNBS). Kenya Demographic and Health Survey 2014. Key Indicators. Nairobi, Kenya: KNBS, 2015. Available: http://dhsprogram.com/pubs/pdf/PR55/PR55.pdf. Accessed: 6 April 2015.

17 National AIDS and STI Control Programme, Ministry of Public Health and Sanitation, Kenya. Guidelines for HIV Testing and Counselling in Kenya, 2008. Available: http://www.ilo.org/wcmsp5/groups/public/—ed_protect/—protrav/—ilo_aids/documents/legaldocument/wcms_127533.pdf. Accessed: 4 April, 2013.

18 Ministry of Health. Kenya. Kenya National Guidelines for Research and Development of HIV/AIDS Vaccines, 2005. Available: http://www.globalgiving.org/pfil/1108/projdoc.pdf. Accessed: 4 April 2013.

19 Ministry of Health, Division of Reproductive Health, Kenya. National Reproductive Health Research Guidelines, 2006. Available: https://www.k4health.org/toolkits/kenya-health/national-reproductive-health-research-guidelines. Accessed: 4 April 2013.

20 Austin A, Langer A, Salam SA, Lassi ZS, Das JK, Bhutta ZA. Approaches to improve the quality of maternal and newborn health care: an overview of the evidence. Reprod Health. 2014;11 Suppl 2:S1. Medline:25209614

21 StataCorp. Stata. College Station, TX, USA: StataCorp.

22 Stanton CK, Rawlins B, Drake M, dos Anjos M, Cantor D, Chongo L, et al. Measuring coverage in MNCH: Testing the validity of women's self-report of key maternal and newborn health interventions during the peripartum period in Mozambique. PLoS ONE. 2013;8:e60694. Medline:23667427 doi:10.1371/journal.pone.0060694

23 Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143:29-36. Medline:7063747 doi:10.1148/radiology.143.1.7063747

24 Eng J. Receiver operating characteristic analysis: a primer. Acad Radiol. 2005;12:909-16. Medline:16039544 doi:10.1016/j.acra.2005.04.005

25 Campbell H, Biloglav Z, Rudan I. Reducing bias from test misclassification in burden of disease studies: use of test to actual positive ratio- new test parameter. Croat Med J. 2008;49:402-14. Medline:18581619 doi:10.3325/cmj.2008.3.402

26 Vecchio TJ. Predictive value of a single diagnostic test in unselected populations. N Engl J Med. 1966;274:1171-3. Medline:5934954 doi:10.1056/NEJM196605262742104

27 Eisele TP, Rhoda DA, Cutts FT, Keating J, Ren R, Barros A, et al. Measuring coverage in MNCH: accuracy of measuring diagnosis and treatment of childhood malaria from household surveys in Zambia. PLoS Med. 2013;10:e1001417. Medline:23667337 doi:10.1371/journal.pmed.1001417

28 Adegoke A, Utz B, Msuya SE, van den Broek N. Skilled birth attendants: who is who? A descriptive study of definitions and roles from nine sub Saharan African countries. PLoS ONE. 2012;7:e40220. Medline:22808121 doi:10.1371/journal.pone.0040220

29 Tunçalp O, Stanton C, Castro A, Adanu R, Heymann M, Adu-Bonsaffoh K, et al. Measuring coverage in MNCH: validating women's self-report of emergency cesarean sections in Ghana and the Dominican Republic. PLoS ONE. 2013;8:e60761. Medline:23667428 doi:10.1371/journal.pone.0060761

30 Yoder P, Rosato M, Mahmud R, Fort A, Rahman F, Armstrong A, et al. Women's recall of delivery and neonatal care: a study of terms, concepts and survey questions. Calverton, Maryland: ICF Macro, 2010. Available: https://dhsprogram.com/pubs/pdf/QRS17/QRS17.pdf. Accessed: 9 September 2014.

journal of global health

PAPERS

Community perspectives on HIV, violence and health surveillance in rural South Africa: a participatory pilot study

Nitya Hullur[1], Lucia D'Ambruoso[1,2], Kerstin Edin[2,3], Ryan G Wagner[2,3,4], Sizzy Ngobeni[3], Kathleen Kahn[2,3], Stephen Tollman[2,3,4], Peter Byass[1,2,3]

[1] Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, Scotland, UK
[2] Umeå Centre for Global Health Research, Division of Epidemiology and Global Health, Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden
[3] MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
[4] INDEPTH – An International Network for the Demographic Evaluation of Populations and Their Health, Accra, Ghana

Background South Africa faces a complex burden of disease consisting of infectious and non–communicable conditions, injury and interpersonal violence, and maternal and child mortality. Inequalities in income and opportunity push disease burdens towards vulnerable populations, a situation to which the health system struggles to respond. There is an urgent need for health planning to account for the needs of marginalized groups in this context. The study objectives were to develop a process to elicit the perspectives of local communities in the established Agincourt health and socio-demographic surveillance site (HDSS) in rural north–east South Africa on two leading causes of death: HIV/AIDS and violent assault, and on health surveillance as a means to generate information on health in the locality.

Methods Drawing on community–based participatory research (CBPR) methods, three village–based groups of eight participants were convened, with whom a series of discussions were held to identify and define the causes of, treatments for, and problems surrounding, deaths due to HIV/AIDS and violent assault. The surveillance system was also discussed and recommendations generated. The discussion narratives were the main data source, examined using framework analysis.

Results The groups identified a range of social and health systems issues including risky sexual health behaviors, entrenched traditional practices, alcohol and substance abuse, unstable relationships, and debt as causative. Participants also explained how compromised patient confidentiality in clinics, insensitive staff, and a biased judicial system were problematic for the treatment and reporting of both conditions. Views on health surveillance were positive. Recommendations to strengthen an already well–functioning system related to maintaining confidentiality and sensitivity, and extending ancillary care obligations.

Conclusion The discussions provided information not available from other sources on the social and health systems processes through which access to good quality health care is constrained in this setting. On this basis, further CBPR in routine HDSS to extend partnerships between researchers, communities and health authorities to connect evidence with the means for action is underway.

Correspondence to:
Lucia D'Ambruoso
R1.078, Polwarth Building
Institute of Applied Health Sciences
University of Aberdeen
Foresterhill Health Campus
Aberdeen AB252ZD
UK
lucia.dambruoso@abdn.ac.uk

South Africa is in a state of a health transition, facing a burden of infectious diseases, characterised by high levels of HIV/AIDS and tuberculosis, emerging epidemics of non–communicable diseases (NCDs) including mental illness, extremely high rates of mortality owing to violence and injuries, as well as existing burdens of maternal, neonatal and child deaths. There are considerable pressures on the health system to deal with a complex and dynamic quadruple burden of disease and mortality [1–4]. This paper reports on a study concerned with two leading causes of death in the country: HIV/AIDS and violent assault.

HIV/AIDS has characterized South Africa's health profile for four decades. Despite constituting 0.7% of the world population, South Africa accounts for 17% of the global burden [5], with an estimated 6.4 million people infected [6]. The distribution of the burden is highly unequal. Prevalence rates in black populations are 40–50 times that of whites, 18–20 times that of Indians and Asians and five–to–six times that of coloreds [6,7]. And in adolescents, risks are eight times higher in females vs males [6]. A host of additional social and structural drivers–mobile populations, over–crowded settlements and exploitative migrant labor–contribute to HIV/AIDS remaining a critical public health challenge [8].

Following an initial period of denial by the Government over the epidemic in the early 2000s, access to antiretroviral therapy (ART) has expanded dramatically through donor initiatives and progressive health policies [9,10]. Currently South Africa operates the world's largest ART initiative, covering 1.8 million people [11]. Despite achievements, pronounced disconnects between policies and implementation, resulting from ineffective leadership, lack of accountability and inadequate financing, have hampered impacts [12].

Violence is a major cause of death in South Africa [13,14]. The murder of Reeva Steenkamp by the Paralympic champion Oscar Pistorius in 2013 and the Marikana miners' massacre in 2012 reflect the normalcy of violence at all levels of South African society. At population level, excess homicide is observed according to residence, age, sex and socio–economic status [15]. Studies also suggest that known community members commit almost half (44%) of violent assaults out of jealousy and anger, and often aggravated by alcohol and substance abuse [16,17]. Entrenched inequalities in income and opportunity clearly push disease burdens towards vulnerable populations in South Africa, a situation to which the health system struggles to respond.

Since the first democratic elections in 1994, South Africa has made radical health reforms: a constitutional commitment to the right to health and with stated aims for equity through universal health coverage [18,19]. Progress towards the Millennium Development Goals (MDGs) has been attributed to the expansion of primary care and free health services for expectant mothers and children under five years [18], and a national health insurance system is currently being implemented [11]. In spite of these achievements however, the health system is deeply unequal. Colonial and apartheid legacies persist in unaccountable governmental systems, inadequate stewardship and financing, inefficient management and insufficient resources [20]. Most recently, neoliberal macroeconomic structural adjustment policies that prioritize growth over redistribution have deepened the divide between public and private care [12,21].

In a context of complex disease burdens, multiple and intersectional health inequities and weak health systems, civil registration and vital statistics (CRVS) play a critical role [22]. Routine health information that is reliable and robust is a critical means to strategize, evaluate and monitor progress [23] as well as foster security and citizenship more broadly [24]. The issue requires special attention following estimates that over two–thirds of deaths worldwide pass without registration [25], with over three quarters of these belonging to regions in sub–Saharan Africa and South–East Asia [26]. Although South Africa has a vital registration system for births, deaths and medical cause of death that is comprehensive in relation to other countries in the region [27], the system does not allow for correction of misclassifications in death certificates and audits have identified errors in up to 94% of records on HIV/AIDS deaths [28].

There is an urgent need for the health of marginalized groups to be accurately represented in this context. The objectives were therefore to develop a method to elicit perspectives of local communities in an established health and demographic surveillance site (HDSS) in rural north–east South Africa on two leading causes of death: HIV/AIDS and violent assault, and on HDSS as the means through which health information is generated in the locality. The broader aim was to demonstrate the utility of routinely consulting communities in HDSS.

METHODS

Study setting

The study setting was the Medical Research Council (MRC)/Wits Rural Public Health and Health Transitions Research Unit in rural northeastern South Africa, which oversees the Agincourt Heath and Socio–Demographic Surveillance Site (HDSS). The Agincourt site was established in 1992 in response to an absence of vital information on rural populations in South Africa, and has conducted annual censuses since collecting information on vital events (births, deaths

bind migrations) in a population of approximately 110 000 occupying 21 000 households across 31 villages (**Figure 1**) [29,30]. Agincourt established the Learning, Information, dissemination and Networking with Communities (LINC) group in 2004 to enable community participation in research and governance. LINC enhances research quality through the feedback of research results to community stakeholders [31]. Through these activities, the Agincourt HDSS tracks population health over time, measures the impact of interventions, supports community research and addresses gaps in population health data [32].

Participatory approach

The research adopted a community–based participatory research (CBPR) approach [33,34]. This was based on the premises that deaths identified through routine health surveillance have social and health systems determinants, and the mechanisms through which these factors influence health outcomes can be reliably identified with local knowledge [35]. Given the time and resources available, communities participated in identifying and defining health problems only. Other than in terms of health surveillance, the communities did not formulate remedial actions, and these were not implemented or evaluated. The research was of a pilot nature exploring feasibility and providing formative data as a basis form which to develop fuller forms of participation in the study setting.

Participant recruitment

Three villages were selected in the surveillance area in which to convene the discussion groups on the bases of demographic variation (**Table 1**) and feasibility within the time and resources available. LINC staff then approached women of reproductive age (WRA), family members, traditional healers, religious leaders, community health volunteers, health workers, village officials, and community leaders in villages to convene discussion groups that broadly represented the community. To mitigate against any po-

Table 1. Characteristics of selected villages*

	Village		
	A	B	C
Number of households	1178	932	647
Population, total	6158	4827	3705
Population, male	3005	2305	1781
Population, female	3147	2522	1924
Population, children under 5	647	513	458
Population, children of school age	1911	1410	1167

*Source: Household data collected by the MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), June 2013 [36].

Figure 1. Agincourt Health and Socio-Demographic Surveillance Site (HDSS) in Bushbuckridge Municipality, Mpumalanga Province, South Africa.

tential bias as for result of social and power differentials in the groups, the group consisted of women only in one village (**Table 2**).

An introductory meeting was held where the purpose, planned activities and outputs of the study were described. Those willing to participate were enrolled, written consent was taken and written information on the study was provided. A total of 24 participants were recruited into three village–based groups of eight participants, which operated independently in a series of weekly discussions on four selected health conditions, plus the introductory meeting described above (**Table 3**). The conditions were selected on the bases of high prevalence (HIV prevalence is 45% among men and 46% among women aged 35–39 in the Agincourt HDSS [37] and mortality from violent deaths is "outstandingly high" [38]) and in terms of socio–cultural relevance [39–42].

Table 2. Composition of discussion groups

	Discussion group		
Participants*	A	B	C†
Women of reproductive age	1	1	2
Family members‡	2	2	2
Traditional healers	1	1	2
Religious leaders / elders	1	1	2
Community health volunteers	1	1	
Community officials	1	1	
Community health providers	1	1	
TOTAL	8	8	8
			24

*All participants recruited were 18 y or older. Although participants typically had >1 role in the community, one primary role per individual was adopted for the purposes of convening the focus groups. Primary roles were also confirmed with participants.It was acknowledged that people with working arrangements, particularly village health workers and village officials' availability for five consecutive weekly meetings could be compromised. We also acknowledged the ethical imperative of engaging participants who would otherwise be involved in earning income and/or in the provision of public services. Participant recruitment was based on the compositions above with a degree of pragmatism and flexibility towards those committing to the process, and with careful consideration of minimising disruption to local public services.

†Group C was a women–only group to mitigate against the power differentials arising from the heterogeneous constituency of the groups.

‡Close relative: parents, grandparents, siblings, children, in–laws, nieces, nephews and cousins.

Table 3. Schedule of focus group discussions

Week/topic	1	2	3	4	5	
Focus Group	Recruitment/ Introduction	Stroke	HIV/AIDS	Violent assault	Epilepsy and feedback	Total meetings, per group
A	A, 1	A, 2	A, 3	A, 4	B, 5	5
B	B, 1	B, 2	B, 3	B, 4	B, 5	5
C	C, 1	C, 2	C, 3	C, 4	C, 5	5
Total number of focus group discussions						15

Data collection

A qualitative approach to data collection and analysis was adopted to elicit the collective perspectives of the village–based groups on the relationships between medical problems and their social and health systems determinants. The focus group discussion (FGD) method was used to encourage participation, to capitalize on communication between participants and to explore people's knowledge to gain an understanding of the collective norms and attitudes surrounding the two conditions [43,44]. A series of five weekly FGDs of 1.5–2 hours were held in each of the three villages, 15 FGDs were held in total.

A senior qualitative investigator (SN) with detailed knowledge of the area facilitated the discussions. SN presented data gained via the annual census on the conditions to the groups and facilitated discussions on this basis. Topic guides were prepared for the meetings in which the conditions, their causes, treatments, and the means through which information on them was generated in the locality (ie, HDSS) were discussed. The discussions were audio recorded and translated from the local language xi–Tsonga into English and transcribed. Two investigators took observational field notes and provided generally assistance during the meetings (LD and KE).

Data analysis

The narratives and field notes were the main data sources. Towards the end of the data collection, the groups were presented with and discussed a preliminary analysis to determine the plausibility and relevance of early interpretations of the discussions (**Panel 1**). Following completion of the data collection, a detailed analysis of the discussion transcripts was undertaken using framework analysis (NH). Framework analysis is a flexible tool to analyze qualitative data with the aim of creating a descriptive overview of an entire data set [45]. This method involved familiarization and coding of the data followed by preparation of summaries/charts to map the range of views on the phenomena of interest [46]. The steps of the framework analysis are summarized in **Table 4**. NVivo software (QSR International, London, UK) was used for data management and coding.

Ethical considerations

Ethical considerations related to the research process and outcomes were integrated into the study design [47]. All participants gave informed consent that guaranteed anonymization of all identifiable data in study reports, and assured participants that they were free to leave the study at

Table 4. An adapted framework analysis approach [45,46]

Stage	Description
1. Immersion and organisation	An initial organisation of data according to pre–determined (deductive) categories, as well as to preliminary emergent (inductive) themes.
2. Development of coding frameworks	The development of thematic, or coding frameworks that resulted from Stage 1.
3. Application of coding frameworks	The thematic frameworks are applied to the data to code or index it. This is done iteratively, until no new themes emerge ("thematic saturation").
4. Preparation of thematic summary grids	Thematic summaries prepared: grids of dominant and recurrent themes prepared with related themes and sub–themes in columns and respondents (or groups of respondents) as rows. This allows large volumes of narrative data to be distilled, and allows for the identification of patterns within and between narratives.
5. Interpretation	Establishing associations between themes to construct descriptive and explanatory accounts of the phenomena of interest.

any time and for any reason. The study protocols were peer reviewed to determine local, methodological and substantive relevance. The Human Research Ethics Committee at the University of the Witwatersrand (clearance #M121039) and the Mpumalanga Province health authority also reviewed and approved the study protocol.

RESULTS

This section presents the results of the analysis illustrated with verbatim quotes from the transcripts of the FGDs.

Causes of HIV/AIDS related mortality

Participants reported that widespread financial insecurity, often coupled with material desires, encourages young people to trade sex for immediate financial gain, often at the expense of their long–term health. The participants also described how gendered roles and expectations in relationships, combined with the need to maintain social and financial security, constrain safe sex practices. The introduction of condoms was specifically reported to induce doubts among men about the fidelity of their partners, which threatened the relationships that are sources of social and financial support for many women.

"Poverty is one of the main issues when it comes to this disease. People still take risks because they want money to make a living." [Group A; Village Official]

"…Most men … don't believe that they should use a condom with a woman that they have a baby with or a wife." [Group A; Village Official]

Traditional practices were also reported as causative. A practice called *Milo* was noted in this regard. *Milo* is a traditional cleansing ceremony performed following the death of a husband. The ceremony involves widows having unprotected sex with a mentally ill man for 7–14 days to prepare for future sexual relationships. Due to the presence of few such individuals in communities, the same people were rotated among widows in villages for *Milo*. Due to the likelihood that women were widowed as a result of HIV/AIDS, *Milo* was identified as a route of transmission in communities.

"…they don't know the health status of [persons involved in the ceremony] and it might happen that he infect[s] my mother … because he does it [with] several people who needs the ceremony." [Group A; WRA]

Treatments for HIV/AIDS related mortality

The use of traditional medicine for HIV/AIDS was described in detail. Traditional healers were often portrayed negatively in these discussions, as money–minded and misguiding. Participants reported instances in which healers, after identifying symptoms of HIV/AIDS, advised patients to undertake training to become a traditional healer themselves. Participants expressed suspicions that this was a strategy to gain revenue, and one that discouraged individuals from seeking and receiving medical treatment.

"…They were taken to the traditional healers and some … say it's a call from the ancestors to become a *Sangoma* [traditional healer]. She will go there and start to be a healer but at the end the traditional healer will chase her when he saw that she is not getting better (participants laugh)." [Group B; WRA]

Participants also reported that HIV patients would often be diagnosed and treated for *Tindzaka*, a disease believed to develop due to unprotected sex (as with HIV/AIDS) when mourning a death. Participants described treatments for *Tindzaka* involving cutting the skin of the patient, putting a mixture of herbs in the cut, and making the patient inhale the smoke of the burnt herbs. The participants also described medical treatments and self–help actions known to help patients to cope with, and prevent worsening of, the condition. The latter included acceptance of the situation, familial support, positive changes in diet and lifestyle, reductions in alcohol consumption, and a complete stop on sexual activity.

"…As a family we need to give moral support to the person and show love to him, ensure that he eats and drinks the medication." [Group A; Village Official]

In terms of medical treatments, participants reported that ART had proved beneficial for the health of people living with HIV/AIDS. Many people however were reported to experience pain, nightmares and excessive hunger with ARTs, which constrained compliance in some cases.

"... ARTs are very painful when you are using them. That is why some are defaulting." [Group B; Village Official]

Problems with patient confidentiality were reported in all groups as a major issue with medical treatment for HIV/AIDS. Although some clinics counselled patients before disclosing test results, elsewhere providers reportedly disclosed patients' status using non–verbal gestures (**Figure 2**). Casual and stigmatising attitudes among nurses when dealing with the personal details of patients were also described. Participants appeared uninformed of the pathways through which disciplinary actions for such breaches of confidentiality could be initiated.

"…HIV/AIDS [is] associated with prostitution, then when at the clinic say those with files [indicating those who are HIV/AIDS positive] come to this side then we look at them and say these are the ones who are not behaving well." [Group B; Village Official]

"…our people don't know about the channels to be used when they want to lay a complaint." [Group A; Village Official]

Figure 2. Focus group discussion (FGD) participant showing hand gestures used by medical staff to disclose HIV/AIDS status. Permissions were secured from participants for the reproduction of this image.

The insensitive handling of patient information had pronounced consequences reported to create fear and discomfort among individuals and often resulted in delayed or no testing. Individuals in communities reportedly sought care long distances from their local facilities to avoid regular breaches in patient confidentiality. Delayed care and transportation costs were recounted as key factors in care seeking decisions in these circumstances.

"… Many people are dying in their houses without consulting the clinic because they don't have faith in the clinic." [Group A; Traditional Healer]

Participants also described how fears over loss of financial security, social and emotional support delayed people seeking testing and treatment. The reported behaviours of providers and patients underscored pronounced stigmatisation of people living with HIV/AIDS.

"... if he can find that he is HIV/AIDS he will think that he is dead, [participants laugh] he will live in fear you see." [Group B; Village Official]

The participants also described unintended consequences of social welfare schemes that make support payments to patients with CD4 counts below a certain threshold. Participants described how patients deliberately discontinue ART to continue to qualify for the welfare grant to safeguard a predictable income, but at the expense of continuing treatment.

"…Some when they realize that his CD4 count is getting better and he realize that they are going to take away the grant and he will struggle … he stops taking his treatment and he will remain sick." [Group B; Village Official]

Causes of mortality owing to violent assault

Poverty was also a recurring theme in the discussions on the causes of violent deaths. Participants explained how situations of poverty force individuals to take loans and incur debt and how the needs for financial and social security encourage women to have multiple partners. Non–repayment of credit and disloyalty in relationships were described to give rise to violent conflict with fatal consequences.

"… some people have a lot of credit and… he is not paying, like the loan sharks if you don't pay them they can beat you until you die." [Group A; WRA]

Participants also recounted excessive drinking and drug abuse as precursors to fatal assaults. These were reported to involve sexual assaults, robberies and criminal gangs resulting in general increase in violence in the community. The narratives also revealed how conflict between couples can result from wives publicly disrespecting their husbands. Alcohol and drug abuse were again reported to exacerbate conflicts.

"… drunk people will fight for something minor; they will even fight just because the other person stepped on his toe." [Group B; Village Official]

The women only group portrayed men as the cause of familial violence. Specific issues reported in this group were the discovery of dishonesty and affairs. The heterogeneous groups, by contrast, reflected on the social acceptability of male violence as a means by which men establish control and dominance within the household and wider community.

"… according to our tradition, a women is always a women and the man has the upper hand at home." [Group A; Village Official]

Conflicts over property ownership, asset distribution, food, money, lack of discipline and guidance between siblings who had lost parents were also reported. Additionally, favoritism of siblings by parents who were present was reported as a source of tension, especially in second marriages.

Treatments for mortality owing to violence

The groups described several traditional and religious practices employed to reduce conflict and violence in families. These included husbands drinking traditional herbs to induce calm. In one village, an instance where the pastor took the name of the Lord, holding the t–shirt of the violent husband to reduce his anger was recounted. The participants expressed confidence in the effectiveness of traditional and religious practices.

Non–disclosure of violence due to stigma and the need to maintain socio–economic security were also recounted. Many participants reported how assaults are hidden from the police with women using home–based therapies to treat wounds so as not to compromise household stability and income. These actions protected violent spouses, again often with fatal consequences.

"…They are scared that they will arrest their husband because they are the ones providing food on the table for the family. That is why some they keep it to themselves until they are beaten to death." [Village B; Village Official]

Views on hospital treatment for assault were generally positive. Participants reported that hospitals provided timely care and engaged with the authorities to report crimes and protect victims. Despite this however, poor quality services in public health facilities were also described. Participants reported insensitive behaviours of nurses, unskilled medical staff and rigid shifts often leading to long waiting hours, sometimes with fatal outcomes.

"… hospital staff will even put a policeman to guard him, this shows that they not only want to heal people but also want the law to take its course." [Group B; Pastor]

"… nurses at the hospital are very cruel and rude… they never cared, they said I must pick the person and lay him in bed and the person died." [Group C; Community Health Provider]

Ambulances were frequently reported as lacking, as were the high transportation costs and poor access to emergency services. Unwillingness to take injured persons to hospitals in private vehicles was expressed in some villages. The participants also reported that at the police station, men who report their wives face disrespectful behaviour from officers.

"…many people are scared of blood … when you call people with cars when they see that the person is heavily bleeding they won't take that person to hospital because they think he might die in their car." [Group C; Traditional Healer]

"…when they arrest a man they will always come to his house while he is in jail to get a police statement and they use that as an excuse to get your wife while you are away [participants laugh]." [Group B; Village Official]

Health surveillance

The discussions also sought views on the procedures and outcomes of longitudinal health surveillance in the communities. The participants acknowledged the benefits of health surveillance in terms of understanding and awareness of health issues in communities.

"Health surveillance staff can help us in understanding more about HIV/AIDS issues and also help those who are left behind in being able to understand more about the disease and if needs be they take the treatment in a correct manner." [Group A; Community Health Provider]

The participants developed recommendations for the surveillance system. These included educational, financial and employment support to families whose needs were identified through routine surveillance. Other suggestions referred to HDSS partnering with community social workers to provide guidance and improve awareness on health and disease in the community.

"…I think if you investigate the cause of death it will be much better if you can come back and offer assistance to the family in a sense of checking their health status." [Group C; WRA]

Participants also stated that field–workers should show sensitivity and patience during household surveys. Additionally, participants reported higher levels of acceptability with field–workers who were not known in their community. As well as for impartiality and confidentiality, this was reported to improve the validity of information provided.

"... you are from [an]other village and come to investigate about death I will tell you the truth because I don't know

you. It will be between you and me." [Group B; Family Member]

The participants stated that the feeding back results of routine surveillance should incorporate an individual approach to personally inform families of the outcomes of cause of death conclusions gained through elements such as Verbal Autopsy (VA). Participants also stated that one–to–one feedback should maintain strictly confidentiality, protecting families from potentially harmful consequences related to the disclosure of stigmatised conditions.

"…if you tell the whole community that so [individual] was killed by so [cause of death] it will be a disaster." [Group A; Pastor]

DISCUSSION

The community views on HIV/AIDS and violence were broadly consistent and common issues were identified across the groups. The discussions on HIV/AIDS revealed serious problems with respectful care, confidentiality and patient dignity, while the discussions on violence reflected a patriarchal society with pervasive use of violence as a means of establishing control, social power and position within households and the wider society. The discussions on both conditions revealed the extent to which economic and social insecurities and traditional beliefs influence health and health behaviours. Specific issues included: norms of unsafe sex, widespread prostitution, debt, acceptance of domestic violence, and stigma around disclosure.

These issues did not exist in isolation. The discussions revealed how, in a context of pervasive vulnerabilities, actions to maintain social support and position in the immediate term (eg, by not practising safe sex or not seeking of HIV testing or treatment) were often necessary to prioritise over actions to safeguard health in the longer term. These actions and norms were further reinforced by the lack of effective health system responses such as the denial of confidentiality of health status and lack of emergency transport. The results can therefore be considered in terms of convergent forms of disadvantage and exclusion that exert strong influences over people's ability to protect their health.

More generally, poverty was repeatedly reported as a root cause for the emergence, transmission and exponentiation of mortality owing to HIV/AIDS and violent assault. Literature on the social diagnosis approach [48,49] and fundamental cause approach [50] assert that poorer households face problems in the availability, accessibility, acceptability and affordability of health care. These issues were clearly observed in the problems reported including unavailable emergency transportation, far–away clinics, expensive transport rental services, poor quality of care reflected through long waiting hours, limited hospital resources and staff, and poor confidentiality and insensitive behaviours on the part of health providers.

The South African health system is deeply divided as a result of historical colonialism, apartheid and, more recently, macroeconomic policies imposing neoliberal structural adjustment [51,52]. A recent assessment of equity in the health system in South Africa concludes that despite progressive financing, the distribution of health benefits remains distinctly pro–rich [53]. Entrenched poverty and social inequality, divisions between public and private care, and disconnects between policy and ineffective implementation, have deteriorated public sector personnel and facilities. The implications of poverty as a root cause of mortality in an unequal health system, in which deep social norms of eligibility for care linked to ability to pay, are important and should be a focus of future research.

The discussions provided information not available from other sources on the social and health systems mechanisms through which access to good quality health care is constrained in this setting. The routine engagement of marginalised group in the development of health information, coupled with HDSS for measuring and attributing progress to interventions developed and implemented is a clear avenue for further research [54]. It is encouraging that South Africa has a constitutional commitment to the right to health that centralises community participation in primary health care [55] that has been institutionalised, albeit with variable success, in Community Health Committees in the Western Cape [55]. Further CBPR in the Agincourt HDSS is currently underway to extend the partnerships initiated in this study between communities, researchers and health authorities. The intention is to develop co-constructed practical knowledge built from multiple perspectives, which can be readily embedded in local policy context [56,57].

The community views on health surveillance were largely positive reflecting established public engagement in Agincourt. Suggestions for modifications to an already well–functioning system related to ensuring that surveillance is respectful of loss, grieving and mourning, and for confidentiality and sensitivity when discussing deaths of relatives at the individual level. Extending ancillary care obligations with the provision of support to families in situations of bereavement and impoverishment were also suggested. A recent study on the cultural acceptability of health surveillance supports this finding, recommending that HDSS sites prioritize community sentiments and traditions in data collection and dissemination [58]. Further work on the balance between collective utilities and the protection of individuals in routine surveillance will further strengthen the activities [59,60]

Strengths and limitations

The degree to which participatory principles were adopted in the study was limited. Given the time and resources available, communities participated in identifying and defining health problems only. Research questions and study designs were largely determined prior to contact with communities, as were the conditions that were discussed. Other than identifying modifications to the health surveillance system, the groups did not discuss remedial actions for services and neither implemented nor evaluated strategies to respond to the issues identified.

Despite limitations in the nature and extent of participation, the study demonstrated that consulting communities offers unique perspectives on the social and health systems components of mortality. Acknowledging principles of CBPR related to developing sustained and authentic partnerships and mutual agendas between communities and researchers [61], the study has served as a basis upon which to design and implement a larger participatory action research (PAR) process in Agincourt. This work will develop a methodology suitable for application in other locations that promotes empowerment and social inclusion in health systems, with capacity building and evidence–based advocacy [62].

Otherwise, the study was conducted in a defined area using qualitative methods and the findings may not necessarily be relevant to different contexts and settings. It is maintained however that qualitative enquiry seeks to provide authentic representations rather than generalizable findings [63,64], and that in terms of the process, CBPR is concerned with changing academic research paradigms through more inclusive collaboration [61], and building partnerships though formative, feasibility and pilot data [65], to which progress has been achieved.

The study also explored perceptions on sensitive issues that were discussed in groups. This may have been subject to limited disclosure and so risk of bias. The differences in perspectives between the discussion groups convened to represent the community vs those that consisted of women only were noteworthy however, and suggest that the engagements may have been to a sufficient degree, authentic and representative of collective viewpoints.

The combined inductive/deductive approach to data collection and allowed flexibility to reveal unanticipated aspects of the community's perspectives on HIV/AIDS, violence and health surveillance, as well as considering issues identified a priori. A notable result in this sense relates to the unanticipated consequences of the welfare grant awarded on the basis of CD4 counts and how this acts as a disincentive to treatment compliance. Finally, the presentation and discussion/confirmation of the preliminary analysis served as a validity/integrity check with the participants to ensure rigor while reflecting on the phenomena of interest (**Figure 3**).

CONCLUSIONS

Eliciting community views on the long–standing challenges of HIV and violence in South Africa provided information not available from other sources on the mechanisms through which social, economic and health systems factors influence

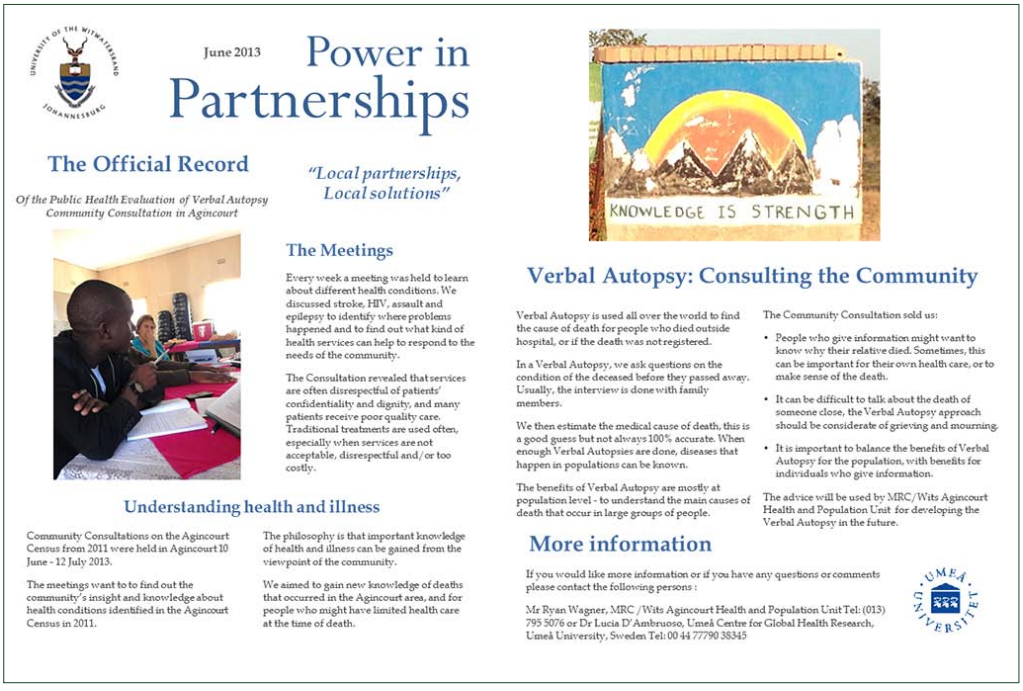
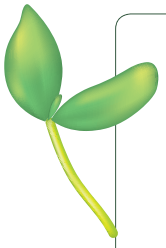


Figure 3. Preliminary analysis presented to discussion groups in the final meeting.

the accessibility and acceptability of heath care. The discussions also revealed how these factors combine and converge to seriously and negatively constrain the extent to which people can engage in behaviours that safeguard long–term health. Health planning must take account of the social aspects of mortality in service organisation and delivery in future. The results also indicate the need to address pervasive disenfranchisement of rural and poor communities.

In the context of HDSS, systematic documentation of population health and demographic data coupled with validation and co–production of health knowledge in a process connected to the health system at different levels may provide a means to improve evidence–based public health care services and address existing knowledge gaps. As stated by Scott–Samuel on Health Impact Assessments (HIA): "The identification and production of evidence that includes the interests of less powerful groups is a priority for HIA and would be furthered if a human rights–based method of HIA were developed." [66]. Further participatory work in the Agincourt HDSS is underway to explore the potential to enhance survey data, and to provide a basis from which to develop partnerships between researchers, communities and health authorities in order to connect robust evidence with the means for remedial action.



Acknowledgements: The authors would like to acknowledge the study participants who contributed time and information and made significant contributions to the analysis, and the staff at the MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.

Funding: The analysis presented in this paper is supported by Health Systems Research Initiative Development Grant from DFID/MRC/Wellcome Trust/ESRC (MR/N005597/1). The fieldwork was completed with the Umeå Centre for Global Health Research, with support from FORTE: Swedish Council for Health, Working Life and Welfare (grant No. 2006–1512). The Agincourt HDSS is supported by the School of Public Health, University of the Witwatersrand, South African Medical Research Council and the Wellcome Trust, UK (Grants 058893/Z/99/A; 069683/Z/02/Z; 085477/Z/08/Z; 085477/B/08/Z).

Authorship declaration: This work was made possible through the following contributions. NH led data analysis and prepared the manuscript; LD conceived of the study and study design, designed and led data collection, oversaw data analysis, and prepared the manuscript; KE contributed to data collection and analysis, and provided comments on the manuscript; SN led data collection and transcription, contributed to data analysis, and provided comments on the manuscript; RW contributed to study design, provided organisational support, oversaw data collection, contributed to analysis, and provided comments on the manuscript; KK, ST and PB oversaw study design, data collection and provided comments on the manuscript.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). The authors declare no competing interests.

REFERENCES

1 Perez AM, Yusuf OAA, Hofman K, Kalideen S, Maker A, Mokonoto D, et al. Establishing a Health Promotion and Development Foundation in South Africa. S Afr Med J. 2013;103:147-9. Medline:23472686 doi:10.7196/samj.6281

2 Mayosi BM, Flisher AJ, Lalloo UG, Sitas F, Tollman SM, Bradshaw D. The burden of non–communicable diseases in South Africa. Lancet. 2009;374:934-47. Medline:19709736 doi:10.1016/S0140-6736(09)61087-4

3 Boutayeb A, Boutayeb S. The burden of non–communicable diseases in developing countries. Int J Equity Health. 2005;4:2. Medline:15651987 doi:10.1186/1475-9276-4-2

4 World Health Organization. Non–communicable diseases country profile. Geneva: World Health Organization, 2014. Available: http://www.who.int/nmh/countries/en/. Accessed: 25 July 2014.

5 Mayosi BM, Benatar SR. Health and health care in South Africa—20 Years after Mandela. N Engl J Med. 2014;371:1344-53. Medline:25265493 doi:10.1056/NEJMsr1405012

6 Shisana O, Rehle T, Simbayi LC, Zuma K, Jooste S, Zungu N et al. South African National HIV Prevalence, Incidence and Behaviour Survey, 2012. Cape Town: HSRC Press, 2014. Available: http://www.hsrc.ac.za/en/research–data/view/6871 Accessed: 25 July 2014.

7 Kenyon C, Buyze J, Colebunders R. HIV prevalence by race co–varies closely with concurrency and number of sex partners in South Africa. PLoS ONE. 2013;8:e64080. Medline:23704973 doi:10.1371/journal.pone.0064080

8 AVERT. HIV and AIDS in sub–Saharan Africa. Brighton: AVERT, 2014. Available: http://www.avert.org/hiv–aids–sub–saharan–africa.htm. Accessed: 25 July 2014.

9 Wouters E, van Rensburg HCJ, Meulemans H. The National Strategic Plan of South Africa: what are the prospects of success after the repeated failure of previous AIDS policy? Health Policy Plan. 2010;25:171-85. Medline:19955092 doi:10.1093/heapol/czp057

10 Bekker LG, Venter F, Cohen K, Goemare E, Van Cutsem G, Boulle A, Wood R. Provision of antiretroviral therapy in South Africa: the nuts and bolts. Antivir Ther. 2014;19:105-16. Medline:25310359 doi:10.3851/IMP2905

REFERENCES

11 Mayosi BM, Lawn JE, van Niekerk A, Bradshaw D, Abdool Karim SS, Coovadia HM. Health in South Africa: changes and changes since 2009. Lancet. 2012;380:2029-43. Medline:23201214 doi:10.1016/S0140-6736(12)61814-5

12 Coovadia H, Jewkes R, Barron P, Sanders D, Mclntyre D. The health and health system of South Africa: historical roots of current public challenges. Lancet. 2009;374:817-34. Medline:19709728 doi:10.1016/S0140-6736(09)60951-X

13 Seedat M, Niekerk AV, Jewkes R, Suffla S, Ratele K. Violence and injuries in South Africa: prioritising an agenda for prevention. Lancet. 2009;374:1011-22. Medline:19709732 doi:10.1016/S0140-6736(09)60948-X

14 Ngobale KN, Ogunbanjo GA, Omole OB. Interpersonal violence amongst primary health care patients in Lesotho: A qualitative study of the reasons for assault. Afr J Primary Health Care & Family Medicine. 2013;5:1-8.

15 Otieno G, Marinda E. Bärnighausen, Tanser F. High rates of homicide in rural South African population (2000–2008): findings from a population–based cohort study. Popul Health Metr. 2015;13:20. Medline:26300697 doi:10.1186/s12963-015-0054-0

16 Statistics South Africa. Victims of crime survey 2012. Statistical release P0341. Pretoria: Statistics South Africa, 2012. Available: https://www.issafrica.org/crimehub/uploads/Victims–of–crime%20survey–2012.pdf. Accessed: 28 November 2014.

17 Africa Check. Factsheet South Africa: Official crime statistics for 2012/13. Johannesburg: Africa Check, 2013. Available: http://africacheck.org/factsheets/factsheet–south–africas–official–crime–statistics–for–201213/. Accessed: 25 July 2014.

18 Benatar SR. Health care reform and the crisis of HIV and AIDS in South Africa. N Engl J Med. 2004;351:81-92. Medline:15229313 doi:10.1056/NEJMhpr033471

19 McIntyre D. What healthcare financing changes are needed to reach universal coverage in South Africa? S Afr Med J. 2012;102:489-90. Medline:22668945

20 Economist Intelligence Unit (EIU). Healthcare outcomes index 2014. London: The Economist, 2014.

21 Kleinert S, Horton R. South Africa's health: departing for a better future? Lancet. 2009;374:759-60. Medline:19709726 doi:10.1016/S0140-6736(09)61306-4

22 Statistics South Africa. Mortality and causes of death in South Africa, 2010: Findings from death notification. Pretoria: Statistics South Africa, 2013. Available: http://www.statssa.gov.za/publications/P03093/P030932010.pdf. Accessed: 27 July 2014.

23 Bird J, Byass P, Kahn K, Mee P, Fottrell E. A matter of life and death: practical and ethical constraints in the development of a mobile verbal autopsy tool. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: ACM 1489–98, 2013.

24 Shibuya K, Gilmour S. Civil registration as a means to promote human security. Lancet. 2015;386:e14-5. Medline:25971221 doi:10.1016/S0140-6736(15)60765-6

25 World Health Organization. Civil registration: why counting births and deaths is important Fact sheet N°324. Geneva: World Health Organization, 2014. Available: http://www.who.int/mediacentre/factsheets/fs324/en/. Accessed: on 31 July 2014.

26 World Health Organization. PMNCH Knowledge summary #17 Civil Registration and Vital Statistics (CRVS). Geneva: World Health Organization, 2012. Available: http://www.who.int/pmnch/knowledge/publications/summaries/knowledge_summaries_17_civil_registration/en/. Accessed: 23 July 2014.

27 Joubert J, Rao C, Bradshaw D, Vos T, Lopez AD. Evaluating the quality of national mortality statistics from civil registration in South Africa, 1997–2007. PLoS ONE. 2013;8:e64592. Medline:23724066 doi:10.1371/journal.pone.0064592

28 Birnbaum JK, Murray CJ, Lozano R. Exposing misclassified HIV/AIDS deaths in South Africa. Bull World Health Organ. 2011;89:278-85. Medline:21479092 doi:10.2471/BLT.11.086280

29 Kahn K, Collinson MA, Gómez–Olivé FX, Mokoena O, Twine R, Mee P, et al. Profile: Agincourt Health and Socio–demographic Surveillance System. Int J Epidemiol. 2012;41:988-1001. Medline:22933647 doi:10.1093/ije/dys115

30 MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt) Johannesburg. University of the Witwatersrand, 2015. Available: http://www.agincourt.co.za/index.php/about/. Accessed: 25 July 2014.

31 Twine R, Tibane J, Khosa A, Kahn K. Building a research smart community (or walking the tightrope). Johannesburg: University of the Witwatersrand, 2011.

32 Kahn K, Tollman SM, Collinson MA, Clark SJ, Twine R, Clark BD, et al. Research into health, population and social transitions in rural South Africa: Data and methods of the Agincourt Health and Demographic Surveillance System. Scand J Public Health Suppl. 2007;69:8-20. Medline:17676498 doi:10.1080/14034950701505031

33 Arnstein SR. A ladder of citizen participation. J Am Inst Plann. 1969;35:216-24. doi:10.1080/01944366908977225

34 Rifkin SB. Paradigms lost: toward a new understanding of community participation in health programmes. Acta Trop. 1996;61:79-92. Medline:8740887 doi:10.1016/0001-706X(95)00105-N

35 Popay J, Williams G. Public Health research and lay knowledge. Soc Sci Med. 1996;42:759-68. Medline:8685744 doi:10.1016/0277-9536(95)00341-X

36 MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt). Routine household data June 2013. Mpumalanga; Agincourt Health and Socio-Demographic Surveillance Site (HDSS), 2013.

37 Gómez–Olivé FX, Angotti N, Houle B, Klipstein–Grobusch K, Kabudula C, Menken J, et al. Prevalence of HIV among those 15 and older in rural South Africa. AIDS Care. 2013;25:1122-8. Medline:23311396 doi:10.1080/09540121.2012.750710

REFERENCES

38 Garenne M, Kahn K, Tollman S, Gear J. Causes of death in a rural area of South Africa: an international perspective. J Trop Pediatr. 2000;46:183-90. Medline:10893926 doi:10.1093/tropej/46.3.183

39 Maredza M, Bertram MY, Tollman SM. Disease burden of stroke in rural South Africa: an estimate of incidence, mortality and disability adjusted life years. BMC Neurol. 2015;15:54. Medline:25880843 doi:10.1186/s12883-015-0311-7

40 Williams N, Nefdt WM, Wilmshurst JM. Epilepsy South Africa: Turning obstacles into true potential. Epilepsia. 2015;56:184-7. Medline:24995481

41 Wagner RG, Ngugi AK, Twine R, Bottomley C, Kamuyu G, Gómez–Olivé FX, et al. Prevalence and risk factors for active convulsive epilepsy in rural northeast South Africa. Epilepsy Res. 2014;108:782-91. Medline:24582322 doi:10.1016/j.eplepsyres.2014.01.004

42 Mangena–Netshikweta ML. Perceptions about epilepsy in the Limpopo Province of the Republic of South Africa. Curationis. 2003;26:51-6. Medline:15027278

43 Kitzinger J. Qualitative research. Introducing focus groups. BMJ. 1995;311:299-302. Medline:7633241 doi:10.1136/bmj.311.7000.299

44 Krueger RA, Casey MA. Focus groups: a practical guide for applied research. 3rd Edition. Thousand Oaks, California: Sage Publications, 2000.

45 Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi–disciplinary health research. BMC Med Res Methodol. 2013;13:117. Medline:24047204

46 Pope C, Mays N. Qualitative research in health care. Oxford; Blackwell Publishing, 2006.

47 Nuffield Council on Bioethics. The ethics of research related to healthcare in developing countries: follow–up paper. London: Nuffield Council on Bioethics, 2005.

48 Brown P, Lyson M, Jenkins T. From diagnosis to social diagnosis. Soc Sci Med. 2011;73:939-43. Medline:21705128 doi:10.1016/j.socscimed.2011.05.031

49 Ghafournia N. Culture, domestic violence and intersectionality beyond the dilemma of cultural relativism and universalism. Int J Crit Cult Stud. 2014;11:23-32.

50 Phelan JC, Link BG, Tehranifar P. Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. J Health Soc Behav. 2010;51 Suppl.:S28-40. Medline:20943581 doi:10.1177/0022146510383498

51 Mooney GH, McIntyre DE. South Africa: a 21st century apartheid in health and health care? Med J Aust. 2008;189:637-40. Medline:19061458

52 Ruff B, Mzimba M, Hendrie S, Broomberg J. South Reflections on health–care reforms in South Africa. J Public Health Pol. 2011;32:S184-S192. Medline:21730990 doi:10.1057/jphp.2011.31

53 Ataguba JE, McIntyre D. Paying for and receiving benfits from health services in South Africa: is the health system equitable? Health Policy Plan. 2012;27:i35-45. Medline:22388498 doi:10.1093/heapol/czs005

54 Bishai D, Ghaffar A, Kelley E, Kieny MP. Honouring the value of people in public health: a different kind of p–value. Bull World Health Organ. 2015;93:661-2. Medline:26478631 doi:10.2471/BLT.14.149369

55 Meier BM, Pardue C, London L. Implementing community participation through legislative reform: a study of the policy framework for community participation in the Western Cape province of South Africa. BMC Int Health Hum Rights. 2012;12:15. Medline:22920557 doi:10.1186/1472-698X-12-15

56 Loewenson R, Laurell AC, Hogstedt C, D'Ambruoso L, Shroff Z. Participatory action research in health systems: a methods reader. Harare: EQUINET, TARS, AHPSR, WHO, IDRC, 2014. Available: www.equinetafrica.org/bibl/docs/PAR%20Methods%20Reader2014%20for%20web.pdf. Accessed: 12 October 2014.

57 D'Ambruoso L. When communities help authorities tally births and deaths, health care equalizes. The Conversation Media Group, October 2015. Available: https://theconversation.com/when–communities–help–authorities–tally–births–and–deaths–health–care–equalises–48507. Accessed: 12 October 2015.

58 Aborigo RA, Allotey P, Tindana P, Azongo D, Debpuur C. Cultural imperatives and the ethics of verbal autopsies in rural Ghana. Glob Health Action. 2013;6:18570. Medline:24054087

59 Bird J, Fottrell E. The challenge of communicating causes of death and their uncertainty. Vienna: GIScience Uncertainty Workshop, 2014. Available: http://cognitivegiscience.psu.edu/uncertainty2014/papers/bird_communicating.pdf. Accessed: 1 March 2015.

60 Carrel M, Rennie S. Demographic and health surveillance: longitudinal ethical considerations. Bull World Health Organ. 2008;86:612-6. Medline:18797619 doi:10.2471/BLT.08.051037

61 Wallerstein NB, Duran B. Using community–based participatory research to address health disparities. Health Promot Pract. 2006;7:312-23. Medline:16760238 doi:10.1177/1524839906289376

62 World Health Organization. Changing mindsets: strategy on health policy and systems research. Geneva: World Health Organization, 2012.

63 Slevin E, Sines D. Enhancing the truthfulness, consistency and transferability of a qualitative study: utilising a manifold of approaches. Nurs Res. 2000;7:79-89.

64 Tobin GA, Begley CM. Methodological rigour within a qualitative framework. J Adv Nurs. 2004;48:388-96. Medline:15500533 doi:10.1111/j.1365-2648.2004.03207.x

65 Rhodes SD, Malow RM, Jolly C. Community–based participatory research (CBPR): a new and not–so–new approach to HIV/AIDS prevention, care, and treatment. AIDS Educ Prev. 2010;22:173-83. Medline:20528127 doi:10.1521/aeap.2010.22.3.173

66 Scott–Samuel A, O'Keefe E. Health impact assessment, human rights and global public policy: a critical appraisal. Bull World Health Organ. 2007;85:212-7. Medline:17486213 doi:10.2471/BLT.06.037069

journal of global health

Electronic supplementary material:
The online version of this article contains supplementary material.

PAPERS

Innovative financing for HIV response in sub–Saharan Africa

Rifat Atun[1], Sachin Silva[2], Mthuli Ncube[3], Anna Vassall[4]

[1] Harvard T.H. Chan School of Public Health, Harvard University, Boston, USA
[2] Health Policy Programme, Imperial College London, London, UK
[3] Blavatnik School of Government, Oxford University, Oxford, UK
[4] London School of Hygiene and Tropical Medicine, London, UK

Background In 2015 around 15 million people living with HIV were receiving antiretroviral treatment (ART) in sub–Saharan Africa. Sustained provision of ART, though both prudent and necessary, creates substantial long–term fiscal obligations for countries affected by HIV/AIDS. As donor assistance for health remains constrained, novel financing mechanisms are needed to augment funding domestic sources. We explore how Innovative Financing has been used to co–finance domestic HIV/AIDS responses. Based on analysis of non–health sectors, we identify innovative financing instruments that could be used in the HIV response.

Methods We undertook a systematic review to identify innovative financing instruments used for (1) domestic HIV/AIDS financing in sub–Saharan Africa (2) international health financing and (3) financing in non–health sectors. We analyzed peer–reviewed and grey literature published between 2002 and 2014. We examined the nature and volume of funds mobilized with innovative financing, then in consultation with leading experts, identified instruments that held potential for financing the HIV response.

Results Our analysis revealed three innovative financing instruments in use: Zimbabwe's AIDS Trust Fund (a tax/levy–based instrument), Botswana's National HIV/AIDS Prevention Support (BNAPS) International Bank for Reconstruction and Development (IBRD) Buy–Down (a debt conversion instrument), and Côte d'Ivoire's Debt2Health Debt Swap Agreement (a debt conversion instrument). Zimbabwe's AIDS Trust Fund generated US$ 52.7 million between 2008 and 2011, Botswana's IBRD Buy–Down generated US$ 20 million, and Côte d'Ivoire's Debt2Health Debt Swap Agreement generated US$ 27 million, at least half of which was to be invested in HIV/AIDS programs. Four additional categories of innovative financing instruments met our criteria for future use: (1) remittances and diaspora bonds (2) social and development impact bonds (3) sovereign wealth funds (4) risk and credit guarantees.

Conclusion A limited number of innovative financing instruments contributed a very modest share of funding toward domestic HIV/AIDS programs. Several innovative financing instruments successfully applied in other sectors could be used to augment domestic financing toward HIV/AIDS programmes.

Correspondence to:
Professor Rifat Atun, FRCP
Professor of Global Health Systems
Harvard T.H. Chan School of Public Health
Harvard University
Boston MA, USA
ratun@hsph.harvard.edu

By 2015, around 15 million individuals were accessing life–saving antiretroviral treatment (ART) [1]. Yet, the "AIDS transition" [2] is not in sight–in 2014, there were 36.9 million people living with HIV, 2 million new HIV infections and 1.2 million AIDS–related deaths [3].

The continuing HIV epidemic requires sustained investment to prevent new infections and to provide treatment to those who need ART now and in the future. However, as more individuals access ART, domestic obligations for financing HIV will increase, reaching an estimated US$ 190 billion between 2015 and 2050 and account for as high as 47% of the GDP in high prevalence sub–Saharan African countries such as Malawi [4], creating long term commitments that have to be met. The financial obligations have major implications for the affected countries and donors–not just in economic terms, but also in the way they manage financing of the HIV response, which requires empowering countries to take greater responsibility for managing funds from all sources.

The return of investment in HIV response as measured by benefit to cost ratio has been estimated at 280% [5], with substantial economic, social and health benefits reported by other studies [6], comparable to benefit to cost ratios reported for maternal and child health investments [7]. However, donor financing, which accounts for a large share of the HIV/AIDS funding in sub–Saharan Africa, is constrained due to the global economic crisis [8]. Domestic financing remains equally challenging, especially in high–prevalence low–income countries that are fiscally constrained. Compounding the financing challenges are inefficiencies in channeling and use of available funds, and the harmful asymmetry between the long–term financing needs for HIV and short–term replenishment cycles of donor institutions [9].

There are opportunities for increasing HIV financing, however. African economies are enjoying economic growth [10]. There are also untapped natural resources that could generate upwards of US$ 4 billion each year [11]. Additional fiscal space could be created in domestic budgets by improving efficiency in allocation both to and within HIV programmes and by co–financing HIV services with funding from other economic sectors [12]. Innovative financing could offer new sources of funding. Conceived as a funding source to meet the Millennium Development Goals (MDGs), innovative financing, which provided around US$ 6 billion in total in 2002–2012 [13], is increasingly an important source of funding for global health [14].

We explore how innovative financing could be used to co–finance the long–term HIV obligations by augmenting domestic contributions in sub–Saharan Africa. We analyze how different innovative financing instruments can be operationalised and the institutional arrangements needed for their effective use.

METHODS

We undertook a systematic review to identify domestic innovative financing used to augment funding for HIV programmes in sub–Saharan African countries. We then extended the search to identify innovative financing used in global health and non–health sectors.

We searched peer–reviewed literature and grey literature published between 2002 and 2014, the period when innovative financing gained prominence following the United Nations International Conference on Financing for Development [15], and the publication of the High Level Taskforce on Innovative International Financing for Health Systems [16]. We also searched web sites of development agencies, and international financing institutions for published reports and to gather data. We provide the framework for the search, a listing of the data sources and the search strings (Figure S1 in **Online Supplementary Document**). In qualifying a financing scheme as innovative, we that took into account the nature of the financing, institutional arrangements and the mode of financing (pooling, channeling and allocation of funds) [13].

We used predetermined criteria (Figure S2 in **Online Supplementary Document**) to systematically examine the nature and magnitude of funding that could be mobilised, the characteristics of the innovative financing and their suitability for long–term funding of HIV obligations. We sought views of leading experts involved in innovative financing and development finance (from the World Bank, International Monetary Fund, The African Development Fund, the Global to Fight AIDS, Tuberculosis and Malaria, and the Bill and Melinda Gates Foundation) and in domestic funding of HIV programmes in sub–Saharan countries (such as from Kenya, South Africa, and Ethiopia). We then categorised the innovative financing instruments with the most potential for raising new funds.

We did not consider funding generated from borrowing, funding from health insurance, reprioritisation of existing budgets and efficiency gains from improvements in health systems [17,18], as these approaches are used routinely in funding and managing health budgets.

We analyzed all search results using the Preferred Items for Systematic Reviews and Meta–Analysis (PRISMA) guidelines [19]. We present all monetary amounts in 2010 US Dollar (US$) equivalents based on World Bank official exchange rates [20] and the GDP deflator indices from the US Department of Commerce, Bureau of Economic Analysis, National Income Product Accounts Tables [21]. We report our findings by calendar year.

RESULTS

Innovative financing from taxes, levies and debt conversion instruments

The systematic review revealed three innovative financing instruments using national taxes and levies and debt con-

version for countries including debt buy–down currently in use for HIV/AIDS, namely: Zimbabwe's AIDS Trust Fund [22], Botswana's National HIV/AIDS Prevention Support (BNAPS) International Bank for Reconstruction and Development (IBRD) Buy–Down [23], and Côte d'Ivoire's Debt-2Health Debt Swap Agreement [24].

We briefly mention three other innovative financing instruments which were intended but were not established as functioning entities, were established and then terminated, or had failed to generate meaningful revenues.

Taxes and levies

Established in 2000, Zimbabwe's AIDS Trust Fund received proceeds from a 3% tax levied on formal sector employers and employees. Of the total funds collected, 50% was earmarked for ART programmes, 10% for prevention programmes, and the remainder toward program administration and support. The funding generated in 2000–2008 was not available as the estimated figures were distorted due to hyperinflation. In 2008–2011, the levy generated US$ 52.7 million (89.8% of the US$ 58.7 million in total domestic public HIV spending [25]), with US$ 5.7 million generated in 2009, US$ 20.5 million in 2010 and US$ 26.5 million in 2011 [26].

We identified other tax/levy–based instruments, which have been proposed, but not fully scaled up or implemented. For example, starting in fiscal year 2015/16, Kenya is establishing an HIV investment unit that will develop a model for resourcing a new HIV Trust Fund which will be created within the National AIDS Control Council, and will seek to mobilise resources from domestic and international sources, including matching funds, corporate social investments, debt swaps, infrastructure bonds and the informal sector [27], but the scheme is not yet operational.

The government of Uganda has proposed to establish a HIV Trust Fund (based on the Zimbabwe model) to provide sustainable financing for HIV programmes. The HIV Prevention and Control Act, 2014 has stipulated the source of income for the Fund from levies (2% of the total tax revenues) on beers, spirits, soft drinks and bottled water, in addition to income from international sources [28].

Tanzania has also established an AIDS Trust Fund in 2015 through the Tanzania Commission for Aids (Amendment) Act 2014 enacted by the Parliament in early 2015 to reduce donor dependence [29].

Trust Funds are a new and a promising new approach that pools funds revenues from multiple sources [30]. Here the innovation is less about the source of financing but more about the pooling and application of funds [13].

Tax and levy–based instruments have the potential increase in revenues in sub–Saharan Africa, (where trade taxes have declined since 1990s and income taxes have stagnated since 2000 [31]) and where revenue streams from extractive industries that could be taxed [11]. Revenues from taxes and levies can be earmarked or ring–fenced for HIV programs within public finance budget (as with Zimbabwe's AIDS Levy). However, ring–fencing or earmarking reduces fungibility of public funds [32] and may limit a government's ability to respond to unexpected shocks and to adjust allocations for short–term priorities [33].

Debt conversion

Buy–downs convert credits to grants, often with conditions [34]. Debt conversion instruments can be operationalised using a combination of schemes. With bilateral conversion, where the lender simply cancels all or a portion of the loan or credit, operationalisation means ensuring that the conditions for cancellation/forgiveness are met. In most instances, program monitoring is mediated via the respective ministries and debt cancellation occurs ex–post [34]. With trilateral conversion, if the lender cancels or forgives all or part of a loan with the expectation that the debtor invest that portion in a multilateral institution, a modified debt arrangement maybe required. If on the other hand, a third party donor purchases all or part of a loan either conditionally or unconditionally, the execution of the debt arrangement must occur ex–ante due to the inclusion of additional institutions [34].

Botswana's National HIV/AIDS Prevention Support IBRD Buy–Down of US$ 50 million was used to address implementation gaps in the domestic HIV response. The program also supported the implementation of a new national operational plan for scaling up prevention as a national "survival strategy". A buy–down of US$ 20 million, supported by the European Commission was later introduced in Botswana to reduce HIV prevalence in young adults with conversion predicated on the HIV program meeting performance objectives [35].

A similar but more recent buy–down under Debt2Health, a novel debt conversion instrument managed by the Global Fund [36], provided US$ 27 million to the domestic HIV response in Côte d'Ivoire. In exchange for the creditor (Germany) forgoing the US$ 27 million debt, Côte d'Ivoire was required to invest at least half of the proceeds on national HIV treatment and prevention programmes [37]. While Debt2Health has financed several HIV/AIDS programmes worldwide, Côte d'Ivoire is the only African country to benefit from the instrument.

Between 2001 and 2011, sub–Saharan Africa received approximately US$ 2 billion in concessional aid for HIV/AIDS programmes [38], which offers potential for the use of Debt2Health. Between 2011 and 2013, debt conversion programmes totaling US $45.7 million were signed using

Debt2Health [39], including US$ 27 million for Côte d'Ivoire.

We summarize in **Table 1** the key features of the innovative financing instruments discussed.

Airline levy and contributions from retail sales

In addition to innovative financing from taxes and levies and debt conversion international innovative financing such as Airline Solidarity Levy (Airline Levy) [40] and PRODUCT(RED)™ [41] have been used to generate financing for HIV/AIDS programmes. Airline Solidarity Levy is domestically implemented with revenues pooled and channeled via UNITAID. In Africa, distinct from the Airline Solidarity Levy, several countries have introduced their own versions of airline levies, including Cameroon, Congo, Madagascar, Mali, Mauritius and Niger, and is under consideration in Benin, Burkina Faso, Central African Republic, Gabon, Guinea, Kenya, Liberia, Namibia, Senegal, São Tomé and Principe, and Togo [40]. Product(RED) generates revenue through direct contributions via retail sales in Western countries the proceeds of which are then channelled to the Global Fund to Fight AIDS, Tuberculosis and Malaria [42].

Innovative financing for other services with potential future use for HIV/AIDS

Our analysis of global health and non–health sectors revealed four additional categories of innovative financing with the potential to expand fiscal space and provide additional funding for HIV/AIDS, namely: remittances and diaspora bonds; social and development impact bonds; sovereign wealth funds; and guarantees (See Panel in **Online Supplementary Document** for a summary of these categories).

Remittances and diaspora bonds

In 2014, remittances accounted for US$ 67.1 billion of the US$ 206.6 billion in external flows to sub–Saharan Africa. Nigeria was the largest recipient with US$ 21 billion in 2013 [31]. Remittances are additional and particularly attractive given their stability in comparison to direct foreign investment or private financing flows.

Remittances can be mobilised via issuance of diaspora bonds, which have been successfully used in India (since 1991), in Israel (since 1951) and in Sri Lanka (since 2001). A diaspora bond can be developed either directly by a government or a state–owned bank. The bond can be setup to raise revenues on a continuous basis (annual issuance) or on an on–demand basis (opportunistic issuance), and can be established as non–negotiable fixed rate or as floating rate bonds or notes in different denominations. Fixed rate bonds, which are inherently less volatile, provide increased predictability to financing [43].

Diaspora bonds offer several benefits to issuers. Due to investors' "patriotic motivations", the issuer could conceivably offer a lower rate of return, thereby gaining a "patriotic discount". The bonds also offer the issuer an opportunity to improve sovereign credit rating by creating a new funding source. For the investor, aside from meeting personal motivations, the bonds offer the flexibility to receive interest and principal in issuer currency, which can be routed to meet liabilities in the issuing country [43].

While revenues from India's diaspora bonds were used to offset the country's balance of payment crisis in 1991 [44],

Table 1. Innovative financing instruments in use in sub–Saharan Africa, those planned, and their features

Instrument name	Operational status	Year established	Instrument type	Source of revenue	Financing agent	Revenues generated/forecast
Botswana's National HIV/AIDS Prevention Support (BNAPS) IBRD Buy–Down	Operational	2009	Debt buy–down	Concessional or non–concessional debt	European Commission (up to US$ 20 million)	US$ 50 million (over five years)
Côte d'Ivoire's Debt2Health Debt Swap Agreement	Operational	2010	Debt swap	Concessional or non–concessional debt	Government of Germany	US$ 27 million
Zimbabwe's AIDS Trust Fund	Operational	2000	Levy	Formal sector employee and employer income	Domestic government	US$ 85.2 million (through 2012)
Kenya's National Aids Control Council (NACC) Tax	Not yet operational	Not yet established	Multiple sources	Multiple sources	Domestic government, augmented with funds from external sources	Unknown
Uganda's HIV Trust Fund	Not yet operational	Established by HIV and AIDS Prevention and Control Act, 2014	Tax	Tax revenue from alcohol, soft drinks and bottled water	Domestic government and international sources	Unknown
Tanzania's AIDS Trust Fund	Not yet operational	Established in 2015	Not specified	Not specified	Domestic government	Unknown

the Israeli diaspora bonds have financed public works such as seawater desalination, housing construction and communication infrastructure [43]. Though we did not uncover evidence of the use in health programs, the characteristics of these instruments suggest that it could be a viable source.

Social Impact Bonds and Development Impact Bonds

Impact bonds have gained prominence as a means to attract and "crowd in" additional private capital to address social challenges. Social Impact Bonds and Development Impact Bonds allow private investors to invest in social causes and generate suitable financial returns, contingent on the quality of the outcomes achieved [45]. In a Social Impact Bond the outcome payer is the government, while in a Development Impact Bond the outcome payer is a donor, development agency or a philanthropic foundation [42]. In incentivising payment based on the quality of the outcomes achieved, the Development Impact Bonds offer the potential to maximise impact underpinned by rigorous monitoring and evaluation. For the sponsoring government or agency (the bond issuer), Development Impact Bonds offer risk protection and potential overall cost savings in implementing programmes.

When implementing a social or a development impact bond however, governments typically focus on social programmes with proven interventions, which fall within investor risk thresholds, can generate cost savings, have well–defined target populations and have quantifiable impacts/outcomes. Thus in the health sector, social or development impact bonds are most appropriate for preventative rather than treatment interventions, as the cost of the latter should be met from operational budgets, and borrowing through bonds should be used to invest and not fund operational expenditures. The ethics of targeting interventions with easy to measure outcomes rather than those with the potential to meet most need is debatable, however.

While social impact bonds have proven successful in non–health sectors, including recidivism reduction in the United Kingdom [46], education and housing in the United States [47], they are yet to be implemented in for HIV prevention and control. Several case studies indicate the potential for the health sector however, including in implementing Treatment as Prevention programmes and tuberculosis control programmes in Swaziland [48] and malaria control programmes in Mozambique [49].

Sovereign Wealth Funds

Sovereign Wealth Funds (SWFs) are special purpose investment funds owned by governments, which are established for creating stable returns on the funds invested for macroeconomic stability, to meet contingent liabilities, and to withstand economic shocks [50]. In 2012, total assets of SWFs accounted for approximately US$ 3 trillion. Based on IMF and the Santiago Principles taxonomy [51], four types of SWFs are distinguishable, namely: stabilization funds, savings funds, development funds (reserve investment fund) and pension reserve funds. Stabilization funds absorb macroeconomic shocks due to commodity price volatility and other external events (eg, Chile's Economic and Social Stabilization Fund). Savings funds preserve wealth for future generations by transforming nonrenewable resources into monetary assets (eg, the Abu Dhabi Investment Authority). Development funds aim to finance social development and infrastructure (eg, Mubadala in United Arab Emirates) [52], whereas pension reserve funds aim to fund pension–related contingent–type liabilities (eg, Norway's Government Pension Fund).

African Sovereign Wealth Funds accounted for US$ 114.3 billion in 2009, approximately 3% of global sovereign wealth funds at the time. By 2012 there were 15 Sovereign Wealth Funds in Africa with the four largest sourced from oil and gas revenues and the fifth sourced from diamonds, minerals and other natural resources, but amounts used as development funds are unknown, due to scarcity of publicly available data [53].

Guarantees

Guarantees can be used to catalyze private financing by mitigating risks, especially those that are political, contractual or regulatory in nature. The largest volume of guarantees originated from the World Bank Group, which provided by 2013 US$ 4.5 billion as 37 guarantees across 30 countries [54]. The guarantees were sourced from International Development Association (IDA), IBRD, International Finance Corporation (IFC) [55] or the Multilateral Investment and Guarantee Agency (MIGA) [56], and structured as Partial Risk Guarantees, Partial Credit Guarantees or Policy Based Guarantees.

Partial Risk Guarantees support private sector investment including in public–private partnerships. Partial Credit Guarantees support commercial borrowing in support of public investment projects and Policy Based Guarantees support commercial borrowing for budget financing or reform programmes. Partial Risk Guarantees are available to all IBRD and IDA countries, while Partial Credit Guarantees and Policy Based Guarantees are only available to IBRD–eligible countries [54]. Aside from the World Bank and other multilateral development banks, private foundations such as the Bill & Melinda Gates Foundation also provide guarantees [57].

Guarantees offer several benefits to borrowers. The reduction of default risk improves potential of the country for securing loans and thereby stimulates additional investment.

Guarantees can also reduce the cost of capital due to lower interest rates afforded to the borrowing government via guarantor's credit worthiness (especially in the case of the World Bank due to the bank's AAA rating) [54]. Guarantees allow governments to share the risk of projects with the private sector. In the case of World Bank guarantees, capacity building is also afforded as an added benefit [58]. While guarantees are beneficial to borrowers they create risks to guaranteeing entities, which with the prevailing global economic crisis may limit the potential for expanding guarantees.

Between 2005 and 2012, the IBRD mobilised US$ 1.2 billion and IDA mobilised US$ 789 million as guarantees. The energy sector received the highest volume of financing (US$ 2 billion) with IFC or MIGA guarantees, with the largest proportion within the African region (US$ 1.1 billion). The health sector however, has yet to receive guarantees from the World Bank Group [54]. By contrast, the Gates Foundation has issued credit enhancement guarantees to enhance affordability of vaccines and health commodities amounting to around US$ 250 000, US$ 500 000 and US$ 400 000 in gross exposure in 2012, 2013 and 2014 respectively [56].

Borrowing increases future financial liabilities for countries, but by reducing the cost of borrowing, guarantees can significantly reduce the cost of servicing the debt burden. The benefits of guarantees can be further augmented if countries demonstrate economic returns from HIV treatment/prevention programmes via reduced HIV incidence and enhanced labour productivity that benefit the economy.

DISCUSSION

We identified limited use of innovative financing instruments in domestic financing of HIV programmes in sub–Saharan Africa. The findings suggest both an opportunity to augment domestic financing, but also a possible hindrance of innovative financing due to weak domestic political or regulatory climates – only three innovative financing instruments were in current use and operationalised to generate meaningful revenues.

The instruments that were successfully implemented were either based on debt conversion or taxes and levies, however with no new instruments that unlocked resources to generate additional predictable revenues from new sources beyond taxes/levies and debt. Innovative financing included both new financing instruments, but also innovative ways of using existing instruments for HIV–for example debt conversion and levies.

The revenues generated through tax/levy mechanisms (US$ 52.7 million in Zimbabwe), accounted for a relatively minor share of the domestic HIV budget [26]. The revenues generated from debt conversion were also modest (US$ 50 million in Botswana [23] and US$ 27 million in Côte d'Ivoire [24]). Similarly, despite the growth of bond issuance by countries of sub–Saharan Africa [10], there are no bond–based instruments for HIV, even though bonds have been used to finance malaria programmes in Mozambique [59].

While the relative absence of innovative financing instruments that generate new sources is a challenge, the nascent space for innovative innovative financing is also an opportunity as countries affected by HIV move to increase domestic share of the financing obligations for HIV. The Global Fund's country graduation and counterpart financing [60] and PEPFAR's multi–year partnership framework agreements [61] will encourage this transition. Recent evidence suggests that 12 high prevalence sub–Saharan African countries would be able to finance as much as 64% of future financing needs in 2014–2018 through economic growth [9]. Much of the additional financing in these countries would be from the expansion of domestic fiscal space from traditional sources [40], which could be further augmented by innovative financing, if constraints for introducing new funding instruments are overcome. In this context, innovative financing instruments that enable securitisation of future income streams offer the most immediate possibility of augmenting existing funding to accelerate the HIV response, though the risk of exaggerating already high future debt obligations has to be carefully considered. Securitisation is an area where guarantees could be leveraged, as guarantees enable the borrowing countries to substantially reduce the interest payable on the debt while reducing default risk to lenders–thereby unlocking new funds from private sector investors by making the benefit–risk calculus more attractive.

Taxes and levies continue to be promising sources of new revenue if implemented as a modest charge on high volume transactions (especially on goods such as tobacco and alcohol with well–proven health harm) to expand the government revenue base with revenues pooled and effectively committed. Diaspora bonds are hitherto untapped sources of funding for the HIV response, due to their long–term nature, which is aligned with the long–term obligations for HIV. However, apart from the diaspora bonds issued in India and Israel, their widespread uptake is limited [43].

Sovereign wealth funds offer potential new funding sources if they could be used to leverage private sector investments, through public–private partnerships to invest in health system infrastructure or in new ventures to create additional capacity for health service provision in the health system, where new providers could be contracted by the government to provide services.

Social impact bonds offer significant potential for mobilizing new and additional external resources from private sources. Social impact bonds would be particularly suitable

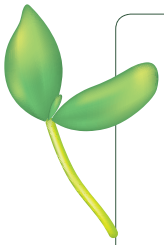
for financing preventative interventions that reduce future burden of disease, especially for effective interventions that are under–utilized or inefficiently delivered (such as prevention of mother to child transmission, harm reduction, voluntary male circumcision, condom distribution and use) to bring health, economic and social benefits and achieve returns beyond the costs. By transferring the risk of success to private investors that bring new funds and innovative service delivery models to the sector, the government (or the public sector funding entity) pays for successful outcomes. The investors, be it philanthropic foundations, high net worth individuals or socially conscious funders, receive financial return and social benefits.

To maximise the benefits of revenues generated from domestic innovative financing, the "partnership" arrangements between donors and HIV–affected countries must be reframed. In spite of the Paris Declaration on Aid Effectiveness and the Accra Agenda for Action [62], country ownership of donor funding and health programmes has yet to be achieved in most settings. Instead, the partnership remains disproportionately weighted toward donors, especially in countries where donor financing outweighs domestic finding. For example, in Uganda domestic financing for HIV accounted for US$ 53 million in 2013, whereas donor financing accounted for US$ 446 million [9]. In Malawi, the donor share of HIV funding was 98% of cumulative spend [63], although the appropriate balance of domestic and international financing is debated [64].

In restructuring the partnership, domestic governments' view that a duty lies with the donor must be balanced with a view that emphasizes mutual responsibility and accountability to ensure sustainable and predictable financing for HIV response. In low income countries with high prevalence of HIV, as HIV financing accounts for a large portion of the health budget and the GDP [9], however, long term financing consideration need to rest with the ministry of finance and the government in general, which have the responsibility for management of debt levels and priorities in the available fiscal space. Similarly, donor perceptions that shifting resources among recipient countries is acceptable, as long as outlay commitments are met, should also be overcome.

"Commitment Technologies", which can be utilized to enforce binding and credible commitments, both from donors and recipients, can capitalise on several key features of the status quo arrangements. Presently, donors reap the most benefit from efficiency gains in development assistance. Recipients on the other hand, are uncertain of benefits, as the returns are intangible in budgetary terms, although health benefits are real. If donors (or investors) make multi–year commitments to invest in a proportion of need, they stand to reap a share of efficiency improvements equal to participation. Similarly, if the incentives for domestic governments can be made more tangible, via the use of mechanisms such as social impact bonds[46], results–based financing [65] or cash–on–delivery [66] for example, where results to be delivered equate to reduction in incidence, then the prospects for reframing the partnership become promising.

The opportunity for innovative financing to augment domestic financing for HIV/AIDS is real and important given the magnitude of the long–term obligations in sub–Saharan Africa estimated at around US$ 180 billion [4]. Sustaining HIV response in the era of sustainable development with competing priorities makes the search for funding from innovative financing all the more pressing.



Acknowledgment: This paper was published as part of RethinkHIV. RethinkHIV is a consortium of senior researchers, funded by the RUSH Foundation, who evaluate new evidence related to the costs, benefits, effects, fiscal implications and developmental impacts of HIV interventions in sub–Saharan Africa.

Funding: The Rush Foundation and Harvard University.

Ethics approval: Not required.

Authorship declaration: RA conceived the study. RA wrote the first draft and subsequent drafts of the manuscript with SS and undertook the analysis with SS. AV and MN contributed to the revisions of the manuscript. All the authors have seen and approved the final manuscript. RA is the guarantor.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). The authors declare no financial relationships with any organizations that might have an interest in the submitted work in the previous four years; and no other relationships or activities that could appear to have influenced the submitted work.

REFERENCES

1 UNAIDS. UNAIDS announces that the goal of 15 million people on life–saving HIV treatment by 2015 has been met nine months ahead of schedule. 2015. Available at: http://www.unaids.org/en/resources/presscentre/pressreleaseandstatementarchive/2015/july/20150714_PR_MDG6report. Accessed: 25 January 2016.

2 Over M. Achieving an AIDS transition. 1st ed. Washington, D.C.: Center for Global Development, 2011.

3 UNAIDS. Core Epidemiology Slides. 2015. Available: http://www.unaids.org/en/resources/documents/2015/20150714_coreepidemiologyslides_ppt. Accessed: 25 January 2016.

4 Atun R, Chang AY, Ogbuoji O, Silva S, Resch S, Hontelez J, et al. Long term financing needs for HIV control in sub-Saharan Africa in 2015–2050: modelling study. BMJ Open. Forthcoming 2016.

5 Resch S, Korenromp E, Stover J, Blakley M, Krubiner C, Thorien K, et al. Economic returns to investment in AIDS treatment in low and middle income countries. PLoS ONE. 2011;6:e25310. Medline:21998648 doi:10.1371/journal.pone.0025310

6 Bor J, Tanser F, Newell M, Barnighausen T. In a study of a population cohort in South Africa, HIV patients on antiretrovirals had nearly full recovery of employment. Health Aff (Millwood). 2012;31:1459-69. Medline:22778335 doi:10.1377/hlthaff.2012.0407

7 Stenberg K, Axelson H, Sheehan P, Anderson I, Gülmezoglu A, Temmerman M, et al. Advancing social and economic development by investing in women's and children's health: a new Global Investment Framework. Lancet. 2014;383:1333-54. Medline:24263249 doi:10.1016/S0140-6736(13)62231-X

8 Healthdata. Financing global health 2012: the end of the Golden Age? 2010. Available: http://www.healthdata.org/policy–report/financing–global–health–2012–end–golden–age. Accessed: 25 January 2016.

9 Resch S, Ryckman T, Hecht R. Funding AIDS programmes in the era of shared responsibility: an analysis of domestic spending in 12 low–income and middle–income countries. Lancet Glob Health. 2015;3:e52-61. Medline:25539970 doi:10.1016/S2214-109X(14)70342-0

10 African Economic Outlook. African economic outlook 2014. Available: http://www.africaneconomicoutlook.org/en/news–events/article/african–economic–outlook–2014–95/. Accessed: 25 January 2016.

11 African Development Bank. Harnessing Africa's extractive industries for structural transformation and inclusive development. 2012. Available: http://www.afdb.org/en/aec–2012/programme/october–31–2012/harnessing–africa's–extractive–industries–for–structural–transformation–and–inclusive–development/. Accessed: 25 January 2016.

12 World Bank. Global economic prospects, January 2015: having fiscal space and using it. 2015. Available: https://openknowledge.worldbank.org/handle/10986/20758. Accessed: 25 January 2016.

13 Atun R, Knaul F, Akachi Y, Frenk J. Innovative financing for health: what is truly innovative? Lancet. 2012;380:2044-9. Medline:23102585 doi:10.1016/S0140-6736(12)61460-3

14 Healthdata. Global health funding reaches new high as funding priorities shift. 2016. Available: http://www.healthdata.org/news–release/global–health–funding–reaches–new–high–funding–priorities–shift. Accessed: 25 January 2016.

15 United Nations. International Conference on Financing for Development. 2003. Available: http://www.un.org/esa/ffd/ffdconf/. Accessed: 25 January 2016.

16 International Health Partnership. High Level Taskforce for Innovative International Financing of Health Systems. Working Group 2 Report. 2009. Available: http://www.internationalhealthpartnership.net/fileadmin/uploads/ihp/Documents/Results___Evidence/HAE__results___lessons/WORKING%20GROUP%202%20REPORT%20–%20Raising%20and%20Channeling%20Funds.pdf. Accessed: 25 January 2009.

17 Heller PS. The prospects of creating 'fiscal space' for the health sector. Health Policy Plan. 2006;21:75-9. Medline:16415338 doi:10.1093/heapol/czj013

18 Tandon A, Cashin C. Assessing public expenditure on health from a fiscal space perspective: Health, Nutrition and Population Discussion Paper. Washington D.C.: World Bank, 2010.

19 PRISMA. Prisma Statement. PRISMA. 2015. Available: http://www.prisma–statement.org. Accessed: 25 January 2016.

20 World Bank. Official exchange rate (LCU per US$, period average). 2016. Available: http://data.worldbank.org/indicator/PA.NUS.FCRF. Accessed: 25 January 2016.

21 Agency for Healthcare Quality and Research. Using appropriate price indices for expenditure comparisons. 2015. Available: http://meps.ahrq.gov/about_meps/Price_Index.shtml. Accessed: 25 January 2016.

22 UNAIDS. Feature stories. Zimbabwe: AIDS levy generates new resources for treatment. 2012. Available: http://www.unaids.org/en/resources/presscentre/featurestories/2012/february/20120. Accessed: 25 January 2016.

23 IBRD. IBRD Botswana Country Brief: Helping Botswana confront a challenging future. 2012. Available: http://siteresources.worldbank.org/NEWS/Resources/BotswanaChallenges4–7–10.pdf. Accessed: 25 January 2016.

REFERENCES

24 AIDSPAN. Two new Debt2Health agreements signed. 2010. Available: http://www.aidspan.org/gfo_article/two–new–debt2health–agreements–signed. Accessed: 25 January 2016.

25 UNAIDS. AIDSinfo. Data. 2015. Available: http://www.aidsinfoonline.org/devinfo/libraries/aspx/dataview.aspx. Accessed: 25 January 2016.

26 National AIDS Council of Zimbabwe. Funding. 2012. Available: http://www.nac.org.zw/about/funding. Accessed: 25 January 2016.

27 National AIDS Control Council. Kenya AIDS Strategic Framework 2014/2015 – 2018/2019. 2014. Available: http://www.nacc.or.ke/images/KASF_F_web.compressed.pdf. Accessed: 25 January 2016.

28 Government of Uganda. HIV Prevention and Control Act. 2014. Available: http://www.parliament.go.ug/new/images/stories/acts/2014/HIV%20and%20AIDS%20prevention%20and%20control%20act%202014.pdf. Accessed: 25 January 2016.

29 The United Republic of Tanzania. Bill Supplement. No. 11. 2014. Available: http://www.tanzania.go.tz/egov_uploads/documents/11–2014–2_sw.pdf. Accessed: 25 January 2016.

30 World Bank. 2013 Trust Fund Annual Report. Washington DC: World Bank, 2013. Available: http://siteresources.worldbank.org/CFPEXT/Resources/299947–1274110249410/CFP_TFAR_AR13_High.pdf. Accessed: 25 January 2016.

31 African Economic Outlook. African economic outlook–measuring the pulse of Africa. 2014. Available: http://www.africaneconomicoutlook.org/en/. Accessed: 25 January 2016.

32 Harper S. The fungibility of aid earmarked for HIV/AIDS control programs. World Dev. 2012;40:2263-74. doi:10.1016/j.worlddev.2012.06.009

33 McCleary W. The earmarking of government revenue: a review of some World Bank experience. World Bank Res Obs. 1991;6:81-104. doi:10.1093/wbro/6.1.81

34 UNAIDS. Documents. Debt for AIDS Swaps: a UNAIDS policy information brief. 2014. Available: http://data.unaids.org/Publications/IRC–pub06/jc1020–debt4aids_en.pdf. Accessed: 25 January 2014.

35 World Bank. IBRD Botswana Country Brief: Helping Botswana confront a challenging future. 2014. Available: http://siteresources.worldbank.org/NEWS/Resources/BotswanaChallenges4–7–10.pdf. Accessed: 25 January 2016.

36 Cassimon D, Renard R, Verbeke K. Assessing debt–to–health swaps: a case study on the Global Fund Debt2Health Conversion Scheme. Trop Med Int Health. 2008;13:1188-95. Medline:18664242 doi:10.1111/j.1365-3156.2008.02125.x

37 AIDSPAN. Debt2Health Initiative launched by Global Fund. 2007. Available: http://www.aidspan.org/gfo_article/debt2health–initiative–launched–global–fund. Accessed: 25 January 2016.

38 World Bank. HIV/AIDS in Africa: Projects and programs. 2015. Available: http://go.worldbank.org/ZC8O2E5ZR0. Accessed: 25 January 2016.

39 The Global Fund to Fight AIDS. Tuberculosis and Malaria. Key Documents. 2016. Available: http://www.theglobalfund.org/en/. Accessed: 25 January 2016.

40 UNITAID. Innovative financing. 2016. Available: http://www.unitaid.eu/en/how/innovative–financing. Accessed: 25 January 2016.

41 The Global Fund to Fight AIDS. Tuberculosis and Malaria. (RED). 2016. Available: http://www.theglobalfund.org/en/privatengo/red/. Accessed: 25 January 2016.

42 Shah S, Costa K. Social Finance: a primer. Center for American Progress, 2013. Available: https://www.americanprogress.org/issues/economy/report/2013/11/05/78792/social–finance–a–primer/. Accessed: 25 January 2016.

43 Ketkar S, Ratha D. Development finance via Diaspora Bonds track record and potential. Policy Research Working Paper. Washington, DC: World Bank, 2014. Available: http://elibrary.worldbank.org/doi/abs/10.1596/1813–9450–4311. Accessed: 25 January 2016.

44 Saxena S, Cerra V. What caused the 1991 currency crisis in India? IMF Working Papers, 2000;00:1-27.

45 Social Fiance UK. Towards a new Social Economy. Blended value creation through Social Impact Bonds. 2016. Available: http://www.socialfinance.org.uk/wp–content/uploads/2014/07/Towards–A–New–Social–Economy–web.pdf. Accessed: 25 January 2016.

46 Social Finance UK. Social Finance. Peterborough Social Impact Bond: one year on. 2016. Available: http://www.socialfinance.org.uk/peterborough–social–impact–bond–one–year–on/. Accessed: 25 January 2016.

47 Social Finance US. Social Impact Bonds. 2016. Available: http://www.socialfinanceus.org/social–impact–financing/social–impact–bonds/history–sib–market/united–states. Accessed: 25 January 2016.

48 Center for Global Development; Development Impact Bond Working Group. 2016. Available: http://www.cgdev.org/working–group/development–impact–bond–working–group. Accessed: 25 January 2016.

49 Social Finance US. Other international activity. 2016. Available: http://www.socialfinanceus.org/social–impact–financing/social–impact–bonds/history–sib–market/other–international–activity. Accessed: 25 January 2016.

REFERENCES

50 International Working Group of Sovereign Wealth Funds. Generally accepted principles and practices. 2015. Available: http://www.iwg–swf.org/pubs/eng/santiagoprinciples.pdf. Accessed: 25 January 2016.

51 International Forum of Sovereign Wealth Funds. IFSWF members' experiences in the application of the Santiago Principles. 2015. Available: http://www.ifswf.org/pst/stp070711.pdf. Accessed: 25 January 2016.

52 Al–Hassan A, Papaioannou M, Skancke M, Sung C. Sovereign Wealth Funds: aspects of governance structures and investment management. IMF Working Papers 2013;13:1.

53 Triki T, Faye I. Africa's quest for development: can sovereign wealth funds help? African Development Bank. 2011. Available: http://www.afdb.org. Accessed: 25 January 2016.

54 World Bank. World Bank Guarantee Program for the consultation of modernizing the World Bank's operational policy on guarantees. 2016. Available: https://consultations.worldbank.org/Data/hub/files/paris–feedback–wbguarantees–consultation.pdf. Accessed: 25 January 2016.

55 International Finance Corporation. About IFC. 2016. Available: http://www.ifc.org/wps/wcm/connect/corp_ext_content/ifc_external_corporate_site/about+ifc. Accessed: 25 January 2016.

56 Multilateral Investment and Guarantee Agency. Overview. 2016. Available: https://www.miga.org/who–we–are. Accessed: 25 January 2016.

57 Bill & Melinda Gates Foundation. Financials. 2016. Available: http://www.gatesfoundation.org/Who–We–Are/General–Information/Financials. Accessed: 25 January 2016.

58 Stemmer E. Contractual structures and risk allocation and mitigation in the context of public private partnerships in the health sector. Washington, DC. World Bank, 2008. Available at: http://siteresources.worldbank.org/Intguarantees/Resources/RiskMitigationHealthSector_Eric_Stemmer.pdf. Accessed: 25 January 2016.

59 Impact D. Goodbye Malaria: Mozambique Malaria Performance Bond. 2016. Available: https://www.devex.com/impact/partnerships/goodbye–malaria–mozambique–malaria–performance–bond–362. Accessed: 25 January 2016.

60 The Global Fund to Fight AIDS, Tuberculosis, and Malaria. Counterpart financing and willingness to pay. 2013. Available: http://www.theglobalfund.org/en/fundingmodel/counterpartfinancingwtp/. Accessed: 25 January 2016.

61 The United States President's Emergency Plan for AIDS Relief (PEPFAR). Guidance for PEPFAR partnership frameworks and partnership framework implementation plans version 2·0. 2013. Available: http://www.pepfar.gov/reports/guidance/framework/index.htm. Accessed: 25 January 2016.

62 Organisation for Economic Cooperation and Development (OECD). Paris Declaration and Accra Agenda for Action. 2015. Available: http://www.oecd.org/dac/effectiveness/parisdeclarationandaccraagendaforaction.htm. Accessed: 25 January 2016.

63 Vassall A, Remme M, Watts C, Hallett T, Siapka M, Vickerman P, et al. Financing essential HIV Services: a new economic agenda. PLoS Med. 2013;10:e1001567. Medline:24358028 doi:10.1371/journal.pmed.1001567

64 Galárraga O, Wirtz V, Santa–Ana–Tellez Y, Korenromp E. Financing HIV programming: how much should low– and middle–income countries and their donors pay? PLoS ONE. 2013;8:e67565. Medline:23861772 doi:10.1371/journal.pone.0067565

65 Oxman AD, Fretheim A. Can paying for results help to achieve the Millennium Development Goals? Overview of the effectiveness of results–based financing. J Evid Based Med. 2009;2:70-83. Medline:21348993 doi:10.1111/j.1756-5391.2009.01020.x

66 Over M, Hallett T. How to pay 'cash–on–delivery' for HIV infections averted: two measurement approaches and ten payout functions. Social Science Research Network. Available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1646617. Accessed: 25 January 2016.

journal of global health

Electronic supplementary material:
The online version of this article contains supplementary material.

PAPERS

Cost of management of severe pneumonia in young children: systematic analysis

Shanshan Zhang[1,2], Peter M. Sammon[1], Isobel King[1,3], Ana Lucia Andrade[4], Cristiana M. Toscano[4], Sheila N Araujo[4,5], Anushua Sinha[6], Shabir A. Madhi[7,8], Gulam Khandaker[9,10], Jiehui Kevin Yin[9,10], Robert Booy[9,10], Tanvir M Huda[11,12], Qazi S Rahman[11], Shams El Arifeen[11], Angela Gentile[13], Norberto Giglio[13], Mejbah U. Bhuiyan[14], Katharine Sturm–Ramirez[14,15], Bradford D. Gessner[16], Mardiati Nadjib[17], Phyllis J. Carosone–Link[18], Eric AF Simões[18,19], Jason A Child[20], Imran Ahmed[21], Zulfiqar A Bhutta[21], Sajid B Soofi[22], Rumana J Khan[23], Harry Campbell[1], Harish Nair[1,24]

[1] Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK
[2] Department of Preventive Dentistry, Peking University, School and Hospital of Stomatology, Beijing, PR China
[3] NHS Grampian, UK
[4] Department of Community Health, Federal University of Goias, Brazil
[5] State University of Maranhão, Brazil
[6] New Jersey Medical School, Rutgers, The State University of New Jersey, Newark, New Jersey USA
[7] Medical Research Council, Respiratory and Meningeal Pathogens Research Unit, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
[8] Department of Science and Technology/National Research Foundation, Vaccine Preventable Diseases, University of the Witwatersrand, Johannesburg, South Africa
[9] National Centre for Immunisation Research and Surveillance, The Children's Hospital at Westmead, NSW, Australia
[10] Sydney School of Public Health, Faculty of Medicine, The University of Sydney, NSW, Australia
[11] Centre for Child and Adolescent Health, icddr,b, Dhaka, Bangladesh
[12] School of Public Health, Sydney Medical School, University of Sydney, Sydney, Australia
[13] Epidemiology Department, Ricardo Gutierrez Children Hospital, University of Buenos Aires, Argentina
[14] Centre for Communicable Diseases, icddr,b, Dhaka, Bangladesh
[15] Influenza Division, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, GA, USA
[16] Agence de Médicine Préventive, Paris, France
[17] Faculty of Public Health, University of Indonesia, Jakarta, Indonesia
[18] Department of Pediatrics, Section of Infectious Diseases, University of Colorado Denver School of Medicine, Denver, CO, USA
[19] Center for Global Health and Department of Epidemiology, Colorado School of Public Health, Aurora, CO, USA
[20] Pharmacy Department, Children's Hospital Colorado, Aurora, CO, USA
[21] Center of Excellence in Women and Child Health, the Aga Khan University, Karachi, Pakistan
[22] Department of Paediatrics & Child Health, The Aga Khan University, Karachi, Pakistan
[23] James P Grant School of Public Health, BRAC University, Dhaka, Bangladesh
[24] Public Health Foundation of India, New Delhi, India

Correspondence to:
Dr Harish Nair
Centre for Global Health Research
Usher Institute of Population Health Sciences and Health Informatics
University of Edinburgh Medical School
Edinburgh EH8 9AG
United Kingdom
harish.nair@ed.ac.uk

Background Childhood pneumonia is a major cause of childhood illness and the second leading cause of child death globally. Understanding the costs associated with the management of childhood pneumonia is essential for resource allocation and priority setting for child health.

Methods We conducted a systematic review to identify studies reporting data on the cost of management of pneumonia in children younger than 5 years old. We collected unpublished cost data on non–severe, severe and very severe pneumonia through collaboration with an international working group. We extracted data on cost per episode, duration of hospital stay and unit cost of interventions for the management of pneumonia. The mean (95% confidence interval, CI) and median (interquartile range, IQR) treatment costs were estimated and reported where appropriate.

Results We identified 24 published studies eligible for inclusion and supplemented these with data from 10 unpublished studies. The 34 studies included in the cost analysis contained data on more than 95 000 children with pneumonia from both low– and–middle income countries (LMIC) and high–income countries (HIC) covering all 6 WHO regions. The total cost (per episode) for management of severe pneumonia was US$ 4.3 (95% CI 1.5–8.7), US$ 51.7 (95% CI 17.4–91.0) and US$ 242.7 (95% CI 153.6–341.4)–559.4 (95% CI 268.9–886.3) in community, out–patient facilities and different levels of hospital in–patient settings in LMIC. Direct medical cost for severe pneumonia in hospital inpatient settings was estimated to be 26.6%–115.8% of patients' monthly household income in LMIC. The mean direct non–medical cost and indirect cost for severe pneumonia management accounted for 0.5–31% of weekly household income. The mean length of stay (LOS) in hospital for children with severe pneumonia was 5.8 (IQR 5.3–6.4) and 7.7 (IQR 5.5–9.9) days in LMIC and HIC respectively for these children.

Conclusion This is the most comprehensive review to date of cost data from studies on the management of childhood pneumonia and these data should be helpful for health services planning and priority setting by national programmes and international agencies.

Pneumonia is one of the leading causes of morbidity and mortality in children under–five globally, and accounted for about 935 000 (15%) deaths in 2013 and 120 million new episodes of illness in this age group in 2010 [1,2]. Appropriate management of childhood pneumonia can reduce pneumonia–specific mortality by 32–72% [3–5] and thus accelerate the progress toward achievement of the Millennium Development Goal 4 (MDG4). Childhood pneumonia places a large economic burden on families and the health care system, especially in resource–constrained low– and middle–income countries (LMIC). Severe ALRI is a substantial burden on health services worldwide and a major cause of hospital referral and admission in young children [6]. Although several studies in high–income as well as low– and middle–income countries have reported the costs associated with an episode of pneumonia (at the individual patient level), there are no published systematic reviews summarizing the evidence from different health systems and settings globally. Bahia et al reviewed pneumococcal disease costs and productivity loss in Latin America and the Caribbean showed variation in unit costs of pneumococcal pneumonia at outpatient and inpatient levels [7]. We aimed to conduct a systematic review of published data on the costs associated with management of pneumonia episodes in children younger than 5 years and to identify unpublished data sets from pneumonia research groups globally. Cost estimates based on these data should be useful to develop models for estimating cost of management of pneumonia in community as well as hospital–based settings.

METHODS

Review of published studies

We aimed to identify all published studies reporting empirical cost data on the treatment of episodes of pneumonia in children aged below 5 years during a 15–year period (1998–2013). We included studies in children younger than 5 years with pneumonia managed as in–patients or out–patients (using standard treatment per local standard) in secondary and tertiary hospitals, first level facility or in community settings. Data on the cost of a single episode of severe pneumonia from the societal and health care perspectives were collected as the primary study outcome. We developed a review protocol at the beginning of this study and followed the same throughout the process.

We undertook a systematic literature review with three reviewers (PS, IK, SZ), and hand searched reference list of all included articles. We searched four databases (with online search tools) to offer maximum coverage of the relevant literature: Medline, EMBASE, The Centre for Review and Dissemination Library (incorporating the DARE, NHS EED, and NHS HTA databases); and The Cochrane Library (via the Wiley Online Library) for the period 1 January 1998 to October 31 2013. (for search strategy, see Appendix S1 in **Online Supplementary Document**).

Three review authors (SZ, PS, IK) independently selected potentially relevant studies based on their title and abstract. Any disagreements in study selection or data extraction were resolved after discussion with SZ and HC. The eligible studies were retrieved electronically for full–text review. We included studies that investigated all–cause pneumonia in a non–selective population sample, reported empirical cost data for pneumonia treatment (using any intervention including, but not limited to, antibiotics), and included only children younger than 5 years or reported data separately for this age group. We excluded review articles, vaccine cost–effectiveness trials, and studies considering specially selected cohorts with severe co–morbidity (Appendix S2 in **Online Supplementary Document**). We developed and piloted a comprehensive data extraction template. We collected data on cost per episode, cost and unit cost of medication and services, duration of hospital stay and direct medical and non–medical costs. Direct medical cost included costs related to medication, diagnostic tests, medical staff time and hospital stay. Direct non–medical costs included those relating to food, transportation and accommodation charges. Any additional data on indirect costs such as care–givers' time and earning loss were also recorded, where available. Additionally, we extracted data on study characteristics including country, treatment setting, study type and sample size. We recorded the cost study perspective only if explicitly stated in the text of the article to avoid subjective influence. For those papers that did not explicitly state the perspective used, we noted "–" for "unstated".

We classified countries into high income and low–and–middle income categories based on the classification adopted by the World Bank and according to 2012 Gross National Income (GNI) per capita, calculated using the World Bank Atlas method. The groups are low–income per capita US$ 1035 or less; lower middle–income US$ 1036–US$ 4085; upper middle income US$ 4086–US$ 12 615; and high income US $12 616 or more [8].

Quality assessment

We assessed the quality of the included studies using a 13 point scale based on a modified Drummond checklist [9] for economic evaluation focusing on the methodological robustness and detail of reporting (Appendix S3 in **Online Supplementary Document**). Studies were considered high quality if more than 10 points were addressed, medium quality studies covered 7–9 points and low quality studies addressed less than 6 points. Studies with all quality levels were included in the final analysis.

Unpublished data collection

We collected unpublished data from 10 collaborating sites that were part of a Severe ALRI Working Group (SAWG) [6]. The study population included children under 5 years of age with a clinical diagnosis of pneumonia. We defined pneumonia using the World Health Organization's (WHO) Integrated Mangement of Childhood Illness (ICMI) definition by three different severity categories: non–severe, severe and very severe pneumonia based on WHO pocket book for hospital care for children 2005 [10]. We included all interventions for pneumonia management as detailed in the WHO pocket–book (for community/and facility–based management) where data were available.

We designed a costing spreadsheet with detailed descriptions of case definitions and methods and used this for data collection. Actual cost of medications, supplies, personnel and average laboratory costs were collected. Methods used to gather primary cost data in these studies were recorded in the spreadsheet. Resource utilization data from patient records were also documented, where available, including length of stay in hospital, the quantity of drugs and supplies utilized by each patient, and the use of diagnostic tests and procedures. We also attempted to collect data on out–of–pocket spending (by patients) on transport and food where possible. Indirect cost of caregivers' time and daily pay rate were also recorded. Primary data collection was conducted using the provided standardized templates and guidelines at individual study site. (Appendix S4 in **Online Supplementary Document**).

We used a bottom–up approach to calculate cost per episode for each level of the intervention (community, first level health facility and hospital). Costs were calculated and presented separately based on severity and service delivery channels: very severe pneumonia at hospital level (defined as pneumonia with central cyanosis, inability to breastfeed or drink, or vomiting everything, convulsions, lethargy or unconsciousness and severe respiratory distress diagnosed by doctor or physicians using WHO IMCI (2005) case definition or pneumonia cases requires critical care); severe pneumonia at hospital level (defined as pneumonia with chest indrawing using WHO IMCI definition or pneumonia need for hospital admission based on physician's assessment); severe pneumonia at community level (based on assessment by a trained health worker at home/first level facility using WHO IMCI (2005) case definition); and non–severe pneumonia at outpatient level (defined as fast breathing for age in children aged 2 to 59 months). The costing model included direct medical cost, direct non–medical cost and indirect costs. We calculated the cost per episode based on the estimates of the unit cost per contact (eg, unit cost of an antibiotic per day) at each management level multiplied by the resource utilization proportions (eg, 80% of children took amoxicillin for 5 days), plus indirect costs. For the mean total cost of treatment per episode we summed the cost of drugs, diagnostic investigations and hospital stay, as well as transportation and opportunity cost for caregivers' time. The formula is given in **Figure 1**.

We reported all cost data in 2013 US$ equivalent prices. We first converted all costs to US$ and then adjusted for inflation to 2013 values. Conversions were made using the Penn World Tables 8.0 (http://www.ggdc.net/pwt) and an online inflation–calculating tool (http://usinflation.org/cpi–inflation–calculator) on 20th October 2013.

Statistical analysis

We have stratified the cost results by country income category: high–income countries (HIC) and low– and middle–income countries (LMIC). As an important input in the costing analysis for in–patient management, length of stay (LOS) in hospital was extracted for severe hospitalized cases. Cost per episode, cost by component (direct medical, direct non–medical and indirect costs), and percentage of total cost per episode in each component were summarized. Cost per episode was synthesized by severity of diseases in each strata. The mean with 95% CI of the treatment costs and the median with interquartile range (IQR) of LOS were estimated and reported where appropriate. The 95% CIs were calculated based on 5000 bootstrap samples. Mean and median values were compared using appropriate statistical tests.

Direct medical cost in studies reported from household perspective were compared with monthly household income in respective countries to evaluate the burden on

$$\text{The mean cost per episode} = \text{Unit cost of drugs/other treatment per dose} \times \text{Dose needed per day} \times \text{No. of days of drug use/ other treatment per episode} \times \text{Resource utilisation proportions} + \text{Cost for hospital stay per day} \times \text{Length of stay in hospital} + \text{Transport -ation + food} + \text{Care -givers time} \times \text{Daily pay rate}$$

Figure 1. Formula for the mean total cost of treatment per episode.

families. Monthly household incomes were derived from Gallup World Poll using annual median household income divided by 12 months [11]. These income results were based on Gallup data gathered between 2006 and 2012 in 131 populations. In two countries which annual household income data was missing, we used GNI per capita from World Bank database times the mean number of people per household instead. The percentages of direct non–medical costs and indirect cost per episode of weekly household income were also assessed to show the economic impact of pneumonia management for families when direct medical cost was not considered.

We conducted all data analyses using SPSS v.19 (IBM, New York City, NY, USA) noting that included studies showed marked heterogeneity of population, methodology, treatment procedure reporting categories and perspectives.

RESULTS

Search results

We identified 789 studies through database searching, of which 60 articles were eligible for full text review on the basis of title and abstract assessment (**Figure 2**). Subsequently, only 24 papers were identified to be eligible for data extraction and analysis. The key reasons for exclusions included: no data for children below 5 years or no cost data on pneumonia management were reported. For unpublished studies, we contacted 16 sites, 10 of which had data that met our eligibility criteria and contributed to the analysis. The unpublished cost data were for the period January 2001 to August 2012. Six of these sites provided cost data using a template and guidelines designed for this project while the remainder provided unpublished data in their own formats.

Characteristics of published and unpublished data

We identified 24 studies from the literature review and collected additional 10 data sets of unpublished studies totalling 34 studies from 21 countries across the 6 WHO regions (**Table 1**). Over 60% of the studies (21 out of 34) were conducted in the South East Asia and Africa Regions. The included studies reported data from a variety of treatment settings: community, out–patient and in–patient care settings in primary, secondary and tertiary hospitals, and at city, district, provincial and national levels. Twenty–seven of the 34 studies were stand–alone primary cost analysis and/or cost–of–illness studies. The remaining 7 studies were designed to collect cost data alongside clinical trials or epidemiological studies.

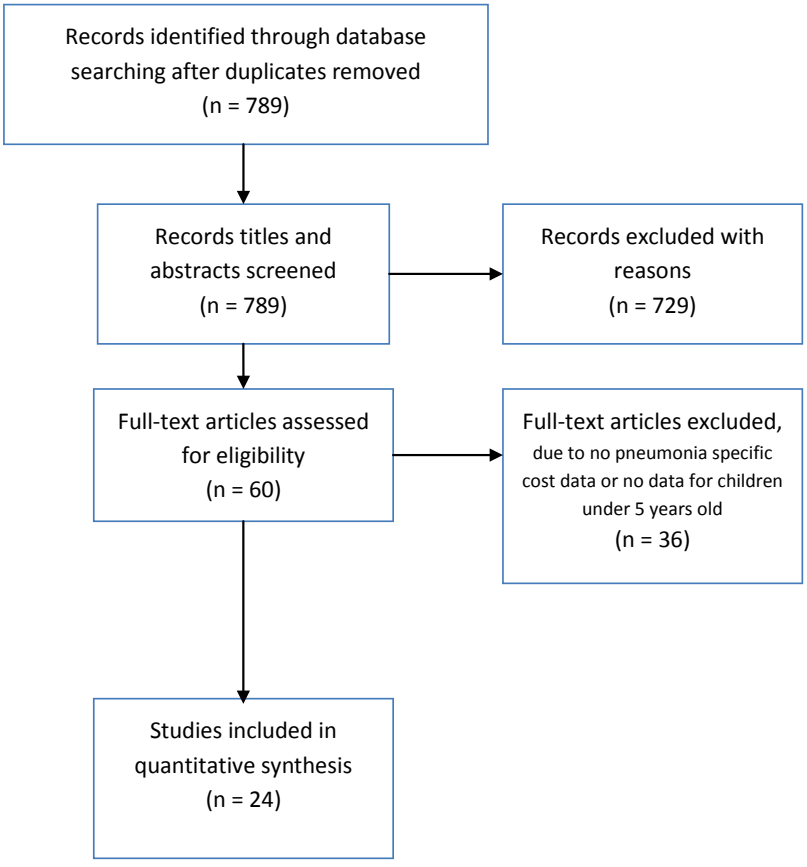


Figure 2. PRISMA flow diagram for severe pneumonia cost systematic review.

Table 1. Characteristics of all studies included*

WHO region	Country, publication year	Study population	Healthcare setting	Severity of pneumonia studied	Study design	Source of case definition	Perspective	Sample size	Mean (SD) / median age of patients (months)	Data source
High–income countries (number of studies = 8)										
EUR	Northern Ireland, 1999 [39][1]	Antrim (urban)	H2	S	QES	PD	N/A–	45	39.60 (16.8)	H
	Spain, 2013 [17]	Barcelona (urban)	H3	S, VS	Cost analysis‡	Culture–proved pneumonia	Healthcare	101	39.60	H
	Germany, 2005 [16]	National	O,H1	S, VS	Cost–of–illness	PD	Societal	402	N/A	N, IQ
AMR	Chile, Uruguay, 2007 [12]	National	O,H1	S, NS	Cost analysis‡	PD, ICD–10	Healthcare	366	N/A	H,IQ
	United States, 2012*	Denver, Colorado (urban)	H3	S, VS, NS	Cost–of–illness	PD by WHO IMCI definition	Societal	940	0–59	H, P
WPR	Australia, 2008 [15]	National	O, H1–3	S	Cost analysis‡	ICD–10	Healthcare	1348	N/A	N
	Australia, 2008 [14]	Melbourne, Victoria (urban)	O,H1	S	Cohort study/ cost–of–illness	Health professional's diagnosis	Societal	528	N/A	N,H,IQ, Pilot
	Australia, 2011*	Sydney (urban)	H3	S, VS	Cost–of–illness	PD by WHO IMCI definition	Societal	N/A	N/A	P, Market price
Low– and middle–income countries (number of studies = 27)										
SEAR	Bangladesh, 2010 [26]	Dhaka (urban)	H3	S	Cost–of–illness	PD	Family	90	5.00	IQ
	Bangladesh, 2005†	Dhaka (urban)	H3	S, VS	Cost–of–illness	PD by WHO IMCI definition	Household	114	70.32	IQ
	Bangladesh, 2010 [24]	Mirpur, Dhaka (urban)	O, H2	S	RCT/CEA	PD by WHO IMCI definition	Societal	360	8.00	–
	Bangladesh, 2010†	Barishal, Bogra, Comilla, Kishoregonj (urban)	H3	S, NS	Cost–of–illness	PD by WHO IMCI definition	Societal	235	N/A	IQ
	Bangladesh, 2012†	Mohakhali, Dhaka (urban)	H3	S, VS, NS	Cost–of–illness	PD by WHO IMCI definition	Societal	340	N/A	H
	India, 2009 [30]	Vellore (rural)	H1, H2	S	Cost–of–illness	PD by WHO IMCI definition	Healthcare/ Household	56	8.8	H, IQ
	India, 2002 [29]	Berhampur, Orissa (urban and rural)	H3	S	Epidemiological study	PD	Societal	52	N/A	H, IQ
	Indonesia, 2001†	Lombok (rural)	H3	S	Cost–of–illness	PD by WHO IMCI definition	Societal	N/A	N/A	H
	Pakistan, 2003 [25]	Peshawar city (urban)	H3	S	RCT/CEA	PD by WHO IMCI definition	–	126	N/A	–
	Pakistan, 2006 [20]	Ghizer district (rural)	O, H1, H2	S, NS	Cost analysis‡	PD	Societal	502	N/A	IQ
	Pakistan, 2008 [19]	Ghizer district (rural)	O, H1, H2	S, VS, NS	Cost analysis‡	PD by WHO IMCI definition	Healthcare	141	N/A	IQ
	Pakistan, 2010†	Matiari (rural)	C	S	Cost–of–illness	PD by WHO IMCI definition	Healthcare	N/A	N/A	Surveillance
	Pakistan, 2012 [23]	Haripur district (rural)	C, H1, H2	S	Cost analysis‡	WHO definition by health worker	Household	423	N/A	H, IQ
	Viet Nam, 2010 [18]	Nha Trang city (urban)	H2	S, VS, NS	Cost–of–illness	PD by WHO IMCI definition	Healthcare	788	12.67	N, H
	Viet Nam, 2001 [28]	Ba Vi district (rural)	C, O, H1	S	Cost analysis‡	WHO definition, self–reported	Household	94	N/A	IQ

Table 1. Continued

WHO region	Country, publication year	Study population	Healthcare setting	Severity of pneumonia studied	Study design	Source of case definition	Perspective	Sample size	Mean (SD) / median age of patients (months)	Data source
AFR	Guinea, 1998 [21]	National	O, H1	S, NS	CEA	PD	–	73650	N/A	H, E
	South Africa, 2011 [33]	Pretoria (urban)	H3	S, VS	Cost analysis‡	WHO definition	–	3014	N/A	H
	South Africa, 2012 [22]	National	H3	S, VS, NS	RCT	PD	Societal/ health care	745	N/A	H, IQ
	South Africa, 2001†	Soweto (urban)	H3	S, VS	Cost–of–illness	PD by WHO IMCI definition	Societal	509	14.00	H,IQ
	Kenya, 2009 [32]	National	H3, H2, H1	S	Cost analysis‡	PD	Societal	205	12.00	H, IQ
	Zambia, 2009 [31]	Kanyama Township (urban)	O,H2	S	Cost analysis‡	PD	Healthcare	9146	N/A	N,H,P,W
AMR	Colombia, 2013 [27]	National	H1,H2,H3	S, VS, NS	Cost–of–illness§	WHO definition, radiographically diagnosed	Healthcare	1545	N/A	I
	Brazil, 2011†	Goiânia (urban)	H3	S, VS	Cost–of–illness§	PD by WHO IMCI definition	Societal	79	0–36	H, N
	Argentina, 2012†	Buenos Aires (urban)	H3	S, VS	Cost–of–illness§	PD by WHO IMCI definition	Societal	N/A	N/A	N
	Brazil,2007 [12]	National	O,H1	S, NS	Cost analysis‡	PD, ICD–10	Healthcare	366	N/A	H,IQ
WPR	Fiji, 2012 [34]	Viti Levu (urban and rural)	O	S	Cost analysis‡	PD by WHO IMCI definition	Societal/ household	390	N/A	N,H, IQ
EMR	Jordan, 2010 [35]	Amman	H1	S	Cohort study	PD	–	728	4.30	N/A

*Severity of pneumonia: NS – non severe, S – severe, VS – very severe. Data source: H – hospital records, N – national data, IQ – interviews and questionnaires, I – insurance database, P – pharmacy database, W – WHO database. Treatment settings: H3 – tertiary hospital in–patient, H2–secondary hospital in–patient, H1 – primary hospital inpatient, O – out–patient care, C – community ambulatory care; PD – physician's diagnosis, CEA – cost effectiveness analysis, RCT – randomized clinical trial, QES– quasi–experimental study, N/A – not available, PD – physician's diagnosis, IMCI – Integrated Management of Childhood Illness, WHO – World Health Organization, SD – standard deviation.

EUR – Europe Region, AMR – the Americas Region, WPR – Western Pacific Region, SEAR – South East Asia, AFR– The Africa Region, EMR– Eastern Mediterranean Region

†Unpublished data.

‡The analysis of the comparative costs of alternative treatments or health care programmes.

§The cost analysis of treatment of a disease.

The included studies reported cost data on a total of 97 062 children treated at facility or community levels, with a median sample size of 378 (IQR 117–741) across all studies. The age of the participants was reported in 12 studies and the median age was 12.3 months (IQR 8.20–33.20). The perspective of costing was explicitly stated in 30 of 34 studies. Of these, the most common perspectives were societal (16 out of 33, 1 study did not specify perspective), health care (11 out of 33) and household (5 out of 33). In most studies, the sources of pneumonia case definition were physician's diagnosis according to WHO IMCI definition (29 out of 34), culture–proved pneumonia was used for case definition in 1 study, self–reported condition according to WHO IMCI definition was adopted in one study, and two studies used health workers' diagnosis. A number of different sources were used for gathering cost data, the most common being through hospital records and costing interviews/questionnaires. Interviews and questionnaires were commonly used in studies with a household perspective to collect data on indirect costs. Other sources included a national database of costs, insurance databases, surveillance data and pharmaceutical databases. The WHO CHOICE database, expert opinion and data from pilot studies were also used to collect data on the unit cost of pneumonia treatment.

The average quality score of 24 published studies was 8.21 out of 13 on scale based on modified Drummond checklist (range 3–12) (Appendix S5 in **Online Supplementary Document**). The majority of the studies failed to consider discounting and did not perform sensitivity analyses. There were 6 studies considered high quality, 14 studies were medium quality, and 4 low quality studies. All studies were included into the final analysis.

Cost of management per episode of pneumonia

Cost results stratified by income category are presented in **Table 2**, **Table 3** and **Figure 3**. In HICs, the mean cost of treatment for an episode of severe pneumonia at the hospital out–patient level was US$ 251.1 in Germany [16]. An Australian study [14] reported similar cost of US$ 254.9 for community management of severe pneumonia. Average costs of facility based case management for young children admitted in primary/sary and tertiary hospitals were US$ 2803.5 (95% CI IQR 2000.6–3683.3), and US$ 7037.2 (95% CI 4028.6–11 311.0) respectively, which was 11–28 fold higher than in those managed as out–patients. The total cost per episode for the management of non–severe pneumonia at out–patient level was also reported for three countries: US$55.8 in Uruguay [12], US$ 272.7 in Chile [12] and US$ 334.6 in the United States [36]. The cost for very severe pneumonia managed in general pediatric wards followed by intensive care unit (ICU) care was reported to be US$9151.3 in a tertiary hospital in Spain [17] and US$ 120 576.3 in the United States, which is nearly 2–17 fold of the cost for severe pneumonia management in hospital settings in HICs. The majority of studies from HIC took only direct medical costs into consideration. Only two studies included direct non–medical costs and indirect costs [16]. The mean proportion of the total cost for direct medical, direct non–medical and indirect costs were 41.5%, 19.5% and 38.5% respectively.

In LMICs, the cost of case management for severe pneumonia was reported across all treatment settings. The community management cost was only reported in studies conducted in South–East Asia region, with a mean cost of US$ 4.3 (95% CI 1.5–8.7) per episode. Out–patient care mean costs were US$ 51.7 (95% CI 17.4–91.0) per case. Costs for in–patient care varied by regions, level of hospitals (primary/sary/tertiary), and levels of care offered at a facility: the mean cost for primary/sary hospital care was 242.7 (95% CI 153.6–341.4) and for tertiary/teaching hospital was 559.4 (95% CI 268.9–886.3). Two groups–severe pneumonia by WHO IMCI definition and hospitalized pneumonia by physician's diagnosis–showed similar costs in all levels of care. The in–patient care costs were 4–11 fold greater than that for out–patient care in the LMICs strata, which in turn was significantly higher than that for community management.

Table 2. Cost per episode for childhood pneumonia management in high–income countries

Severity	WHO region	Country, publication year	Perspective	Sample size	Cost per episode (2013 US$)				Cost component, % of total cost per episode		
					Tertiary/teaching hospital in–patient care	Secondary/ primary hospital in–patient care	Out–patient care	Community care	Direct medical	Direct non–medical	Indirect
Non–severe pneumonia	AMR	US, 2012*	Societal	940			334.6				
		Chile, 2007 [12]	Healthcare	366			272.7				
		Uruguay, 2007 [12]	Healthcare	366			55.8				
Non–severe pneumonia mean cost (95% CI)						**221.0 (55.8–334.6)**					
Severe pneumonia by WHO IMCI Definition	AMR	Chile, 2007 [12]	Healthcare	366		4316.7			100		
		Uruguay, 2007 [12]	Healthcare	366		1421.6			100		
		US, 2012*	Societal	940	15029.2						
	EUR	North Ireland, 1999 [13]	NA	45	5733.8, 2716.8				100		
	WPR	Australia, 2011*	Societal	NA	6,259.1				93.1		6.9
Hospitalised pneumonia	WPR	Australia, 2008 [14]	Societal	528		2813.1		254.9	100		
		Australia, 2008 [15]	Healthcare	1348		2307.8			100		
	EUR	Germany, 2005 [16]	Societal	402		3158.6	251.1		41.5	19.5	38.5
		Spain, 2013 [17]	Healthcare	101	5447.3				100		
Severe pneumonia mean cost (95% CI)					**7037.2 (40 286–11 311.0)**	**2803.5 (2000.6–3683.3)**	**251.1**	**254.9**			
Very severe pneumonia by IMCI	AMR	US, 2012*	Societal	940	120 576.2						
Very severe pneumonia requiring critical care	AMR	Spain, 2013 [17]	Healthcare	101	9151.3						
Very severe pneumonia mean cost (95% CI)					**64 863.8 (9151.3–120 576.3)**						

NA – Information not available, EUR – Europe Region, AMR – The Americas Region, WPR – Western Pacific Region, CI – confidence interval, IMCI – Integrated Management of Childhood Illness

*Unpublished data.

Table 3. Cost per episode for childhood pneumonia management in low– and middle–income countries

Severity	WHO region	Country, publication year	Perspective	Sample size	Cost per episode (2013 US$)				Cost component, % of total cost per episode		
					Tertiary/teaching hospital in–patient care	Secondary/ primary hospital in–patient care	Out–patient care	Community care	Direct medical	Direct non–medical	Indirect
Non–severe pneumonia	SEAR	Viet Nam, 2010 [18]	Healthcare	788			28.6				
		Pakistan, 2008 [19]	Healthcare	141			29.4				
		Pakistan, 2006 [20]	Societal	502			94.1–17.8				
		Bangladesh, 2012*	Societal	340			5.7				
	AFR	Guinea, 1998 [21]	NA	73650			3.2				
		South Africa, 2012 [22]	Societal/ health care	745			263.1				
	AMR	Brazil, 2007 [12]	Healthcare	366			93.0				
Non–severe pneumonia mean cost (95% CI)							**66.9 (21.7–129.7)**				
Severe pneumonia by WHO IMCI Definition	SEAR	Pakistan, 2010*	Healthcare	NA				8.7	100		
		Pakistan, 2012 [23]	Household	423			7.9	1.5	89.1	1.3	9.6
		Bangladesh, 2012*	Societal	340			5.7				
		Bangladesh, 2010 [24]	Societal	360	193.6		124.0		Y	Y	
		Viet Nam, 2010 [18]	Healthcare	788		39.5			Y	Y	Y
		Pakistan, 2008 [19]	Healthcare	141		186.0			64.1	35.9	
		Pakistan, 2003 [25]	NA	126	20.3				100		
		Bangladesh, 2005*	Household	114	80.6†	62.6#			70.9†	29.1†	
		Bangladesh, 2010 [26]	Household	90	124.2				67.6	32.4	
		Indonesia, 2001*	Societal	NA	135.2				75	25	
	AFR	Guinea, 1998 [21]	NA	73650		110.6			69	30	
		South Africa, 2001*	Societal	509	480.9§	110.0					
	AMR	Brazil, 2007 [12]	Healthcare	366		461.0			100		
		Brazil, 2011*	Societal	79	1474.1†,‡	594.5#			94†	1†	5†
		Colombia, 2013[27]	Healthcare	1545		517.6			100		
		Argentina, 2012*	Societal	NA	1648.0				100		
Hospitalised pneumonia	SEAR	Viet Nam, 2001 [28]	Household	94				2.7	56–88	Y	
		Pakistan, 2006 [20]	Societal	502		310.8	127.6		45.3	55	
		India, 2009 [29]	Healthcare/ household	56	145.7	44.7			45.7	5.3	47.4
		India, 2002 [30]	Societal	52	23.9				100		
	AFR	Zambia, 2009 [31]	Healthcare	9146		249.7	55.7		100		
		Kenya, 2009 [32]	Societal	205	236.8	162.1, 89.5			86	14	Y
		South Africa, 2011 [33]	Societal	509	491.4†, 1553.2‡				100		
		South Africa, 2012 [22]	Societal/ health care	745	1223.1				98	2	0.2
	WPR	Fiji, 2012 [34]	Societal/ household	390			25.7, 15.6		61.9	33.2	4.9
	AMR	Colombia, 2013 [27]	Healthcare	1545		304.4	76.2				
	EMR	Jordan, 2012*	NA	728		563.4			100		
Severe pneumonia mean cost (95% CI)					**559.4 (268.9–886.3)**	**242.7 (153.6–341.4)**	**51.7 (17.4–91.0)**	**4.3 (1.5–8.7)**			
Very severe pneumonia by IMCI	SEAR	Bangladesh, 2012*	Societal	340			15.7				
		Viet Nam, 2010 [18]	Healthcare	788		61.2					
		Pakistan, 2008 [19]	Healthcare	141		81.3					
Very severe pneumonia requiring critical care	AFR	South Africa, 2011 [33]	NA	3014	849.0† 14795.4‡						
		South Africa, 2012 [22]	Societal/ health care	745	6696.2						
	AMR	Colombia [27]	Healthcare	1545	3643.4						
Very severe pneumonia mean cost (95% CI)					**6496.0 (2246.2–12007.4)**	**71.3 (61.2–81.3)**	**15.7 (15.7–15.7)**				

NA – information not available, Y – authors considered the cost component, but the proportion was unknown, EUR – Europe Region, AMR – The Americas Region, WPR – Western Pacific Region, SEAR – South East Asia, AFR – The Africa Region, EMR – Eastern Mediterranean Region

*Unpublished data.

†Public health care.

‡Private health care.

§Pediatric ward.

#Supplementary health system.

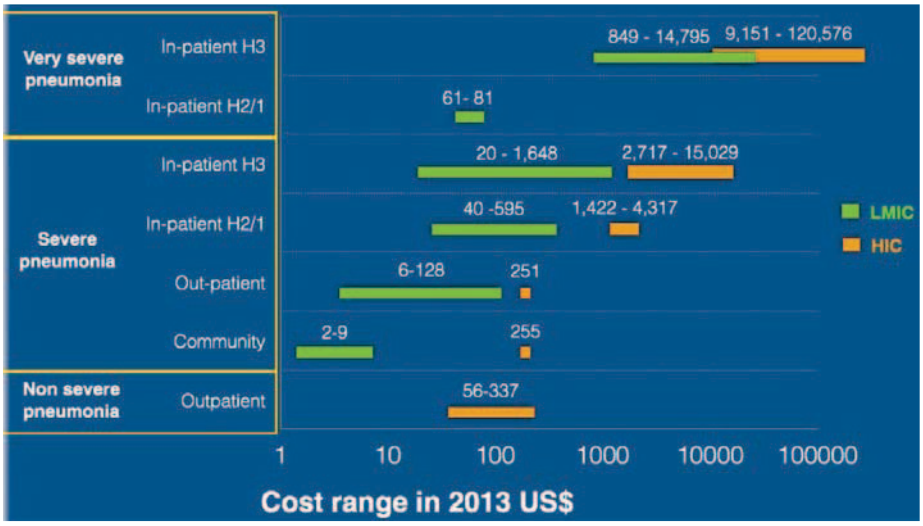


Figure 3. Range of cost per episode for pneumonia management (2013 US$).

The costs for management of non–severe pneumonia managed at outpatient level were US$ 66.9 (95% CI 21.7–129.7), which was slightly higher than for severe cases at outpatient level. This is because the hospital short stay for non–severe pneumonia in South Africa [22] was relatively high at US$ 263.1 per episode. The mean cost for very severe cases was US$ 6496.0 (2246.2–12 007.4), which is nearly 10-fold of severe case management cost.

There were 6 studies reporting cost from household perspective, mainly from LMICs in South East Asia Region. Direct medical cost for severe pneumonia in hospital inpatient settings were 26.6%–115.8% of the monthly household income, thus demonstrating that severe pneumonia management in hospital placed a significant financial burden on families. On the other hand, outpatient and community management of severe pneumonia accounted for only 0.4%–4.1% of family's monthly income indicating decreased burden in these settings. (**Table 4**).

Of the papers reporting cost component of severe pneumonia management, direct medical cost was reported across all studies and accounted for 45%–100% of the total cost. The mean direct non–medical cost and indirect cost were US$ 22.0 (11.8–32.7) and US$ 27.0 (4.0–54.3) respectively, which account for 0.5%–31.0% of weekly household income (**Table 5**).

Length of stay in hospital

The in–patient cost was determined primarily by the length of stay (LOS) and the average cost per bed day. In this review, we extracted length of stay for severe pneumonia for future costing analysis reference (**Tables 6–8**).

Table 4. Direct medical cost for severe pneumonia management in low– and middle–income countries reported from household perspective

Country, publication year	Direct medical cost (2013 $US; % of direct medical cost to monthly household income)				Monthly household income (2013 $US)*
	Tertiary/teaching Hospital in–patient care	Secondary/primary hospital in–patient care	Out–patient care	Community ambulatory care	
Bangladesh, 2010 [26]	124.2 (52.9%)	–	–	–	234.9
Bangladesh, 2005†	80.6‡ (34.3%)	62.6§ (26.6%)		–	234.9
India, 2009 [30]	305.8 (115.8%)	135.1 (64.7%)	–	–	264.0
Pakistan, 2012 [23]	–	–	7.9 (2.3%)	1.5 (0.4%)	338.3
Viet Nam, 2001 [28]	–	–	–	2.7 (0.7%)	398.6
Fiji, 2012 [34]	–	–	25.7 (4.1%)/15.6 (2.5%)	–	632.5‡

*Monthly household income (2913 US$) were derived from Gallup World Poll annual median household income, equals annual median household income divided by 12. These results were based on Gallup data gathered between 2006 and 2012 in 131 population. Source: http://www.gallup.com/poll/166211/worldwide-median-household-income-000.aspx.

†Numbers used GNI per capita based on purchasing power parity (PPP) due to lack of monthly household income data. Source: http://data.worldbank.org/indicator/NY.GNP.PCAP.PP.CD/countries.

‡Public health care.

§Supplementary health system.

Table 5. Direct non–medical cost and indirect cost per episode for severe pneumonia management in low– and middle–income countries

Country, publication year	Direct non–medical cost*			Indirect cost†	% of non–medical cost to monthly household income	Monthly household income (2013 US$)
	Total	Transportation	Food	Total		
Bangladesh, 2010 [24]	32.4				13.8%	234.9
India, 2002 [29]	5.3	5.3		47.5	2.0%	264.0
Pakistan, 2008 [19]	35.9	12.2	23.7		10.6%	338.3
Pakistan, 2006 [20]	55.0				16.3%	338.3
Pakistan, 2012 [23]	3.3	2.3	1.0	9.6	1.0%	338.3
Kenya, 2009 [32]	14.0				9.0%	155.8
Guinea, 1998 [21]	30.0				31.0%	96.7‡
South Africa, 2012 [22]	2.0	1.4		0.2	0.5%	434.8
Fiji, 2012 [34]	33.2	33.2		4.9	5.3%	632.5‡
Brazil, 2011*	9.7	8.41	1.31	73.1	1.6%	626.8
Mean (95% CI)	**22.0 (11.8–32.7)**	**10.5 (3.5–22.3)**	**8.7 (1.0–23.7)**	**27.0 (4.0–54.3)**		
Median (IQR)	**22.0 (4.6 –33.9)**	**6.9 (2.1–11.5)**	**1.3 (–)**	**9.6 (2.5–60.3)**		

CI – confidence interval, IQR – interquartile range

*Direct medical costs include medications and consultation, non–medical cost includes transportation, food and accommodation.

†Indirect cost refers to parental loss of earnings in the period of illness. Monthly household income (2013 US$) was derived from Gallup World Poll annual median household income, equals annual median household income divided by 12. These results were based on Gallup data gathered between 2006 and 2012 in 131 populations. Source: http://www.gallup.com/poll/166211/worldwide-median-household-income-000.aspx.

‡Numbers used GNI per capita based on purchasing power parity (PPP) due to lack of monthly household income data. PPP GNI is gross national income (GNI) converted to international dollars using purchasing power parity rates. An international dollar has the same purchasing power over GNI as a US dollar has in the United States. GNI is the sum of value added by all resident producers plus any product taxes (less subsidies) not included in the valuation of output plus net receipts of primary income (compensation of employees and property income) from abroad. Source: http://data.worldbank.org/indicator/NY.GNP.PCAP.PP.CD/countries.

Table 6. Length of stay of very severe pneumonia and severe pneumonia in hospital in high–income countries

Country and year	Setting	LOS (SD) days	Sample size
Very severe pneumonia:			
Germany, 2005 [16]	ICU	7.4 (6.1)	2039
Spain, 2013 [17]	ICU	18.0	99
USA, 2012*	ICU	18.3 (43.1)	1116
Australia, 2011*	ICU	11.0	–
Median (IQR)		14.5 (10.1–18.1)	
Severe pneumonia:			
Ireland, 1999 [13]	Control group	8.3 (7.7–9.0)	44
	New treatment group	4.0 (3.5–4.6)	45
Germany, 2005 [16]	Hospitalised cases	7.4 (6.1)	2039
Australia, 2008 [14]	Without impact diary	8.8	202
	With impact diary	13.5	523
Australia, 2011*	Non–ICU	6.0	–
Spain, 2013 [17]	Non–ICU	10.5	99
USA, 2012*	Non–ICU	2.7 (2.3)	940
Median (IQR)		**7.9 (5.5–9.2)**	

IQR – interquartile range, ICU – intensive care unit

*Unpublished data.

The mean LOS for severe pneumonia reported in individual studies ranged from 4–13.5 days, with a mean LOS 7.7 (95% CI 5.5–9.9) days and median 7.9 (IQR 5.5–9.2) days in HIC, and mean LOS 5.8 (95% CI IQR 5.3–6.4) days and median 6.4 (IQR 4.1–7.1) days in LMIC. For very severe pneumonia management in intensive care unit (ICU), LOS ranged from 7.4 to 18.3 days. The mean and median LOS were 13.7 (95% CI IQR 9.2–18.2) and 14.5 (IQR 10.1–18.1) days in HIC, and 9.5 (95%CI, 7.4–11.8) and 9.2 (IQR 6.1–12.6) days in LMIC.

Unit cost of case management

Unit cost of treatment and resource uptake should be routinely reported in cost studies. However, only 13 of the 34 included studies reported these data. Since treatment protocols (use of antibiotics, diagnostic tests, procedures and levels and intensity of care) varied between studies, this contributed to variations in costs across studies. For example, the average cost of chest radiograph in LMIC was US$ 8.4 (95% CI 4.3–27.0), which was significantly lower than US$ 185.5 (95% CI 66.3–357.7) in high income countries (**Table 9**). We attempted to abstract unit cost data but were unable to include it in the presented direct medical costs because of paucity of information.

DISCUSSION

This is the first attempt to conduct a systematic review of all published and available unpublished cost data on the management of childhood pneumonia. Costs per episodes in HICs were 5–13–fold higher in all delivery channels than those in LMICs. The review demonstrates that the magnitude of cost per episode increases markedly as the level of treatment delivery rises. Community management for severe pneumonia was less than 10% that of the cost of out–patient management among all levels of management in LMICs. Thus, there are strong economic reasons for considering community case management as a central strategy for pneumonia case management in low income countries; this merits further evaluation which should include consideration of medical outcomes. The mean lengths of stay

Table 7. Length of stay of severe pneumonia in hospital in low– and middle–income countries

Country, year	Description	Length of stay (SD) in days‡	Sample size
Viet Nam, 2010 [18]	Probable pneumonia	7.2 (5.0)	40
	Radiograph confirmed	6.7 (3.8)	426
	Probable severe pneumonia	6.2 (3.3)	59
	Radiograph confirmed severe pneumonia	6.4 (2.7)	193
Bangladesh, 2010 [24]	Hospital care	6.0 (5.0–7.0)	180
Bangladesh, 2005*	Public health care	7.1	73
	Private health care	6.4	41
Bangladesh, 2010*	Hospital stay	7.0 (3.0)†	93
Pakistan, 2003 [25]	Antibiotic use duration	Approx. 8	124
Kenya, 2009 [32]	National hospitals	8.2	49
	District hospitals	6.7	30
	District hospitals	4.8	29
	District hospitals	4.2	17
	Provincial hospitals	6.6	31
	Mission Hospitals	7.8	30
	Mission Hospitals	3.4	19
Zambia, 2009 [31]	Tertiary health center	4.0	221
Pakistan, 2008 [19]	Time spent at health facility for severe pneumonia	3.3	65
Pakistan, 2006 [20]	Secondary hospital	3.0	502
Jordan, 2010 [35]	In–patient days	4.0–5.0	728
India, 2009 [30]	Secondary hospital	3.5 (2.9–4.1)	31
	Tertiary hospital	3.7 (3.0–4.4)	25
India, 2002 [29]	Tertiary hospital	6.5 (2.5)	52
Brazil, 2011*	Public health system	3.9 (2.2)	59
	Supplementary health system	5.3 (4.7)	20
Colombia, 2013 [27]	Primary	2.0 (1.0–2.0)	247
	Secondary hospital	4.0 (1.0–5.0)	1208
	Tertiary hospital	6.0 (3.0–9.0)	47
South Africa, 2011 [33]	Public sector ward	8.7	86
	Fee for service sector	5.6	7786
South Africa, 2012 [22]	Paediatric ward	8.1 (7.4–8.8)	513
Indonesia, 2001*	Non–ICU	6.7	–
Argentina, 2012*	Severe pneumonia	7.5(8.5)	42
	Unilateral focal pneumonia without complications	7.4 (6.0)	1994
	Multifocal pneumonia without complications	8.0 (6.5)	323
Median (IQR)		**6.4 (4.1–7.1)**	

ICU – intensive care unit

*Unpublished data.

†Combined HIV+ and HIV–, HIV+ had longer stay in ward (9.3 vs 7.0 days).

‡Length of stay (LOS) reported as mean, mean (standard deviation) or median (interquartile range). When stratified LOS available, then stratified LOS was reported, not average length of stay of all pneumonia.

in hospital for severe pneumonia were 1.8–4.6 days less in LMIC compared to HIC, and at a mean of 5.8 and median of 6.4 days, were close to the WHO recommendation of 5 days in–patient treatment [10].

We demonstrated that the cost (per episode) for the management of severe pneumonia varied greatly by unit cost of intervention, disease severity and treatment procedures in different settings. The review also demonstrated that major factors governing the total cost per episode were length of stay in the hospital, countries income level and the presence or absence of community case management for pneumonia. Many other studies have also found GDP per capita to be the main driver of costs [37]. These findings demonstrate that choosing the appropriate value for these inputs will have a significant influence on the total cost. Existing studies calculated pneumonia management costs in many countries assuming the same treatment procedure and unit cost of medicine. However, the cost data we collected demonstrate that this method may have limitations; the uncertainty in the traditional estimates can be measured using the cost data reported in this review.

Our results showed that direct medical costs for childhood pneumonia management, especially inpatients, represent a significant proportion of the average monthly household income for families in LMICs. This is often compounded by further direct non–medical cost and indirect cost ie, loss

Table 8. Length of stay of very severe and non–severe pneumonia in hospital in low– and middle–income countries

Country, year	Description	Length of stay (SD) days	Sample size
Very severe pneumonia			
Viet Nam, 2010 [18]	Very severe pneumonia	6.4(2.7)	26
	Confirmed very severe pneumonia	5.8 (3.0)	44
Colombia, 2013 [27]	ICU	13.0 (6.0–14.0)	43
South Africa, 2011 [33]	ICU	9.4	46
	ICU	10.5	93
South Africa, 2012 [22]	ICU	14.4(10.3–18.5)	7
Pakistan, 2008 [19]	Time spent at health facility for very severe pneumonia	3.9	35
Argentina, 2012†	Very severe pneumonia	8.9	–
	Unilateral focal pneumonia without complications	17.2	–
	Multifocal pneumonia without complications	11.5	–
Brazil, 2011*	Public health system	6.9	–
	Supplementary health system	6	–
Median (IQR)		9.2 (6.1–12.6)	
Non severe pneumonia			
Pakistan, 2008 [19]	Time spent at health facility for pneumonia	0.3	41
South Africa, 2012 [22]	Short stay	1.4 (1.3–1.6)	338
Median (IQR)		**0.9 (0.3–1.4)**	

ICU – intensive care unit, IQR – interquartile range

*Unpublished data.

†Note added in proof: The data from this study are unpublished but the data on the length of stay are published in Giglio ND, Cane AD, Micone P, Gentile A. Cost-effectiveness of the CRM-based 7-valent pneumococcal conjugated vaccine (PCV7) in Argentina. Vaccine. 2010;28:2302-10. Medline:20064478

Table 9. Chest Radiography cost per episode

Country, year		Cost per episode (US$, 2013)
High income countries	Australia, 2011*	129.8
	Chile, 2007 [12]	135.1
	Uruguay, 2007 [12]	43.4
	United States, 2012*	433.7
Mean(SD)		**185.5 (66.3–357.7)**
Median(IQR)		**132.5 (108.2–209.8)**
Low– and middle–income countries	Argentina, 2012*	26.7
	Brazil, 2011*	10.7
		6.0
	Brazil, 2007 [12]	13.63
	Bangladesh, 2010*	2.3
	India, 2009 [30]	5.4
	Pakistan, 2008 [19]	3.2
	Indonesia, 2001*	4.6
	Kenya, 2009 [32]	2.3
	South Africa, 2001*	29.7
	South Africa, 2011 [33]	59.7
		137.2
	South Africa, 2012 [22]	27.7
Mean (SD)		**25.3 (9.8–47.3)**
Median (IQR)		**8.4 (4.3–27.0)**

*Unpublished data.

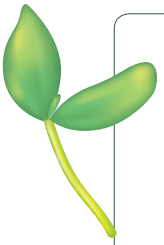
of earnings when caring for the sick child. In countries where these families were uninsured, health payments for pneumonia management were a heavy burden on household and can have a significant impact on the family, particularly when the payments for care were out–of–pocket in most LMIC countries. Alamgir et al investigated the impact that this strain had on families and how they source the funds: many borrow or take high–interest loans [26]. Furthermore, Ayieko et al found that 10% of the patients in district hospitals and up to 25% of children in tertiary hospitals wait in hospital beds after medical discharge while families source the fees. The latter translates to an additional cost of US$ 17.46 to the public provider and US$ 5.32 to the family [32], resulting in a drain on both the resources of the family and the health care provider, as well as denying a bed to another sick child. It is therefore important that national strategies for pneumonia management in LMIC are not only cost–effective for the national program but also give attention to the burden of costs on families so that these are maintained at a level that is affordable.

The data in this review comprise "actual" cost data measured in cost studies conducted in many LMIC and HIC. We believe that these represent a fair first approximation of true costs in these countries. It is noteworthy that the resulting cost estimates are higher than those currently contained in the WHO–CHOICE estimates [38]. Three factors could have contributed to this variation. First, we identified longer facility and hospital stays compared to standard treatment protocols recommended by the WHO [39]. Moreover, most existing cost studies were conducted at tertiary level hospitals where out–patient and in–patient treatments carry a much higher cost compared to the community or first level facility. Third, the wide variety of antibiotics (including variations in dosage, route of administration and duration) across the sites, as well as the heterogeneity in the costing methodology and the cost components in existing studies may have led to higher estimates.

This review has several limitations. First, the primary goal of the systematic review was to obtain data on cost of management (per episode) of severe pneumonia. However, the lack of any standard management protocols (which varied widely across the included studies) and the general lack of service uptake data, may have contributed to the substantial uncertainty around the estimates. Second, we did not include costs of diagnostic investigations in the cost modeling in some study sites, because country–specific unit prices and utilization data were not available. Therefore, the true economic burden resulting from the management of childhood pneumonia could be considerably higher. Third, costs were highly dependent on level of care offered at facility and LOS could be skewed to longer period if high level of care (such as intensive care unit (ICU) care) was offered to severe and very severe cases. In this review, severe cases were all managed at non–ICU hospital settings, and very severe cases were managed at both non–ICU and ICU care in hospital settings. We were able to report LOS separately for non–ICU and ICU care but this stratification was not possible for total cost per episode. A further limitation was that the definition of ICU and ICU care may vary by country. Fourthly, we limited the search to English articles only, which may exclude some cost reported in other languages, however only 13 studies out of 789 articles in other languages were found. We tried to compliment this with unpublished data from non–English speaking countries. Lastly, there was a wide range in per capita income and health care system and payment schemes within LMIC category and the existing cost data may only reflect the situation when and where the data were collected and may not be representative of the whole country or the current situation. Furthermore, there were not sufficient studies to address all possible cost scenarios to facilitate international and public/private comparisons.

Our results demonstrate that further research on the economic burden due to the management of childhood pneumonia is needed, with clear reporting of data on unit cost of intervention, dosage of various drugs and information on health care utilization, such as length of stay in hospital. We recommend that standard reporting of unit cost of intervention with direct medical and non–medical costs and indirect costs, standard treatment protocols and health resource utilization in conjunction with the total cost per episode in any cost–of–illness studies would facilitate economic estimates of national scale–up and international comparisons. Further studies on the cost–effectiveness of standardized IMCI protocol against other treatment protocols could be expected to find a cost–saving management strategy for high burden countries.

Identifying the most cost–effective interventions for pneumonia management is essential for achieving the goal of further reducing child mortality. Our study demonstrated that early treatment in the community costs less (per event) than late treatment in the hospital. This finding suggests that the public health community should explore ways for community outreach for early diagnosis and treatment before severe pneumonia sets in. The results from this systematic review provide important missing information on the cost of pneumonia treatment in children across many countries. These data and the cost estimates should provide important information useful to program managers and policy makers at national and regional levels, international agencies, and donor organisations to aid resource allocation, program planning and priority setting. The estimates presented in this review could enable a more detailed economic evaluation of the revised WHO pneumonia management guidelines [39], and help identify the most cost–effective preventive and treatment interventions for reducing the burden of childhood pneumonia.



Acknowledgments: The authors thank the Maternal Neonatal Child and Adolescent Health (MNCAH) division of the WHO for funding the study and China Scholarship Council for the fellowship to SZ.

Funding: Shanshan Zhang is supported by WHO Maternal, Newborn, Child and Adolescent Health division and China Scholarship Council. Harish Nair and Harry Campbell received funding support from WHO Maternal, Newborn, Child and Adolescent Health division and Bill and Melinda Gates Foundation (OPP 1096225).

Authorship declaration: HN and HC conceived the study. SZ and HN designed the study. PMS, IK, SZ, CMT, SNA, AS, JKY, TMH, NG, MUB, BDG, PJCL, JAC, IA and RK performed the experiments. SZ analysed the data. SZ, HN and HC wrote the paper. AA, CMT, AS, SAM, GK, RB, SEA, QSR, AG, KSR, BDG, MN, PJCL, EAFS, ZAB and SBS critically reviewed the manuscript:. All authors have read, and confirm that they meet, ICMJE criteria for authorship.

Competing interests: HC is an editor–in–chief of the *Journal of Global Health*. To ensure that any possible conflict of interest relevant to the journal has been addressed, this article was reviewed according to best practice guidelines of international editorial organizations. All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). The authors have declared that no competing interests.

REFERENCES

1 Rudan I, O'Brien KL, Nair H, Liu L, Theodoratou E, Qazi S, et al. Epidemiology and etiology of childhood pneumonia in 2010: estimates of incidence, severe morbidity, mortality, underlying risk factors and causative pathogens for 192 countries. J Glob Health. 2013;3:010401. Medline:23826505

2 Liu L, Oza S, Hogan D, Perin J, Rudan I, Lawn JE, et al. Global, regional, and national causes of child mortality in 2000—13, with projections to inform post–2015 priorities: an updated systematic analysis. Lancet. 2015;385:430-40. Medline:25280870

3 Sazawal S, Black RE, Pneumonia Case Management Trials G. Effect of pneumonia case management on mortality in neonates, infants, and preschool children: a meta–analysis of community–based trials. Lancet Infect Dis. 2003;3:547-56. Medline:12954560 doi:10.1016/S1473-3099(03)00737-0

4 Bhutta ZA, Das JK, Walker N, Rizvi A, Campbell H, Rudan I, et al. Interventions to address deaths from childhood pneumonia and diarrhoea equitably: what works and at what cost? Lancet. 2013;381:1417-29. Medline:23582723 doi:10.1016/S0140-6736(13)60648-0

5 Rakha MA, Abdelmoneim A-NM, Farhoud S, Pičche S, Cousens S, Daelmans B, et al. Does implementation of the IMCI strategy have an impact on child mortality? A retrospective analysis of routine data from Egypt. BMJ Open. 2013;3:e001852. Medline:23355663 doi:10.1136/bmjopen-2012-001852

6 Nair H, Simões EA, Rudan I, Gessner BD, Azziz Baumgartner E, Zhang JSF, et al. Global and regional burden of hospital admissions for severe acute lower respiratory infections in young children in 2010: a systematic analysis. Lancet. 2013;381:1380-90. Medline:23369797 doi:10.1016/S0140-6736(12)61901-1

7 Bahia L, Toscano CM, Takemoto MLS, Araujo DV. Systematic review of pneumococcal disease costs and productivity loss studies in Latin America and the Caribbean. Vaccine. 2013;31S:C33-44. Medline:23777689 doi:10.1016/j.vaccine.2013.05.030

8 WorldBank. Country and lending groups 2014. Available: http://data.worldbank.org/about/country-and-lending-groups#high_income. Accessed: 22 October 2014.

9 Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G. Critical assessment of economic evaluation. 3rd edition. Oxford: Oxford University Press; 2005.

10 WHO. Hospital care for children: World Health Organization; 2005.

11 Gallup. Worldwide median household income 2013. Available: http://www.gallup.com/poll/166211/worldwide-median-household-income-000.aspx. Accessed: 13 February 2015.

12 Constenla D. Evaluating the costs of pneumococcal disease in selected Latin American countries. Rev Panam Salud Publica. 2007;22:268-78. Medline:18078589 doi:10.1590/S1020-49892007000900007

13 Al–Eidan FA, McElnay JC, Scott M, Kearney M, Troughton K, Jenkins J. Sequential antimicrobial therapy: treatment of severe lower respiratory tract infections in children. J Antimicrob Chemother. 1999;44:709-15. Medline:10552992 doi:10.1093/jac/44.5.709

14 Lambert SB, Allen KM, Carter RC, Nolan TM. The cost of community–managed viral respiratory illnesses in a cohort of healthy preschool–aged children. Respir Res. 2008;9:11. Medline:18215329 doi:10.1186/1465-9921-9-11

15 Newall AT, Scuffham PA. Influenza–related disease: the cost to the Australian healthcare system. Vaccine. 2008;26:6818-23. Medline:18940222 doi:10.1016/j.vaccine.2008.09.086

16 Ehlken B, Ihorst G, Lippert B, Rohwedder A, Petersen G, Schumacher M, et al. Economic impact of community–acquired and nosocomial lower respiratory tract infections in young children in Germany. Eur J Pediatr. 2005;164:607-15. Medline:15965766 doi:10.1007/s00431-005-1705-0

17 Brotons P, Gelabert G, Launes C, Sicuri E, Pallares R, Munoz–Almagro C. Cost of hospitalizing children with invasive pneumococcal pneumonia. Vaccine. 2013;31:1117-22. Medline:23262168 doi:10.1016/j.vaccine.2012.12.025

18 Anh DD, Riewpaiboon A, Le Huu Tho S, Kim BN, Kilgore P. Treatment costs of pneumonia, meningitis, sepsis, and other diseases among hospitalized children in Viet Nam. J Health Popul Nutr. 2010;28:436-42. Medline:20941894

19 Hussain H, Waters H, Khan AJ, Omer SB, Halsey NA. Economic analysis of childhood pneumonia in Northern Pakistan. Health Policy Plan. 2008;23:438-42. Medline:18755733 doi:10.1093/heapol/czn033

20 Hussain H, Waters H, Omer SB, Khan A, Baig IY, Mistry R, et al. The cost of treatment for child pneumonias and meningitis in the Northern Areas of Pakistan. Int J Health Plann Manage. 2006;21:229-38. Medline:17044548 doi:10.1002/hpm.847

21 Jha P, Bangoura O, Ranson K. The cost–effectiveness of forty health interventions in Guinea. Health Policy Plan. 1998;13:249-62. Medline:10187595 doi:10.1093/heapol/13.3.249

22 Sinha A, Kim S, Ginsberg G, Franklin H, Kohberger R, Strutton D, et al. Economic burden of acute lower respiratory tract infection in South African children. Paediatr Int Child Health. 2012;32:65-73. Medline:22595212 doi:10.1179/2046905512Y.0000000010

23 Sadruddin S, Shehzad S, Bari A, Khan A, Khan A, Qazi S. Household costs for treatment of severe pneumonia in Pakistan. Am J Trop Med Hyg. 2012;87:137-43. Medline:23136289 doi:10.4269/ajtmh.2012.12-0242

24 Ashraf H, Mahmud R, Alam NH, Jahan SA, Kamal SM, Haque F, et al. Randomized controlled trial of day care versus hospital care of severe pneumonia in Bangladesh. Pediatrics. 2010;126:e807-15. Medline:20855397 doi:10.1542/peds.2009-3631

25 Aurangzeb B, Hameed A. Comparative efficacy of amoxicillin, cefuroxime and clarithromycin in the treatment of community–acquired pneumonia in children. J Coll Physicians Surg Pak. 2003;13:704-7. Medline:15569557

REFERENCES

26 Alamgir NI, Naheed A, Luby S. Coping strategies for financial burdens in families with childhood pneumonia in Bangladesh. BMC Public Health. 2010;10:622. Medline:20955627 doi:10.1186/1471-2458-10-622

27 Alvis–Guzman N, Orozco–Africano J, Paternina–Caicedo A, Coronell–Rodríguez W, Alvis–Estrada L, Jervis–Jálabe D, et al. Treatment costs of diarrheal disease and all–cause pneumonia among children under–5 years of age in Colombia. Vaccine. 2013;31:C58-62. Medline:23777692 doi:10.1016/j.vaccine.2013.05.035

28 Toan NV, Khe ND, Hojer B. Expenditure and payment sources for treating acute respiratory infections in rural Vietnam. Southeast Asian J Trop Med Public Health. 2001;32:682-8. Medline:12041539

29 Sahu S, Satapathy D, Sahu T, Tripathy R, Das B, Pradhan S. A study of acute respiratory tract infection cases admitted to a tertiary level health centre. Health and Population. 2002;25:186-93.

30 Madsen HO, Hanehřj M, Das AR, Moses PD, Rose W, Puliyel M, et al. Costing of severe pneumonia in hospitalized infants and children aged 2–36 months, at a secondary and tertiary level hospital of a not-for-profit organization. Trop Med Int Health. 2009;14:1315-22. Medline:19719464 doi:10.1111/j.1365-3156.2009.02374.x

31 Chola L, Robberstad B. Estimating average inpatient and outpatient costs and childhood pneumonia and diarrhoea treatment costs in an urban health centre in Zambia. Cost Eff Resour Alloc. 2009;7:16. Medline:19845966 doi:10.1186/1478-7547-7-16

32 Ayieko P, Akumu AO, Griffiths UK, English M. The economic burden of inpatient paediatric care in Kenya: household and provider costs for treatment of pneumonia, malaria and meningitis. Cost Eff Resour Alloc. 2009;7:3. Medline:19161598 doi:10.1186/1478-7547-7-3

33 Kitchin OP, Wessels F, Masekela R, Becker P, Green RJ. Costs of admission for paediatric pneumonia in a setting of human immunodeficiency virus infection. Int J Tuberc Lung Dis. 2011;15:1702-7. Medline:22118183 doi:10.5588/ijtld.11.0167

34 Temple B, Griffiths UK, Mulholland EK, Ratu FT, Tikoduadua L, Russell FM. The cost of outpatient pneumonia in children<5 years of age in Fiji. Trop Med Int Health. 2012;17:197-203. Medline:22008519 doi:10.1111/j.1365-3156.2011.02897.x

35 Khuri–Bulos N, Williams JV, Shehabi AA, Faouri S, Jundi EA, Abushariah O, et al. Burden of respiratory syncytial virus in hospitalized infants and young children in Amman, Jordan. Scand J Infect Dis. 2010;42:368-74. Medline:20100116 doi:10.3109/00365540903496544

36 Carosone-Link P, Simoes E, Child J. Cost of management of severe pneumonia in young children in Colorado. 2014. Unpublished data.

37 Adam T, Evans D, Murray C. Econometric estimation of country–specific hospital costs. Cost Eff Resour Alloc. 2003;1:3. Medline:12773218 doi:10.1186/1478-7547-1-3

38 Cost effectiveness and strategic planning (WHO–CHOICE).–Quantities and unit prices (cost inputs). Available: http://www.who.int/choice/cost-effectiveness/inputs/en/. Accessed: 22 October 2014.

39 WHO. Pocket book of hospital care for children. WHO 2013. Available: http://apps.who.int/iris/bitstream/10665/81170/1/9789241548373_eng.pdf. Accessed: 22 October 2014.

journal of global health

Moving global health forward in academic institutions

Didier Wernli[1], Marcel Tanner[2,3], Ilona Kickbusch[4], Gérard Escher[5], Fred Paccaud[6], Antoine Flahault[7]

[1] Global Studies Institute, University of Geneva, Switzerland
[2] Swiss Tropical & Public Health Institute, Basel, Switzerland
[3] University of Basel, Switzerland
[4] Global Health Programme, Graduate Institute of International and Development Studies, Geneva, Switzerland
[5] Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland
[6] Institute of Social and Preventive Medicine, Lausanne, Switzerland
[7] Institute of Global Health, Faculty of Medicine, University of Geneva, Switzerland

Global health has attracted growing attention from academic institutions. Its emergence corresponds to the increasing interdependence that characterizes our time and provides a new worldview to address health challenges globally. There is still a large potential to better delineate the limits of the field, drawing on a wide perspective across sciences and geographical areas. As an implementation and integration science, academic global health aims primarily to respond to societal needs through research, education, and practice. From five academic institutions closely engaged with international Geneva, we propose here a definition of global health based on six core principles: 1) cross–border/multilevel approach, 2) inter–/trans–disciplinarity, 3) systems thinking, 4) innovation, 5) sustainability, and 6) human rights/equity. This definition aims to reduce the century–old divide between medicine and public health while extending our perspective to other highly relevant fields. Overall, this article provides an intellectual framework to improve health for all in our contemporary world with implications for academic institutions and science policy.

Health and well–being are major challenges for the 21st century. While these key areas of societal development have gained prominence worldwide by receiving more political attention and funding than ever, the expression 'global health' has emerged to describe the profound shift in the nature of health within the context of globalization. Becoming ubiquitous, global health has generated increasing interest from academic institutions, which, as places of knowledge innovation, validation, transmission, and application, have a critical role to play in global health education, research, and practice [1,2]. In this article we use the terms 'academic global health' (AGH) to focus on the key role of academic institutions including university hospitals in the global health system. As an integration and implementation science [3], the primary goal of AGH is to foster transformative knowledge, which implies both new models of thinking and new types of research. At the operational level, this translates into a process of mutual learning for change and health improvement, through sharing and comparing across systems and cultures, using both qualitative and quantitative methods, validating new evidence internally and externally, and making interdisciplinary and international collaborations a prerequisite. From the viewpoint of five academic institutions closely engaged with Geneva, a leading city in global health and global governance, the present article attempts to reflect on the core principles, definition, and significance of AGH.

THE NEW CONTEXT FOR HEALTH

AGH integrates the three traditional areas of health care, international health, and public health and reflects global changes in five key dimen-

Correspondence to:
Didier Wernli
Global Studies Institute
University of Geneva
8G, rue des Vieux-Grenadiers
1205 Genève
Switzerland
didier.wernli@unige.ch

sions. First, from a predominantly local or national issue, health has become more transnational as the scope and velocity of the transmission of diseases and their determinants have increased, thereby making broad international collaborations and partnerships indispensable. Second, as epitomized by HIV–AIDS, the distinction between curative individual–based medicine and preventive population–based public health has blurred, requiring a rethink of the provision of public health services and health care delivery as a continuum rather than separate entities. Third, the governance of health and social systems has come to include a broad range of actors beyond governments such as charity, civil society, and the private sector, making a continuous assessment of roles and responsibilities of all actors a necessity. Fourth, the biomedical paradigm rooted dominantly in reductionism and biological determinism has failed to provide sustainable solutions for health and well–being, implying the need to develop broader transdisciplinary approaches. Fifth, the interdependence of health with other sectors, together with foreign policy agendas such as trade, security, human rights, environment, and development, has been increasingly recognized, requiring systemic approaches through which diseases and health problems are positioned within broader social, ecological and political systems. Clearly, contemporary global changes have decreased the capacity of the 20th century dominant conceptualizations of international health, and to a lesser extent health care and public health, to address current health challenges effectively. As AGH emanates from this new context for health, we propose six core principles to guide global health research, education, and practice.

SIX CORE PRINCIPLES FOR AGH

AGH addresses cross–border and multi–level health issues

As health issues increasingly cross national boundaries, we need to understand how phenomena occurring at different spatial and temporal scales interface. For example, multilevel geo–ecological frameworks explore how determinants shape health from micro/local to macro/global levels drawing on the progress of scientific knowledge in many fields. Globalization is not a simple process and not everything is global: we constantly face complex "fragmegrative" dynamics where globalizing forces are counteracted by localizing ones [4]. Thus, continuous communication between local communities and academics or professionals working at different levels and diverse geographic areas is crucial to optimize the tailoring of local interventions while avoiding fragmentation of the strategies regionally and globally. A comparable relevant example is the interrelated area of climate change. As addressing the root global causes of climate change is impossible at the local/national level alone, an understanding of scientific evidence facilitated by the UN's Intergovernmental Panel on Climate Change can direct these constituencies toward well–tailored adaptation and mitigation policies.

AGH mobilizes all relevant academic disciplines

While traditional academic disciplines identify, delimit, and analyze phenomena, they tend to produce hyper–specialization, which in turn can result in fragmented understanding and actions in silos. Fragmentation is amplified by the enormous amount of knowledge produced within and outside the academy. As philosopher of science Karl Popper put it: "we are not students of some subject matter, but students of problems" [5], meaning that addressing complex societal problems such as the Ebola crisis in West Africa in 2014 transcends the boundaries of academic disciplines. Consequently, AGH should not be conceptualized as a new discipline but rather as a "transdiscipline" that seeks to integrate knowledge from different sources. Although currently AGH is, still mostly multidisciplinary, corresponding to a juxtaposition of disciplinary perspectives, it should become more interdisciplinary integrating insights from all relevant academic disciplines. Even better it should aim to become transdisciplinary, integrating insights from all relevant disciplines and actors outside academia to address problems too complex for a single discipline or sector [6].

AGH studies complex systems in the real world

Systems science, which encompasses a broad set of theories and methods developed in life sciences, social sciences and engineering during the 20th century, focuses on the principles that govern living and social systems. From cells to global governance, global health refers to complex systems. These systems are constituted of multiple components interacting through reinforcing or inhibiting feedback loops, they operate in constantly evolving contexts, and they typically exhibit properties that result not from specific components of the system but from their interactions, such as nonlinear behavior, self–organization, and emergence [7]. By analyzing the roles, positions, responsibilities and interdependencies of the different building blocks of global health systems, systems science modifies our mental boundaries, generates new questions and hypotheses, and improves our models which can in turn reduce policy failure. While a system perspective is critical to address major health problems in the real world, AGH does not rule out reductionist and selective approaches, as basic reductionist research is a major driver of scientific progress and as selective approaches has been highly successful in some cases (eg, eradication of smallpox). Overall, AGH aims to provide an integrated intellectual framework for debating, experimenting, and implementing options.

AGH seeks to provide affordable, effective, and integrated innovation

The exponential growth of scientific and technological knowledge is key to improving health and well–being both in high–income countries (HIC) and low– and middle–income countries (LMIC). Technologies for global health include both health technologies (ie, vaccines, e–health, genomics) and technologies that "*have health benefits that arise from use outside of health, such as the Internet or irrigation*" considering that "*most health problems are best addressed by a combination of technologies*" [8]. In addition global health relies strongly on computer technologies to use and model the increasing amount of data (data science) for example in worldwide disease surveillance. Innovation in global health also takes place at the social and policy levels [9]. Social and policy innovation for health encompasses all strategies to improve the uptake of technological innovation, to promote health and well–being, and to address broader problems such as access to education. While in the 20th century health innovation used to flow exclusively from HIC to LMIC, HIC can also benefit from innovation in LMIC (reverse innovation) including from the social innovation capacity of communities, and promotes mutual learning for change.

AGH is concerned with sustainability

With rapid population growth combined with unsustainable modes of production and consumption, ever growing constraints apply on the planet. In this early 21st century, humanity is facing the huge challenge to learn to live within planetary boundaries [10]. From anthropocentric models of socio–economic development, we need to include the environmental dimension into the equation and shift to sustainable development. Sustainability science, defined as the study of "*the interactions between natural and social systems, and with how those interactions affect the challenge of sustainability: meeting the needs of present and future generations while substantially reducing poverty and conserving the planet's life support systems*" [11] is an integral part of global health. Health is a prerequisite–good health and well–being are required for people to achieve their full potential–and an outcome of sustainable development. As the health and fate of humanity ultimately depends on Earth's natural systems, AGH is thus essential to shape sustainable development goals, to measure progress toward human well–being, and to improve our understanding of how environmental, social, economic, and health goals can be integrated to preserve planetary health [10,12].

AGH is committed to the normative framework of human rights and equity

Health is an essential part of the broad normative framework of human rights and social justice as affirmed by the World Health Organization preamble. Indeed, several international treaties consider health a human right, which imposes obligations on states to respect, to protect, and to contribute to its progressive realization. Beyond access to health care, the right to health covers social determinants of health, since living conditions are broadly shaped by the distribution of resources and power [13], the rule of law, and levels of liberty, security, and dignity. Central to AGH is understanding the distribution and impact of the unfair and avoidable differences (inequities) in health status between population, genders, and countries, and the reduction of these inequities through action within and beyond the health sector. Universal Health Coverage, the provision of health services with adequate financial protection for all, should thus be enshrined within the broader right to health [14] and the overarching goal of reducing poverty, the main single obstacle to health with 896 million people living with less than a US$ 1.90/day and 2.1 billion below US$ 3.1/day according to the World Bank in 2012 [15].

DEFINITION AND CHALLENGES FOR ACADEMIC INSTITUTIONS

Based on the six principles above, we propose the following definition of AGH: *Within the normative framework of human rights, global health is a system–based, ecological and transdisciplinary approach to research, education, and practice which seeks to provide innovative, integrated, and sustainable solutions to address complex health problems across national boundaries and improve health for all.* This definition first underlines the dynamic complexity which results from our era of interdependence [16]. Within the progressive differentiation of scientific knowledge, it aims to reconcile the century–old divide between medicine and public health, while extending our perspective to other highly relevant fields such as engineering and international relations. Second, this definition corresponds to the perspective of five Swiss academic institutions closely engaged with international Geneva as the main hub of global health governance. While we believe that it reflects the challenges associated with addressing health issues across the world, we consider our work as a proposal to foster further debate with researchers in other countries especially from the global South. Third, translating this definition into concrete projects regarding education, research, and partnerships is key to move AGH forward. **Table 1** summarizes projects based in our five academic institutions which contribute to the conceptualization of global health presented here.

In education, the main challenge is to extend the topics and methods taught both in the curricula of global health in medicine, public health, and engineering, and in other programs granting global health degrees, while maintaining sufficient coherence and disciplinary depth. Mixing students from diverse backgrounds is paramount to foster col-

Table 1. Examples of programs in global health based at five Swiss academic institutions

Program name and institution	Short description
Master of Science and PhD in global health, Institute of Global Health and Global Studies Institute, University of Geneva	As innovative educational programs in global health, the PhD is an executive program based on blended learning (residential weeks in Switzerland, highly intensive distance learning, and accredited MOOCs) while the Master is a transdisciplinary two–year full time program based in Geneva with specialization in other training programs.
EssentialTech Initiative, Swiss Federal Institute of Technology in Lausanne (EPFL)	The aim of this cooperation and research initiative is to foster the development and implementation of essential technologies including medical equipment, water, and sanitation, which can contribute to improve health in LMIC.
Long term partnership with Ifakara, Tanzania, Swiss Tropical and Public Health Institute in Basel (SwissTPH)	The SwissTPH has a long–term collaboration with the Ifakara Health Institute (IHI) in Tanzania, a successful institution for basic and translational health research, education and support in public health. While the IHI has been a Tanzanian institution since 1996, the model of building comparable centers has spread through SwissTPH and partners to other countries in Africa.
Research on chronic diseases, Institute of social and preventive medicine (IUMSP) in Lausanne	The IUMSP specializes in research on epidemiology and prevention of chronic diseases, particularly cancers and cardiovascular diseases as the burden of these conditions are growing in aging societies and requires new public health responses.
Executive Training in Global Health Diplomacy, Global Health Programme, Graduate Institute of International and Development Studies in Geneva	Since 2007, the Global Health Programme offers executive training in global health diplomacy around the world with the aim of bringing together diplomats and health decision–makers to understand their common interests in health as a goal of foreign policy.

laboration across disciplines and to develop the reflexive and synthesizing mind in a competence–based education. Educational models such as interdisciplinary co–teaching and the introduction of existing textbooks for interdisciplinary teaching in the curricula can help. In addition, advances in e–learning and particularly massive open online courses (MOOCs) can effectively complement curricula. MOOCs offer unprecedented opportunities to create large scale horizontally and vertically integrated learning communities.

In research, this definition requires collaborative or transdisciplinary 'team science' with knowledge increasingly produced through teams and networks of scholars. Some key areas of interdisciplinary enquiry are mentioned in **Table 2** while their scope is presented within the wider context of global health in **Figure 1**. As differences of disciplinary cultures and paradigms are common obstacles for interdisciplinary research, dedicated support from academic institutions, funding agencies, and governments can help alleviate these barriers. Leading medical journals already play their part by publishing perspectives from non–medical disciplines although the format for research submissions often still remains too rigid [18]. More importantly, the obstacles associated with an interdisciplinary academic career pathway remain a major issue almost everywhere. Traditional disciplinary candidates are favored when it comes to promotion and tenure for faculty position [19]. As AGH needs to work across academic disciplines, AGH programs may be organized in interfaculty or interdisci-

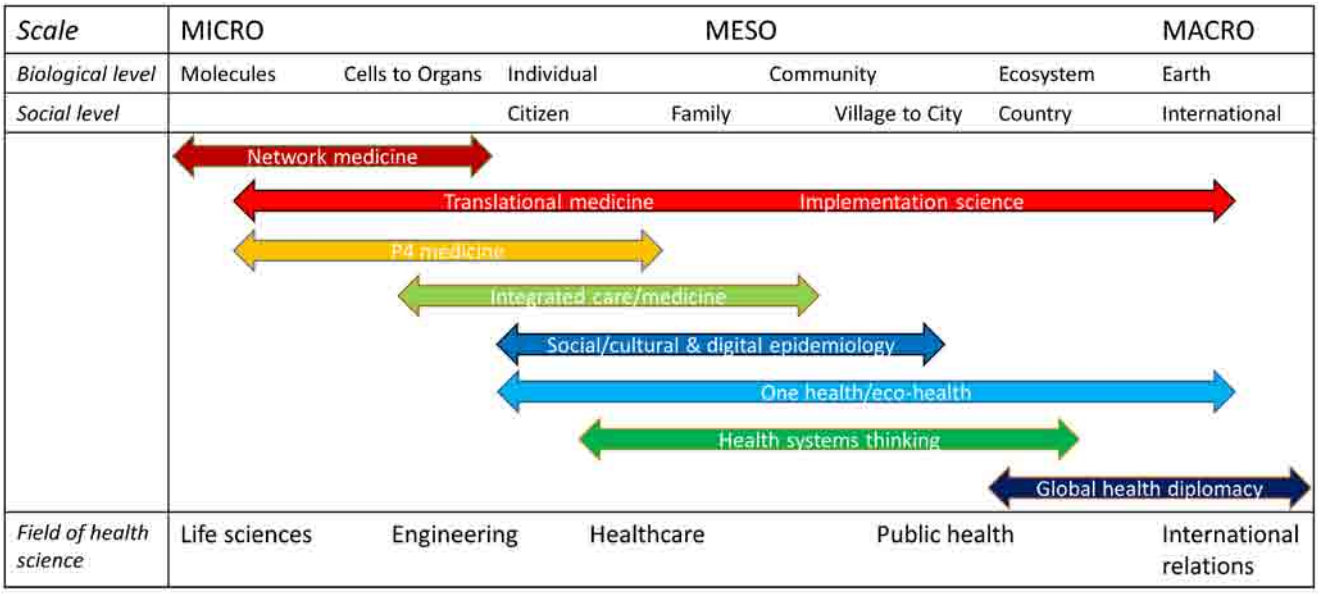


Figure 1. Scope of selected interdisciplinary research and education approaches in academic global health.

Table 2. Selected interdisciplinary research and education approaches relevant to global health

Approaches	Description
Network medicine	Part of network science, network medicine seeks to improve our understanding of disease mechanisms and pathways. It focuses on measuring and analyzing the structures and dynamics of complex molecular networks, which entails relationships between multiple components at the cellular level. Network medicine contributes to better understanding the genetic interlinkages between diseases, and provides insight for new treatments and diagnostics. It also provides a basis to place disease systems into the context of health and social systems.
P4 medicine	Progress and cost reduction in biotechnologies are enabling a more predictive, preventive, personalized, and participatory medicine (P4), which takes into account the genetic background and other specificities of each patient as well as their economic context. The ambition of P4 medicine is to offer customized treatment and improve the detection of diseases before symptoms appear. While medicine has been largely reactive to diseases, P4 is proactively garnering a range of data to maintain well–being.
Translational medicine/ Implementation science	Implementation science (IS) is firmly based on evidence from basic science and corresponds to a continuum of knowledge translation activities, which aims to reduce the science to policy and practice gaps. In medicine, translational medicine is the processes of transforming basic science and technologies from bench to bedside and population. In public health, IS plays a key role in validating health interventions seeking to reach all those who need them in order to improve population/ community health effectively and equitably.
Integrated care/medicine	Integrated care (IC) seeks to address patient problems in holistic ways rather than only through specialized care to improve health care delivery (eg, quality, satisfaction, access). As a bottom–up person–centered perspective, IC responds to the fragmentation of health care delivery due to progressive hyper–specialization of medicine. An example of integrated care is the development of family medicine where the general practitioners play the role of gatekeeper.
Health and social systems thinking	Health systems are complex open systems with several blocks. Thus, health systems thinking focuses on understanding the roles, functions and positions of the systems' building blocks as well as the complex positive or negative feedback loops between these blocks [22]. It provides a framework to strengthen health and social systems, for example through integrated locally tailored interventions between vertical programs and primary health care.
One Health/eco–health	A "One Health" approach seeks to address, in an integrated way, health issues that result from the interplay of multiple human, animal, and environmental factors within a given socio–ecological context. This approach is timely as zoonoses are the main source of emerging and re–emerging infectious diseases (eg, bird flu, SARS, HIV or Ebola) due to several factors such as the ever increasing mobility of human population, disruptions of ecosystems, industrialization of food systems, and socio–political fragility.
Social/cultural and digital epidemiology	While social/cultural epidemiology mixes epidemiology with social theories, digital epidemiology uses a broad range of digital data sources and computer science. Social/cultural epidemiology establishes causal relationships between economic, social and political conditions in which people live as well as health status over their life–course. Digital epidemiology not only provides information about outbreaks and diseases dynamics but also examines and predicts how health and diseases are spread through social ties and networks.
Global health diplomacy	Global health diplomacy (GHD) is concerned with understanding how we collectively deal with cross–border health issues and global challenges through bilateral or multilateral negotiations across different countries, actors, levels and systems. GHD sheds light on the political nature of health, the competing social norms, the evolving role of myriad actors and the complex scientific and political processes that surround any health issue.

plinary centers with joint appointments [20] and/or work as network of actors across institutions and disciplines.

Currently AGH attracts more attention in HIC than in LMIC, reflecting a wider gap in research and education capacities. While the concept of global health originated and diffused widely in HIC, the long term relevance and success of AGH depends on its use and appropriation by academic institutions in LMIC. Three components are essential in this regard. First, international collaboration is critical for both teaching/learning and research in global health but should not "be a one–way street" [21] and should benefit all partners in HIC and LMIC. One challenge is to depart from a long (neo–) colonialist tradition associated with international and tropical medicine. The development of ethical guidelines for educational exchange is a step in the right direction [22]. In addition, AGH requires more South–South collaboration under the leadership of countries such as China or Brazil whose size are critical for capacity building and outreach. Finally, there is a role to play for international academic bodies such as the World Federation of Academic Institutions for Global Health in promoting an inclusive vision of global health and in reflecting on the future of the field based on a broad geographic representation of academic institutions.

CONCLUSION

Within the knowledge society one of the most important challenges faced by academic institutions is to keep their societal relevance. One way forward is to create and develop new intellectual spaces to pursue the production of knowledge across disciplines while drawing on the achievements of two centuries of disciplinary organization of science. As an integration and implementation science, AGH offers such a space to advance our understanding of complex problems and comply with the social responsibility of academic institutions to contribute to societal well–being and sustainable development.

Acknowledgments: This article has benefited from previous discussions with Dr Slim Slama and Dr David Beran. We gratefully acknowledge institutional support from Prof. Jean–Dominique Vassalli, Prof. Yves Flückiger, and Prof. Nicolas Levrat at the University of Geneva.

Authorship declaration: DW and AF conceived and wrote the paper; MT, IK, FP, and GE provided text and edited the manuscript; DW produced the figure and tables.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). The authors declare no competing interests.

REFERENCES

1 Merson MH. University engagement in global health. N Engl J Med. 2014;370:1676-8. Medline:24785204 doi:10.1056/NEJMp1401124

2 Koplan JP, Bond TC, Merson MH, Reddy KS, Rodriguez MH, Sewankambo NK, et al. Towards a common definition of global health. Lancet. 2009;373:1993-5. Medline:19493564 doi:10.1016/S0140-6736(09)60332-9

3 Bammer G. Disciplining interdisciplinarity: integration and implementation sciences for researching complex real–world problems. Canberra: ANU E Press, 2012.

4 Rosenau JN. Distant proximities: Dynamics beyond globalization. Princeton: Princeton University Press, 2003.

5 Popper KR. Conjectures and refutations: The growth of scientific knowledge. New York: Routledge, 2002.

6 Thompson Klein J, Newell WH. Advancing Interdisciplinary Studies. In: Gaff JG, Ratcliff JL, editors. Handbook of the undergraduate curriculum: a comprehensive guide to the purposes, structures, practices, and change. Jossey–Bass Higher and Adult Education Series. San Francisco: ERIC; 1997.

7 Peters DH. The application of systems thinking in health: why use systems thinking? Health Res Policy Syst. 2014;12:51. Medline:25160707 doi:10.1186/1478-4505-12-51

8 Howitt P, Darzi A, Yang GZ, Ashrafian H, Atun R, Barlow J, et al. Technologies for global health. Lancet. 2012;380:507-35. Medline:22857974 doi:10.1016/S0140-6736(12)61127-1

9 Kickbusch I. Policy innovations for health. In: Kickbusch I, editor. Policy innovation for health. New York: Springer, 2009.

10 Whitmee S, Haines A, Beyrer C, Boltz F, Capon AG, de Souza Dias BF, et al. Safeguarding human health in the Anthropocene epoch: report of The Rockefeller Foundation–Lancet Commission on planetary health. Lancet. 2015;386:1973-2028. Medline:26188744 doi:10.1016/S0140-6736(15)60901-1

11 Kates RW. What kind of a science is sustainability science? Proc Natl Acad Sci U S A. 2011;108:19449-50. Medline:22114189 doi:10.1073/pnas.1116097108

12 Haines A, Alleyne G, Kickbusch I, Dora C. From the Earth Summit to Rio+20: integration of health and sustainable development. Lancet. 2012;379:2189-97. Medline:22682465 doi:10.1016/S0140-6736(12)60779-X

13 Ottersen OP, Dasgupta J, Blouin C, Buss P, Chongsuvivatwong V, Frenk J, et al. The political origins of health inequity: prospects for change. Lancet. 2014;383:630-67. Medline:24524782 doi:10.1016/S0140-6736(13)62407-1

14 Ooms G, Brolan C, Eggermont N, Eide A, Flores W, Forman L, et al. Universal health coverage anchored in the right to health. Bull World Health Organ. 2013;92:2–A. Medline:23397341 doi: 10.2471/BLT.12.115808

15 The World Bank. Poverty overview. Available: http://www.worldbank.org/en/topic/poverty/overview. Accessed: 28 January 2016.

16 Frenk J, Gomez–Dantes O, Moon S. From sovereignty to solidarity: a renewed concept of global health for an era of complex interdependence. Lancet. 2014;383:94-7. Medline:24388312 doi:10.1016/S0140-6736(13)62561-1

17 de Savigny D, Adam T. Systems thinking for health systems strengthening. Geneva: World Health Organization, 2009.

18 Ooms G. From international health to global health: how to foster a better dialogue between empirical and normative disciplines. BMC Int Health Hum Rights. 2014;14:36. Medline:25494979 doi:10.1186/s12914-014-0036-5

19 Rhoten D, Parker A. Risks and rewards of an interdisciplinary research path. Science. 2004;306:2046. Medline:15604393 doi:10.1126/science.1103628

20 Committee on Facilitating Interdisciplinary Research, National Academy of Sciences, National Academy of Engineering, Institute of Medicine. Facilitating interdisciplinary research. Washington DC: The National Academies Press, 2004.

21 Miranda JJ, Garcia PJ, Lescano AG, Gotuzzo E, Garcia HH. Global health training–one way street? Am J Trop Med Hyg. 2011;84:506;author reply 507. Medline:21363994 doi:10.4269/ajtmh.2011.10-0694a

22 Peluso MJ, Encandela J, Hafler JP, Margolis CZ. Guiding principles for the development of global health education curricula in undergraduate medical education. Med Teach. 2012;34:653-8. Medline:22830323 doi:10.3109/0142159X.2012.687848

journal of global health

Can innovative ambulance transport avert pregnancy–related deaths? One–year operational assessment in Ethiopia

Hagos Godefay[1], John Kinsman[2], Kesetebirhan Admasu[3], Peter Byass[2,4]

[1] Tigray Regional Health Bureau, Mekele, Ethiopia
[2] Umeå Centre for Global Health Research, Epidemiology and Global Health, Dept of Public Health and Clinical Medicine, Umeå University, 90187 Umeå, Sweden
[3] Federal Ministry of Health, Addis Ababa, Ethiopia
[4] Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Background To maximise the potential benefits of maternity care services, pregnant women need to be able to physically get to health facilities in a timely manner. In most of sub–Saharan Africa, transport represents a major practical barrier. Here we evaluate the extent to which an innovative national ambulance service in Ethiopia, together with mobile phones, may have been successful in averting pregnancy–related deaths.

Methods An operational assessment of pregnancy–related deaths in relation to utilisation of the new national ambulance service was undertaken in six randomly selected Districts in northern Ethiopia. All 183 286 households in the six randomly selected Districts were visited to identify live–births and deaths among women of reproductive age that occurred over a one–year period. The uptake of the new ambulance transport service for women's deliveries in the same six randomly selected Districts over the same period was determined retrospectively from ambulance log books. Pregnancy–related deaths as determined by the World Health Organization (WHO 2012) verbal autopsy tool [13] and the InterVA–4 model [14] were analysed against ambulance utilisation by District, month, local area, distance from health facility and mobile network coverage.

Findings A total of 51 pregnancy–related deaths and 19 179 live–births were documented. Pregnancy–related mortality for Districts with above average ambulance utilisation was 149 per 100 000 live–births (95% confidence interval CI 77–260), compared with 350 per 100 000 (95% CI 249–479) for below average utilisation ($P=0.01$). Distance to a health facility, mobile network availability and ambulance utilisation were all significantly associated with pregnancy–related mortality on a bivariable basis. On a multivariable basis, ambulance non–utilisation uniquely persisted as a significant determinant of mortality (mortality rate ratio 1.97, 95% CI 1.05–3.69; $P=0.03$).

Conclusions The uptake of freely available transport in connection with women's obstetric needs correlated with substantially reduced pregnancy–related mortality in this operational assessment, though the design did not allow cause and effect to be attributed. However, the halving of pregnancy–related mortality associated with ambulance uptake in the sampled Districts suggests that the provision of transport to delivery facilities in Africa may be a key innovation for delivering maternal health care, which requires wider consideration.

Correspondence to:
Hagos Godefay
Tigray Regional Health Bureau
P.O. Box 7
Mekele
Ethiopia
hgodefay@yahoo.com

Despite major international concerns around maternal health and institutional delivery rates [1], little innovative thought has been given to the logistic issues of getting African women in to appropriate institutions in a timely fashion. Expecting rural women in labour to walk several kilometres to a facility, possibly at night and in bad weather, is unrealistic. Thirty years ago a startling but small–scale finding from The Gambia found that there had been no maternal deaths for eight years in a group of small villages where resident midwifery services and immediate access to referral transport had been made freely available, when otherwise 16 maternal deaths might have been expected [2]. Though that innovation was widely considered unscalable and unsustainable, millions of pregnant African women have died in the intervening decades, partly from not being able to reach maternity services [3,4].

The principle that effective transport for obstetric health emergencies is essential is not a matter for debate in most settings. However, that thinking has not translated widely into sub–Saharan Africa, where access to obstetric care remains a major barrier, with a lack of transportation and other infrastructure. Only a few sub–Saharan countries have considered and evaluated the provision of ambulances to facilitate access to obstetric care even in emergencies, for example in Burundi [5], Uganda [6,7], and South Africa [8].

In Ethiopia, the 2011 DHS report found 9.9% of births nationally during the previous five years were delivered at a health facility (10.6% in Tigray Region) and 71.1% of women mentioned lack of transport to a facility as a major barrier (52.4% in Tigray Region) [9]. Just 0.1% of rural households owned any kind of motorised transport [9]. The challenges of increasing institutional delivery rates and access to emergency obstetric care have now been recognised. The Ethiopian government health service is now unique in sub–Saharan Africa in providing four–wheel drive ambulances in every rural District (areas each covering around 150 000 people), and in making the ambulances available on a 24–hour, 7–day basis to transfer any woman in labour or experiencing other obstetric difficulties to appropriate health facilities. Parallel innovation in mobile telephony, which has brought widespread network coverage to rural Ethiopia, completes the picture by providing a means to call ambulances when needed [10]. This highly innovative approach to improving maternity care has been rolled out nationally in Ethiopia since 2012.

A total of 1250 ambulances have been distributed, with at least one ambulance per District and nearly half of Districts (larger ones) getting two. The total investment to achieve this was about US $50 million. Before ambulances were deployed, the Ethiopian Federal Ministry of Health signed an agreement with governments of the nine regional states and the two autonomous city administrations (Addis Ababa and Dire Dawa) to regulate use. The agreement entailed three important commitments. First the regional governments committed to allocate budgets to cover the running costs for the ambulances; second they committed to replacing the ambulances after five years; and third they agreed to make ambulance services available free of charge. Furthermore, a number of town–hall and community meetings were held to inform communities about these commitments. Once the ambulances were delivered, frequent reminders about the commitments were sent out through local mass media. Laws to enact the ambulance service provision and ensure and safeguard the proper utilisation of ambulances have also been passed by the regional cabinets.

These developments in Ethiopia therefore provided a unique opportunity to contribute to filling the current evidence gap on the provision of non–emergency obstetric transport in Africa. An operational assessment of the effectiveness of the innovative ambulance service for transporting women to facilities and averting pregnancy–related deaths was conducted, based on two interlinked data sources. Tigray Regional Health Bureau, in the north of Ethiopia, had previously undertaken a one–year representative randomised population survey of pregnancy–related mortality [11], which could be linked at the local community (*tabia*) level to data on actual ambulance utilisation for obstetric care, as a means of assessing the extent to which ambulance utilisation may have averted pregnancy–related deaths.

METHODS

The detailed methodology for the pregnancy–related mortality survey design has been presented elsewhere [12]. Briefly, the 34 rural Districts (*woreda*) in Tigray Region form six geographic Zones, and one District per Zone was randomly selected as a stratified sample, covering a population of 843 115. A two–stage retrospective household mortality survey was carried out in mid–2013 by community health staff in the selected Districts, following up deaths among women of reproductive age (15–49 years) over a one–year period (from the ninth month of Ethiopian year 2004 to the eighth month of Ethiopian year 2005, corresponding to 9 May 2012 to 8 May 2013 in the international calendar) using the WHO 2012 verbal autopsy tool [13], and deriving cause of death using the corresponding InterVA–4 model [14]. The date, place and personal details were recorded for all deaths. The same survey captured the corresponding number of live–births. Maternal mortality ratio (MMR) is defined here as pregnancy–related deaths per 100 000 live–births, as adopted by the Demographic and Health Survey programme [9].

As a separate exercise, the vehicle log books for the ambulances in the six randomly selected districts underwent retrospective data capture for the same one–year period. The data for each trip included whether it was connected with a delivery; from which community it originated; patient's name; distance travelled; destination and date. The completeness of the vehicle logs was verified using the odometer readings for the start and end of every journey in each vehicle, before the journeys specifically relating to women delivering were extracted.

Descriptive statistics were used to characterise the patterns of ambulance transportation in relation to pregnancy–related deaths and determinants including District, month, local area, distance from health facility and mobile telephone coverage. Multivariable Poisson regression modelling was used to assess competing contributions of distance to facility, mobile network availability and ambulance utilisation on pregnancy–related mortality at local area level (typically neighbourhoods of around 5000 people).

RESULTS

During the one–year period, 51 pregnancy–related deaths and 19 179 live–births were identified in the community survey of 183 286 households in the six study Districts, as previously described [11]. Corresponding data from ambulance log books detailed 4779 trips related to deliveries, covering 178 736 km. **Figure 1** shows MMR in six study Districts within Tigray Region, also showing numbers of pregnancy–related deaths that occurred in each local area during the study period.

Table 1 summarises births, deaths and MMR across the six study Districts. In two Districts, Alamata and Hintalo Wajirat, ambulances were not yet available at the start of the

Table 1. Ambulance utilisation in relation to delivery care across six rural Districts in Tigray Region, Ethiopia, over a one–year period

Utilisation	Alamata*	Hintalo Wajirat*	Laelay Adiyabo	Saesi Tsaedaemba	Tahtay Maychew	Welkayat
Live–births	2364 (2021)	3516 (934)	2637	2990	2697	4975
Pregnancy–related deaths	8 (5)	8 (1)	7	3	1	24
MMR per 100 000 live–births	338 (247)	228 (107)	265	100	37	482
Ambulance trips for deliveries	1205	268	496	1590	914	306
% deliveries using ambulances	51.0 (59.6)	7.6 (28.7)	18.8	53.2	33.9	6.2
Ambulance km for deliveries	43 633	12 024	20 574	58 562	36 977	6966
Ambulance mean km per delivery	36.2	44.9	41.5	36.8	40.5	22.8
Ambulance mean km per live–birth	18.5 (21.6)	3.4 (12.9)	7.8	19.6	13.7	1.4
% ambulance trips to hospitals	27.8	1.1	no data	6.7	47.0	4.6

*Figures in brackets for Alamata and Hintalo Wajirat Districts reflect only the part of the year during which ambulances were available.

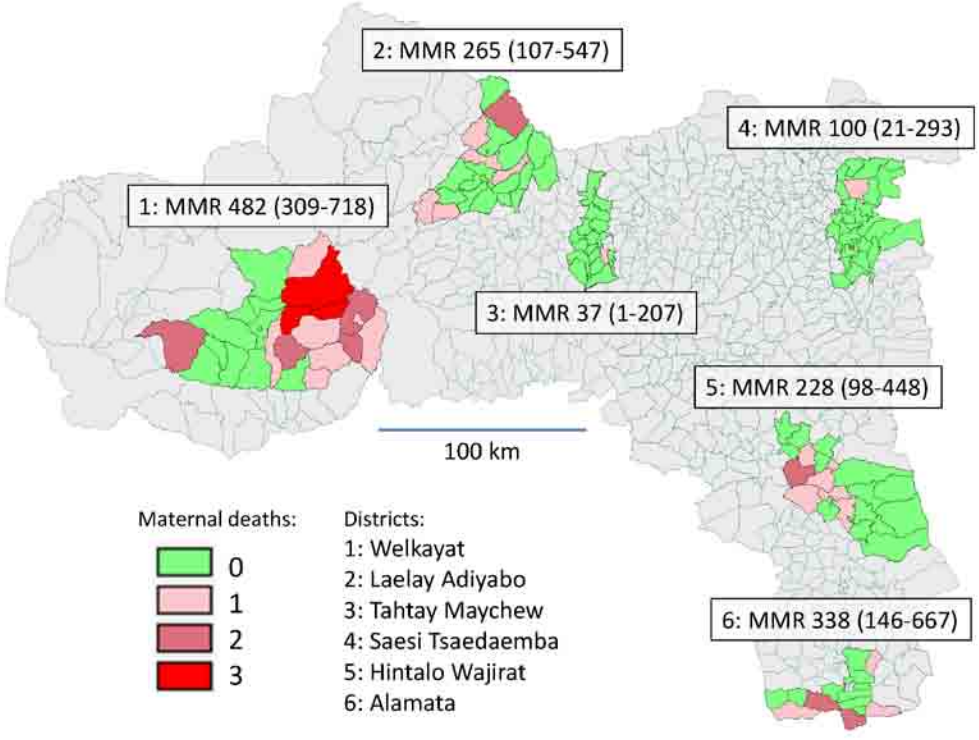


Figure 1. Map of Tigray Region,Ethiopia, showing the six study Districts and maternal mortality ratios (MMR) with 95% confidence intervals for each District. Shading indicates the numbers of maternal deaths in each local area (*tabia*) within each District.

study, so the parameters for those Districts are shown both for the whole year and, in italics, for only the period during which ambulances were available. The overall percentage of deliveries using ambulances was 24.9%, ranging across the Districts from 6% to 53%. The mean distance per delivery where ambulances were used was 37.4 km, ranging from 23 to 45 km across the Districts. The proportions of ambulance trips going to a hospital rather than a health centre were higher in Alamata District, which contains a hospital, and in Tahtay Maychew District, which is close to Axum hospital.

The proportion of deliveries using an ambulance was calculated for each month and District as a means of tracking patterns in ambulance utilisation over the one–year period. This is shown in **Figure 2** for each District, together with the numbers of pregnancy–related deaths that occurred in each month. **Figure 3** shows a comparison between Districts above and below the overall 24.9% level of ambulance utilisation. Aggregated MMRs and 95% CIs for the two groups of Districts are shown by the red bars, 350 per 100 000 (95% CI 249–479) and 149 per 100 000 (95% CI 77–260) respectively; $P=0.01$.

Figure 4 shows the104/131 local areas utilising ambulance trips in connection with deliveries during the year, together with MMR by utilisation. **Figure 5** shows the 103/131 local areas with mobile telephone network coverage. Pregnancy–related mortality was significantly lower both in the local areas using ambulances (MMR 202, 95%CI 135–291 vs 468, 95%CI 293–709; $P=0.006$) and in the local areas covered by mobile telephone networks (MMR 209, 95%CI 141 to 299 vs 447, 95%CI 277–683; $P=0.014$).

A Poisson regression model was constructed including all 131 local areas in the survey, with the number of pregnancy–related deaths as the dependent variable and the number of live–births as the exposure term for each area. Independent variables were distance from the District Health Centre, availability of the mobile telephone network, and ambulance utilisation. **Table 2** shows bivariable and multivariable results from this modelling. In the bivariable model, all the independent variables were significantly associated with pregnancy–related deaths. In the multivariable model, ambulance non–utilisation emerged as the overall significant factor associated with pregnancy–related deaths (mortality rate ratio 1.97, 95% CI 1.05–3.69; $P=0.03$), while distance to the District Health Centre lost most of its effect.

By considering **Figure 3**, it is possible to extrapolate to putative changes in ambulance utilisation in below–average Districts to bring them to above–average levels. The lower group had 1070 utilisations out of 11 128 deliveries, compared with 3709/8051 in the higher group. Thus an additional 3211 deliveries in the lower group would be needed to achieve the same rate of utilisation. Hypothetically assuming that this would have the same effect on pregnancy–related mortality as the observed difference between the lower and higher utilisation groups, a reduction from an

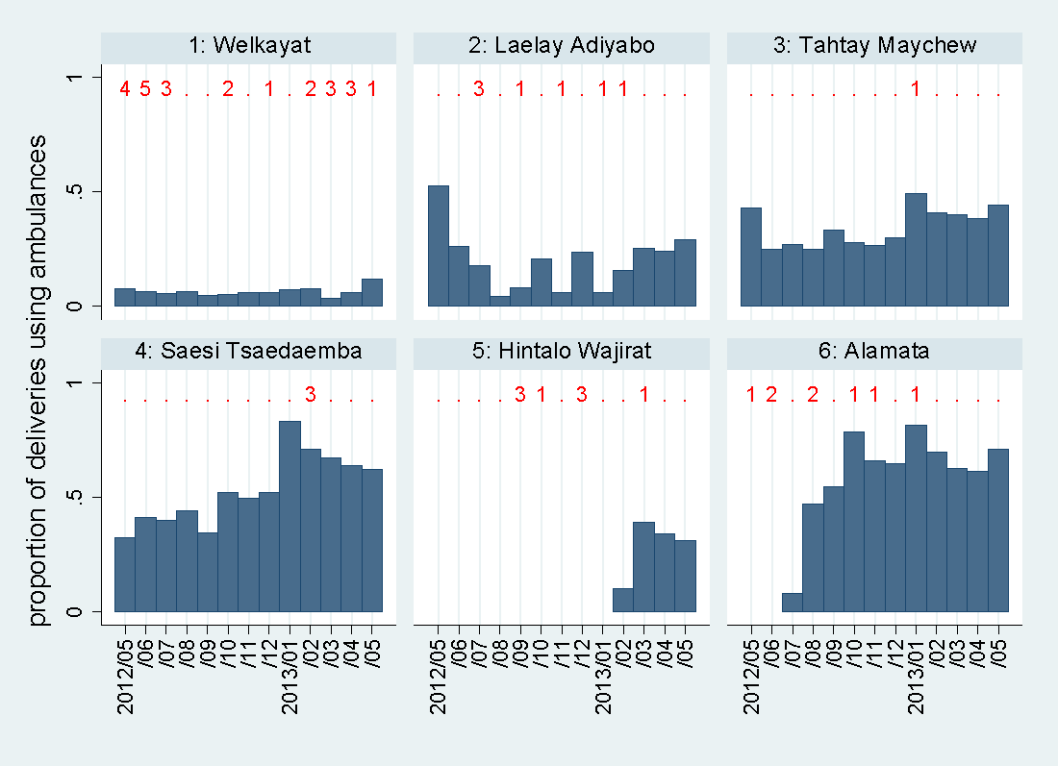


Figure 2. Proportions of deliveries using ambulance transport and numbers of pregnancy–related deaths (figures in red), by month and District.

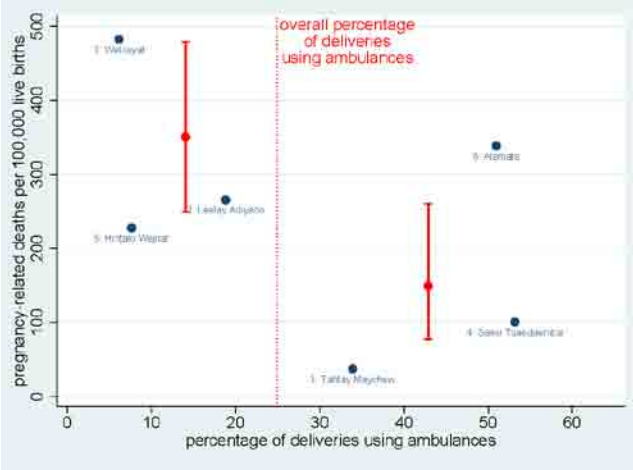


Figure 3. Maternal mortality ratios (MMR) and proportions of deliveries using ambulances for each District. Red bars indicate aggregated MMRs and 95% confidence intervals for Districts achieving above and below the overall percentage (24.9%) of deliveries using ambulances.

MMR of 350 to 149 in the group observed to have lower utilisation would avert 39 × (149/350) = 17 deaths. If each additional ambulance trip involved the mean observed distance of 37.4 km, usage per death averted would be approximately 7000 km. Conversely, therefore, the 178 736 ambulance kilometres that were actually deployed during the study might have averted around 26 pregnancy–related deaths. If that were the case, then the internally adjusted overall MMR for Tigray Region in the absence of the new ambulance service would have been 401 per 100 000 live–births, rather than the 266 per 100 000 observed [11].

DISCUSSION

These results clearly show substantially lower pregnancy–related mortality in places and periods where free ambulance transport was used by women in connection with their deliveries. We entirely accept that an operational assessment of this kind cannot demonstrate statistically that ambulance utilisation caused reductions in pregnancy–related mortality. Nevertheless, observed variations in pregnancy–related mortality were very substantial, and highly correlated with ambulance utilisation. Since, in most of the world, women's means of transport to health care facilities for delivery are taken for granted as an essential component of health systems, it is reasonable to suppose that the availability of transport might be just as essential in sub–Saharan Africa.

The physical obstacles to reaching health facilities, and the lack of available transport options, are probably most extreme in Africa. The adjusted MMR estimate of 401 per 100 000 for Tigray in the absence of the ambulance service was consistent with international estimates of MMR for Ethiopia before ambulances were deployed [3], and also similar to MMR survey results from the Southern Nations Nationalities and Peoples' Regional State in the pre–ambulance period [15]. The lower MMR observed in Tigray when

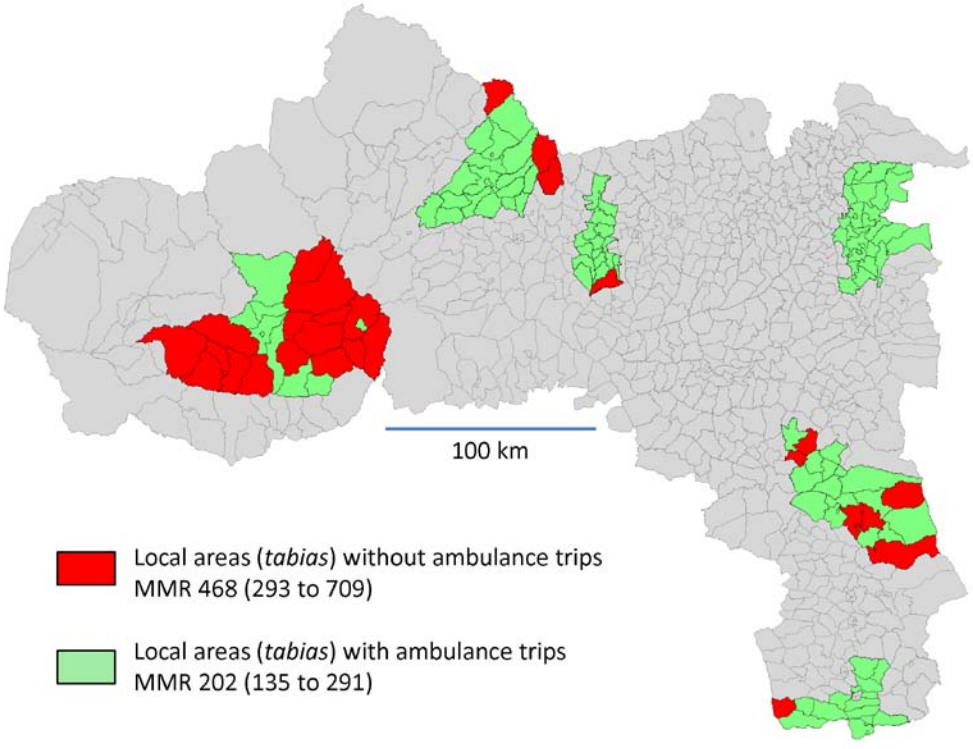


Figure 4. Ambulance utilisation within the six study Districts in Tigray Region, Ethiopia, by local area (*tabia*) and associated maternal mortality ratios (MMR) with 95% confidence intervals.

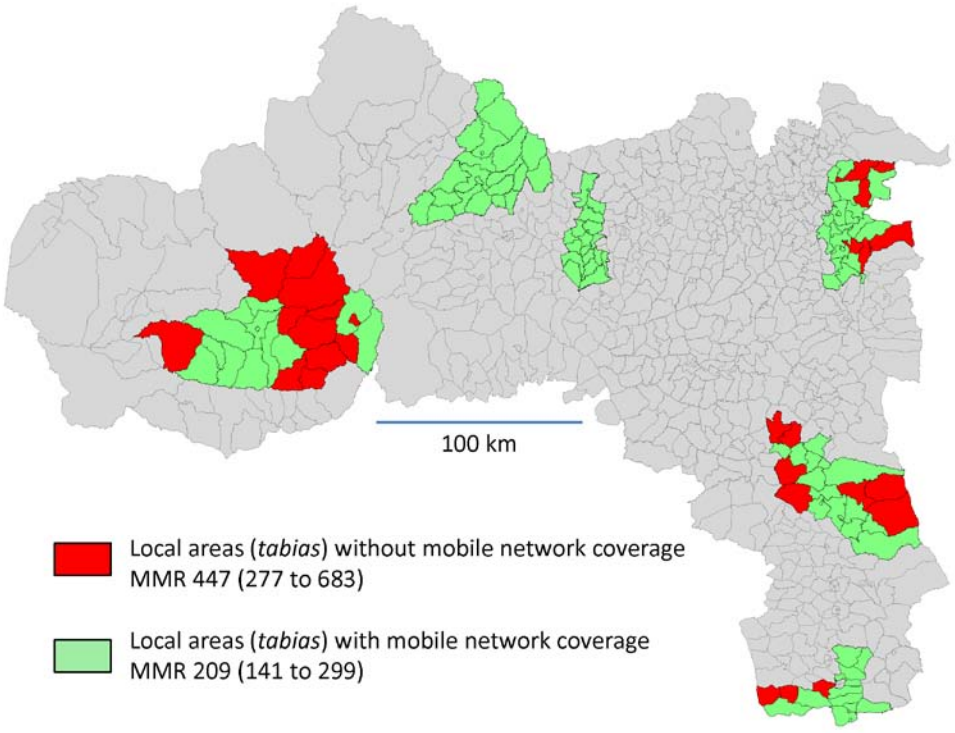


Figure 5. Mobile telephone network coverage within the six study Districts in Tigray Region, Ethiopia, by local area (*tabia*) and associated maternal mortality ratios (MMR) with 95% confidence intervals.

Table 2. Associations between maternal deaths and distance to District Health Centre, availability of mobile telephone network and utilisation of ambulance transport for deliveries in 131 local areas (*tabia*) in Tigray Region, Ethiopia, using a Poisson regression model.

Factor	Level	Number of local areas	Number of pregnancy–related deaths	Bivariable maternal death rate ratio (95% CI)	Multivariable maternal death rate ratio (95% CI)
Distance to District Health Centre	<15 km	32	8	Ref	Ref
	15–30 km	69	23	1.45 (0.65–3.24)	1.23 (0.54–2.79)
	>30 km	30	20	2.33 (1.03–5.28)*	1.15 (0.43–3.11)
Mobile telephone network	Available	103	30	Ref	Ref
	Not available	28	21	2.14 (1.22–3.73)*	1.78 (0.91–3.41)
Ambulance transport for deliveries	Utilised	104	29	Ref	Ref
	Not utilised	27	22	2.31 (1.33–4.03)*	1.97 (1.05–3.69)*

CI – confidence interval

*95% confidence interval excludes unity, P<0.05.

ambulances were available probably accrued from a combination of substantially increasing the proportion of institutional deliveries by transporting women and diminishing transport delays for a smaller proportion of women in acute difficulties. The remote and difficult terrain in Tigray Region also makes it very difficult to undergo transfer to a health facility once complications become evident, and so the provision of a universal service before complications arise may avoid potentially life–threatening transport delays at a later stage. The detailed individual data that would be needed to tease out these different factors were not available.

Ideally this study would have been designed as an intervention with randomised allocation to ambulance transport. However, there would be very serious ethical difficulties in designing such an evaluation, given that the world in general implicitly assumes that availability of effective transportation to access obstetric services is a basic requirement, which may even be considered as a human rights issue [16]. Given the impossibility of mounting a controlled trial, the only available option was to make a transparent operational assessment in the context of the introduction of the new Ethiopian national policy, designed as a retrospective observational study [17]. That said, serendipitous operational delays in ambulance deployment in two of the six Districts, as well as varying effectiveness and coverage by the ambulance service between the randomly selected Districts, enhanced opportunities for comparison.

This operational assessment was only carried out in one Region, where survey data on maternal deaths were already available [11]. Nevertheless it is important to note that Ti-

gray covers some of the most mountainous and hard–to–reach areas of Ethiopia. The ambulance programme in Tigray, by taking a quarter of all delivering women to facilities, made a major impact on previously low institutional delivery rates [9]. Districts included in this operational assessment had been randomly selected for the previous mortality survey [12], and results showed that there were substantial variations between Districts in many parameters. Even though 95% confidence intervals around MMRs are fairly wide in some instances, reflecting the relative rarity of maternal deaths, sufficient numbers were included in the survey to detect important differences. The previous mortality survey and the ambulance log book data capture were undertaken as two completely independent exercises, with neither making use of routine data reporting systems, to ensure integrity and independence. Although retrospective mortality surveys always carry some risk of under–reporting, there is no reason to suppose that any such bias would be correlated with ambulance utilisation. No direct national assessment of pregnancy–related mortality has been made since the ambulance programme was introduced, but our estimates after adjusting for ambulance utilisation reflect pre–ambulance levels as both estimated nationally and as surveyed in another Region [3,15].

Previous studies of obstetric emergency transport in sub–Saharan Africa have been very limited, and none have involved non–emergency provision. A small–scale study in Burundi estimated that a large proportion of cases with complications were transferred by ambulance, but was not able to measure the effect on maternal mortality [5]. Similarly in Ruhira, Uganda, a small study of the use of a single ambulance for emergencies only, with two–thirds of call–outs being for complicated obstetric cases, concluded that it could be cost–effective to provide such a service, but was not able to evaluate changes in maternal mortality [6]. A larger pre– and post– intervention study in Oyam, Uganda, looked at the effect of introducing a single ambulance, and concluded that there was an increased rate of Caesarean sections after the ambulance became available, but also did not evaluate changes in maternal mortality [7]. In Free State, South Africa, the effect of introducing emergency obstetric transport across the Province was evaluated in terms of institutional maternal mortality, with MMR falling from 279 per 100 000 to 152, a risk ratio of 0.54 (95% CI 0.40–0.74) [8], a similar reduction as observed here.

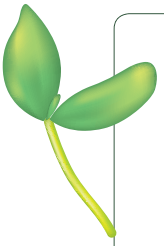
Two factors which might reasonably be supposed to influence pregnancy–related mortality in the absence of a free transport service were distance to District health facilities and the availability of mobile telephone networks. In Tanzania, maternal mortality and distance to hospital were strongly related [18]. In the United States, where availability of transport is presumably not a major issue, distance from hospital has still been shown to be highly associated with maternal deaths [19]. In Oxfordshire, United Kingdom, the use of mobile phones in emergency situations was shown to reduce the risk of death at the scene, though not specifically for pregnancy–related deaths [20]. In the current study, both of these factors could be evaluated on a local area basis, and both individually were highly associated with pregnancy–related deaths, as shown in **Table 2**. However, introducing ambulance use into a multivariable model with these two factors resulted in transport service utilisation emerging as the strongest, and only significant, factor associated with reductions in pregnancy–related deaths. Mobile phone availability retained some of its effect, not surprisingly given that mobile phones are an essential component of the overall ambulance transport policy, providing the only opportunity to call ambulances to communities with no other means of communication. Mobile telephony is a very significant development which has spread across rural Africa relatively recently, and which must be regarded as a huge public health gain. In Ethiopia, EthioTelecom is the sole network provider, with widespread coverage other than in particularly hard–to–reach areas. It may be appropriate for Ministries of Health to put pressure on mobile communications operators to extend coverage into all inhabited areas.

While extrapolating from these results might be speculative, our estimate that approximately 7000 ambulance–kilometres averted one pregnancy–related death suggests that ambulances may be a practical and cost–effective means to substantially reduce the persistently high levels of pregnancy–related mortality across sub–Saharan Africa. Assuming that a free ambulance service such as that now implemented in Ethiopia might cost around US $1 per kilometre to provide (fuel, servicing, depreciation and staff), then an annual budget of US $1 billion for free ambulance transport might avert a substantial proportion of the estimated 180 000 maternal deaths in Africa. Of course this would not simply be a matter of money–resources would have to be translated into a committed and effective transport service, which may not be easy, and local conditions and circumstances would have to be considered. But it seems clear that effective obstetric transport can make a substantial and cost–effective difference to pregnancy–related mortality–possibly achieving as much as two–thirds of the 75% reduction called for by Millennium Development Goal 5 (MDG5) [3].

CONCLUSIONS

Without doubt the possibility for women to be readily transported to health facilities, both for routine deliveries and in obstetric emergencies, is a critical component of providing effective maternal care. Our results show that transport and communication innovations in Ethiopia cor-

related with appreciably reduced pregnancy–related mortality. This was achieved through the provision of four–wheel drive ambulances on a 24/7 basis, which could be called via the mobile telephone network. The magnitude of mortality differences correlated with ambulance utilisation amounted to a considerable proportion of the MDG5 target for maternal mortality reduction. Although this assessment only covered one Region in Ethiopia, the magnitude of the observed reduction in pregnancy–related deaths, plus the commonplace notion that women have a right to be able to physically get to maternity services, underlines the urgent necessity of considering the provision of innovative obstetric transport across Africa.



Acknowledgements: We acknowledge the contribution of Ezedin Adem and Gebreamlak Gidey at Tigray Health Bureau for supervising field work and data entry, and appreciate the cooperation of all survey participants.

Ethical approval: Ethical approval for the study was granted by the Institutional Review Board (IRB) of the College of Health Sciences of Mekele University, Ethiopia, and participants gave verbal informed consent to verbal autopsy interviews.

Funding: Tigray Regional Health Bureau funded this operational assessment as part of its routine work to assess health services in the Region, including the pregnancy–related mortality survey and capturing the ambulance data. The Swedish Research Council provided a Swedish Research Links collaboration grant C0615601 which facilitated external participation in this operational assessment.

Authorship declaration: HG & PB substantially contributed to the conception and design of the work, the acquisition, analysis, and the interpretation of data; all authors contributed to drafting the work and revising it critically for important intellectual content; and all authors gave final approval to the version to be published and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Competing Interests: HG is employed by the Government of Ethiopia as Head of the Health Bureau for Tigray Region. KA is employed by the Government of Ethiopia as Minister of Health at the Federal Ministry of Health. All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). The authors declare no competing interests.

REFERENCES

1 World Health Organization. Strategies toward ending preventable maternal mortality (EPMM). WHO: Geneva, 2015. Available: http://apps.who.int/iris/bitstream/10665/153544/1/9789241508483_eng.pdf. Accessed: 15 September 2015.

2 Lamb WH, Foord FA, Lamb CMB, Whitehead RG. Changes in maternal and child mortality rates in three isolated Gambian villages over ten years. Lancet. 1984;ii:912-4. Medline:6148628 doi:10.1016/S0140-6736(84)90664-0

3 WHO, UNICEF, UNFPA, The World Bank, United Nations Population Division. Trends in maternal mortality: 1990 to 2013. Geneva: WHO, 2014. Available: http://apps.who.int/iris/bitstream/10665/112682/2/9789241507226_eng.pdf. Accessed: 15 September 2015.

4 Pacagnella RC, Cecatti JG, Parpinelli MA, Sousa MH, Haddad SM, Costa ML, et al. Delays in receiving obstetric care and poor maternal outcomes: results from a national multicentre cross–sectional study. BMC Pregnancy Childbirth. 2014;14:159. Medline:24886330

5 Tayler-Smith K, Zachariah R, Manzi M, Van den Boogaard W, Nyandwi G, Reid T, et al. An ambulance referral network improves access to emergency obstetric and neonatal care in a district of rural Burundi with high maternal mortality. Trop Med Int Health. 2013;18:993-1001. Medline:23682859 doi:10.1111/tmi.12121

6 de Ramirez SS, Doll J, Carle S, Anest T, Arii M, Hsieh YH, et al. Emergency response in resource–poor settings: a review of a newly–implemented EMS system in rural Uganda. Prehosp Disaster Med. 2014;29:311-6. Medline:24735913 doi:10.1017/S1049023X14000363

7 Mucunguzi S, Wamani H, Lochoro P, Tylleskär T. Effects of improved access to transportation on emergency obstetric care outcomes in Uganda. Afr J Reprod Health. 2014;18:87-94. Medline:25508044

8 Schoon MG. Impact of inter–facility transport on maternal mortality in the Free State Province. S Afr Med J. 2013;103:534-7. Medline:23885734 doi:10.7196/samj.6828

9 Central Statistical Agency [Ethiopia] and ICF International. Ethiopia Demographic and Health Survey 2011. Addis Ababa, Ethiopia and Calverton, Maryland, USA: Central Statistical Agency and ICF International, 2012. Available: http://dhsprogram.com/pubs/pdf/FR255/FR255.pdf. Accessed: 15 September 2015.

REFERENCES

10 Byass P, D'Ambruoso L. Cellular telephone networks in developing countries. Lancet. 2008;371:650. Medline:18295022 doi:10.1016/S0140-6736(08)60301-3

11 Godefay H, Byass P, Kinsman J, Mulugeta A. Understanding maternal mortality from top–down and bottom–up perspectives: the case of Tigray Region, Ethiopia. J Glob Health. 2015;5:010404. Medline:25674351 doi:10.7189/jogh.05.010404

12 Godefay H, Abrha A, Kinsman J, Myléus A, Byass P. Undertaking cause–specific mortality measurement in an unregistered population: an example from Tigray Region, Ethiopia. Glob Health Action. 2014;7:25264. Medline:25217254

13 Leitao J, Chandramohan D, Byass P, Jakob R, Bundhamcharoen K, Choprapawon C, et al. Revising the WHO verbal autopsy instrument to facilitate routine cause–of–death monitoring. Glob Health Action. 2013;6:21518. Medline:24041439

14 Byass P, Chandramohan D, Clark SJ, D'Ambruoso L, Fottrell E, Graham WJ, et al. Strengthening standardised interpretation of verbal autopsy data: the new InterVA–4 tool. Glob Health Action. 2012;5:1. Medline:22944365

15 Yaya Y, Data T, Lindtjørn B. Maternal mortality in rural South Ethiopia: outcomes of community–based birth registration by Health Extension Workers. PLoS ONE. 2015;10:e0119321. Medline:25799229 doi:10.1371/journal.pone.0119321

16 Hammonds R, Ooms G. The emergence of a global right to health norm–the unresolved case of universal access to quality emergency obstetric care. BMC Int Health Hum Rights. 2014;14:4. Medline:24576008 doi:10.1186/1472-698X-14-4

17 The PLOS Medicine Editors. Transparency in reporting observational studies: reflections after a year. PLoS Med. 2015;12:e1001896. Medline:26506432 doi:10.1371/journal.pmed.1001896

18 Hanson C, Cox J, Mbaruku G, Manzi F, Gabrysch S, Schellenberg D, et al. Maternal mortality and distance to facility–based obstetric care in rural southern Tanzania: a secondary analysis of cross–sectional census data in 226 000 households. Lancet Glob Health. 2015;3:e387-95. Medline:26004775

19 Frölich MA, Banks C, Brooks A, Sellers A, Swain R, Cooper L. Why do pregnant women die? A review of maternal deaths from 1990 to 2010 at the University of Alabama at Birmingham. Anesth Analg. 2014;119:1135-9. Medline:25329025 doi:10.1213/ANE.0000000000000457

20 Wu O, Briggs A, Kemp T, Gray A, MacIntyre K, Rowley J, et al. Mobile phone use for contacting emergency services in life–threatening circumstances. J Emerg Med. 2012;42:291-8.e3. Medline:22142669 doi:10.1016/j.jemermed.2011.02.022

journal of global health

PAPERS

Setting health research priorities using the CHNRI method: IV. Key conceptual advances

Igor Rudan[1]

[1] Centre for Global Health Research, The Usher Institute for Population Health Sciences and Informatics, The University of Edinburgh, Edinburgh, Scotland, UK

Introduction Child Health and Nutrition Research Initiative (CHNRI) started as an initiative of the Global Forum for Health Research in Geneva, Switzerland. Its aim was to develop a method that could assist priority setting in health research investments. The first version of the CHNRI method was published in 2007–2008. The aim of this paper was to summarize the history of the development of the CHNRI method and its key conceptual advances.

Methods The guiding principle of the CHNRI method is to expose the potential of many competing health research ideas to reduce disease burden and inequities that exist in the population in a feasible and cost–effective way.

Results The CHNRI method introduced three key conceptual advances that led to its increased popularity in comparison to other priority–setting methods and processes. First, it proposed a systematic approach to listing a large number of possible research ideas, using the "4D" framework (description, delivery, development and discovery research) and a well–defined "depth" of proposed research ideas (research instruments, avenues, options and questions). Second, it proposed a systematic approach for discriminating between many proposed research ideas based on a well–defined context and criteria. The five "standard" components of the context are the population of interest, the disease burden of interest, geographic limits, time scale and the preferred style of investing with respect to risk. The five "standard" criteria proposed for prioritization between research ideas are answerability, effectiveness, deliverability, maximum potential for disease burden reduction and the effect on equity. However, both the context and the criteria can be flexibly changed to meet the specific needs of each priority–setting exercise. Third, it facilitated consensus development through measuring collective optimism on each component of each research idea among a larger group of experts using a simple scoring system. This enabled the use of the knowledge of many experts in the field, "visualising" their collective opinion and presenting the list of many research ideas with their ranks, based on an intuitive score that ranges between 0 and 100.

Conclusions Two recent reviews showed that the CHNRI method, an approach essentially based on "crowdsourcing", has become the dominant approach to setting health research priorities in the global biomedical literature over the past decade. With more than 50 published examples of implementation to date, it is now widely used in many international organisations for collective decision–making on health research priorities. The applications have been helpful in promoting better balance between investments in fundamental research, translation research and implementation research.

Correspondence to:
Professor Igor Rudan
Chair in International Health and Molecular Medicine
Joint Director
Centre for Global Health Research and WHO Collaborating Centre for Population Health Research and Training
Director of Research, The Usher Institute for Population Health Sciences and Informatics
The University of Edinburgh
Teviot Place
Edinburgh EH8 9AG
Scotland, UK
Igor.Rudan@ed.ac.uk

Child Health and Nutrition Research Initiative (CHNRI) started as an initiative of the Global Forum for Health Research in Geneva, Switzerland [1]. Its aim was to develop a method that could assist priority setting in health research investments [2]. The first version of the CHNRI method was published in 2007–2008 [3–6]. The aim of this paper was to summarize the history of the development of the CHNRI method and its key conceptual advances [7].

The history of the development of the CHNRI method

In 2005, CHNRI was funded by the World Bank to develop a method that could assist priority setting in health research investments. In March 2005, Professor Robert E. Black, from Johns Hopkins University in Baltimore, USA, Dr Shams El Arifeen, Director of the CHNRI Secretariat from the International Centre for Diarrheal Disease Research (ICCDR,B) in Dhaka, Bangladesh, and Nancy Hughart, Secretary of the CHNRI office met in Geneva and appointed me to lead the process of methodology development for CHNRI. Professors Jennifer Bryce and Robert E. Black from the Child Health Epidemiology Reference Group (CHERG) recommended me for this role based on my previous work and contributions to CHERG. In May 2005, at a meeting at Johns Hopkins University, I presented the first background review on different approaches to research priority setting and an early conceptual framework for the future CHNRI method. I received feedback from world–renowned experts in global health, such as Professors Dean Jamison, Ok Pannenborg and Mary Ann Lansang; and from priority–setting experts Jennifer Gibson, Lydia Kapiriri and Craig Mitton.

In June 2005, assisted by the new CHNRI secretary, Ms Deborah Horner, I invited a larger group of global health experts to Dubrovnik, Croatia, to help me develop the method further and plan its implementation in several fields of global health: newborn health (Joy E. Lawn and Zulfiqar A. Bhutta), childhood pneumonia and diarrhea (Harry Campbell and Claudio F. Lanata), child development (Maureen Black and Julie Meeks Gardner), childhood accidents (Shanthi Ameratunga and Adnan A. Hyder) and zinc (Kenneth H. Brown and Sonja Y. Hess). In September 2005, at the *9th Annual Meeting of the Global Forum for Health Research* in Mumbai, India, I presented the first draft version of the CHNRI method and an example of its application in the field of childhood pneumonia. I did this together with Dr Shams El Arifeen, Professor Robert E. Black and Professor Harry Campbell, who were mentoring and supporting me throughout the process of methods development. In December 2005, at the launch of the *Child Survival: Countdown to 2015* conference in London, UK, I presented the key concepts of the new CHNRI method at the plenary session on health research agenda for child survival. Following the feedback from the audience, I revised and improved the method.

In April 2006, I visited Cape Town to conduct the first national–level implementation of the CHNRI exercise – to set research priorities for child health in South Africa. I was supported by Dr Mickey Chopra and Dr Mark R. Tomlinson, from MRC's Health Systems Unit in Cape Town. At this point, the first exercises on childhood pneumonia and zinc were already being piloted at the global level, by Professor Harry Campbell from the University of Edinburgh, UK and Professor Kenneth Brown from the University of California in Davis, USA. At this point, I suggested that two more consultants should be contracted to assist me with preparations for publishing a series of four papers that would describe the CHNRI method: Drs Jennifer L. Gibson and Lydia Kapiriri from the University of Toronto.

In May 2006, the CHNRI Foundation organized a meeting at Johns Hopkins University. The meeting had a wider participation, aiming to include several representatives from donor agencies who could potentially be interested in the implementation of the method. The most recent version of the method and the examples of its implementation were presented and discussed in detail. In June 2006, following the meeting in Baltimore, Dr Jose Martines, the representative of the World Health Organization's Child and Adolescent Health Department (WHO CAH) arranged a meeting in Geneva, Switzerland, where he commissioned a series of 5 CHNRI exercises that would be co–ordinated by WHO CAH and focus on research priorities for five major causes of child deaths: childhood pneumonia, diarrhea, neonatal infections, preterm birth/low birth weight and birth asphyxia. Those exercises were going to be well aligned with UN's Millennium Development Goal 4 – a political commitment made by world's nations to reduce global child mortality by two thirds between 1990 and 2015. This was the first major uptake of the CHNRI method by an international organization.

Further refinements of the CHNRI method were introduced based on the feedback received following the presentations at the *International Conference on Priorities in Health Care* in Toronto, Canada, in September 2006 and at the *10th Annual Meeting of the Global Forum for Health Research* in Cairo, Egypt. A parallel session on priority setting in health research investments was organized by CHNRI at the latter conference, with outstanding secretarial support from Ms Carolina Cueva Schaumann, the CHNRI Secretary. The steering committee for the development of the CHNRI methodology approved the publication of the method, allowing for introduction of all the feedback received to date. I led the writing of a series of four papers that described the CHNRI method. I also presented the final revision of

the CHNRI method at the *11th Annual Meeting of the Global Forum for Health Research* in Beijing, China, in October 2007. The first two papers of the series that introduced the CHNRI method were published in parallel with the Beijing meeting, in October 2007 [3,4], with the remaining two following in June 2008 [5] and December 2008 [6]. In preparing the four papers, I received large help from Professors Robert E. Black, Shams El Arifeen and Harry Campbell, and further assistance from Drs Lydia Kapiriri, Jennifer Gibson, Mickey Chopra, Kit Yee Chan, Mary Ann Lansang, Ilona Carneiro, Shanthi Ameratunga, Alexander C. Tsai, Mark Tomlinson and Sonja Y. Hess.

An important recognition of the CHNRI method came with an invitation from the World Health Organization's Cluster on Information, Evidence and Research (IER), its Department for Research Policy and Cooperation (RPC) and the Special Programme for Research and Training in Tropical Diseases (TDR). Those WHO Clusters and Departments convened a workshop in April 2008 to review the available priority setting methodologies for health research. I presented the CHNRI method, which received endorsement for the uptake at the national level through the meeting's official report. Results of this meeting were later summarized and reported by Tomlinson et al. [7]. I also gave two plenary presentations on the CHNRI method at the *12th Annual Meeting of the Global Forum for Health Research* in Havana, Cuba, in November 2009 and *XIX World Congress of Epidemiology* in Edinburgh in August 2011 [8], where the method was presented to large international audiences and its uptake enhanced.

With a considerable uptake and more than 50 published examples of implementation to date, the CHNRI method is now widely used in many international organisations and professional societies for setting health research priorities. Two recent reviews showed that the CHNRI method has become the dominant approach to setting health research priorities in the global biomedical literature over the past decade [9,10]. Its applications have been helpful in promoting better balance between investments in fundamental research, translation research and implementation research.

Setting health research priorities: universal challenges and CHNRI's key conceptual advances

For anyone interested in setting health research priorities at any level, I recommend several comprehensive reviews of the principles, methods, approaches and tools [3,7–11]. Based on those readings, it should become apparent that the CHNRI method proposed its own definition of health research. In CHNRI method's conceptual framework, "health research" should be regarded as a process that begins with a research question and undertaken to generate new knowledge that will eventually be translated and/or implemented to reduce the existing disease burden (or other health–related problem) in the population [5].

Based on the above definition of health research, the group that developed the CHNRI method identified a considerable number of challenges that will be inherent to any process of setting health research priorities (**Table 1**). In attempts to address those challenges, the CHNRI method

Table 1. A list of twenty "universal challenges" in setting priorities in health research investments, according to the CHNRI method's conceptual framework [5]

1. Deciding who should be involved in the process of setting health research priorities
2. Defining what constitutes a health research investment option opportunity
3. Defining what constitutes the expected "*return*" on the investment
4. Defining what constitutes a potential "*risk*" of the investment
5. Defining health research, its boundaries, and its levels of "*depth*"
6. Systematic listing of a very large number of competing research investment options
7. Defining what is meant by "*priority setting*" in the context of health research
8. Finding a way to address the uncertainty of health research outcomes
9. Defining criteria relevant to priority setting in health research investments
10. Comparing different instruments of health research using the same criteria
11. Development of a simple quantitative way to rank competing research options
12. Limiting the potential of personal biases to substantially influence the outcome
13. Ensuring that priority–setting process is fully transparent
14. Ensuring that it can be repeated and validated
15. Ensuring that it is flexible and adjustable to all contexts and levels of application
16. Ensuring that it is iterative with a feedback loop, instead of a one–way process
17. Ensuring that it is perceived by the users as legitimate and fair
18. Ensuring that it is simple and intuitive, to become popular among the users
19. Linking quantitative ranks of research options with specific investment decisions
20. Involving stakeholders from the wider community into the process

introduced three key conceptual advances that led to its increased popularity in comparison to other priority–setting methods and processes.

First, it proposed a solution to the problem of addressing a potentially endless spectrum of research ideas. It proposed a systematic approach to listing a large number of feasible research ideas. To this end, it uses the "*4D framework*" ("description", "delivery", "development" and "discovery" research). "*Description*" research includes any proposed health research that would allow researchers to assess the burden of health problems in the population of interest and understand its determinants – ie, negative effects of risk factors and positive effects of delivered health interventions. This is typically achieved through epidemiological research. "*Delivery*" research includes all research questions that allow researchers to optimise health status of the population using the means that are already available. This is typically achieved through implementation research, operations research and/or health policy and systems research. "*Development*" research is focused on improving health interventions that already exist, but could be made more effective, affordable or sustainable. Finally, "*discovery*" research includes all research questions that would lead to innovation, ie, generation of new knowledge to develop entirely new health interventions.

Within each of those four main "*instruments*" of health research – the four D's – research questions of different "depth" could be posed: very broad "*research avenues*" (which correspond to research fields), more specific "*research options*" (which correspond to a typical research program of about 5 years in duration), and very specific "*research questions*" (which correspond to a title of a typical research paper). Based on this framework, a very large number of proposed research ideas can be systematically assembled and prepared for prioritization (**Table 2**).

The second key conceptual advance was defining the context and criteria for prioritization among many research ideas based on a sound framework. The five "*standard*" components of the context in which priority–setting is taking place are the population of interest, the disease burden of interest, geographic limits, time scale and the preferred style of investing with respect to risk (**Table 3**). Depending on who the funders are – government, private sector (eg, pharmaceutical industry and/or biotechnological industry), or philanthropic foundations – their choices of the target population and the health problem of interest, geographic limits, time scale and attitude to risk may be very different. Thus, the elements of the context need to be carefully defined and transparently communicated to scorers before the CHNRI prioritization exercise takes place.

Once the context was carefully defined according to **Table 3**, and many competing research ideas systematically categorised using the "*4D framework*" in **Table 2**, the next challenge was finding an optimal set of criteria that could distinguish and discriminate between the proposed research ideas, expose their key strengths and weaknesses

Table 2. Child Health and Nutrition Research Initiative's (CHNRI) proposed framework for systematic listing of research ideas in health research, which takes into account the "instruments" of health research (rows) and the "depth" of proposed research ideas (columns)

Research instrument	Research avenue	Research option	Research question
"Description": research to assess the burden of health problem (disease) and its determinants, ie, negative effects of risk factors and positive effects of delivered health interventions	• Measuring the burden • Understanding risk factors (in terms of their relative risks) • Measuring prevalence of exposure to risk factors • Evaluating the efficacy and effectiveness of interventions in place • Measuring prevalence of coverage of interventions in place	Many research options within each of the avenues; research options should correspond to a research program of up to 5 years in duration	Specific research questions within each of the research avenues should correspond to the title of individual research papers
"Delivery": research to assist in optimising of the health status of the population using the means that are already available	• Health policy analysis • Health system structure analysis • Financing/costs analysis • Human resources • Provision/infrastructure • Operations research • Responsiveness/recipients		
"Development": research to improve health interventions that already exist, but could be improved	• Improving existing interventions (their affordability, deliverability, sustainability, acceptability, etc.)		
"Discovery": research that leads to innovation, ie, entirely new health interventions	• Basic, clinical, and public health research to advance existing knowledge to develop new capacities • Basic, clinical, and public health research to explore entirely novel ideas to develop new capacities		

Table 3. Elements of the context in which health research prioritization takes place; they need to be clearly defined and communicated to invited technical experts prior to listing and scoring health research ideas

(i) **Population of interest** This element of the context defines the main groups in the society whose health problems are being addressed through health research priority setting.
(ii) **Disease, disability, and death burden** This element of the context defines what is known about the burden of disease, disability, and death that will be addressed by supported health research – e.g., can it be measured and quantified (in disability–adjusted life years–DALYs – or in some other way).
(iii) **Geographic limits** This element of the context defines boundaries in terms of space, which may be global, regional, national, sub–national, etc.
(iv) **Time scale** This element of the context defines the level of urgency, ie, in how many years are the first results of the proposed research expected (they may be defined as reaching the endpoints of the research process, or translating and implementing them, or achieving detectable disease burden reduction).
(v) **Preferred style of investing** This element of the context defines investment strategy in health research with respect to risk preferences; it defines whether most of the funding would support a single (or a few) expensive high–risk research ideas (eg, vaccine development), or will the risk be balanced and diversified between many research options which will have different levels of risk and feasibility.

and assign them an overall "value" according to which they could all be ranked and compared between each other. The chosen set of criteria should be aligned with the guiding principle of the CHNRI method – to expose the potential of many competing health research ideas to reduce disease burden and inequities that exist in the population in a feasible and cost–effective way.

Table 4 shows a larger number of the possible criteria that could be used to discriminate between the values of any two (or more) competing research ideas. Using such a large number of criteria is clearly impractical, and many of them overlap to a degree and capture similar information about the proposed research idea. Based on CHNRI's definition of the process of health research, as described earlier in the text, and the likelihood of this process to progress from one stage to another, the five "standard" criteria proposed by the CHNRI method for prioritization between research ideas are: (i) answerability, (ii) effectiveness, (iii) deliverability, (iv) maximum potential for disease burden reduction and (v) the effect on equity.

However, an advantage of the CHNRI method is that both the elements of the context and the number and the composition of the criteria can be flexibly changed to meet the specific needs of each priority–setting exercise. Further elements may be added to the context description, or some of the proposed ones can be dropped or replaced. The same is true for the priority–setting criteria, and I encourage the users of the method to take advantage of this flexibility to meet the goals of their specific exercise. I believe that the CHNRI method owes its uptake and implementation in a

Table 4. Some of the possible priority–setting criteria (and related questions) proposed by Child Health and Nutrition Research Initiative (CHNRI) that can be used to discriminate between any two (or more) health research ideas to set research priorities; the outcomes of the application of different criteria will necessarily conflict each other

Answerability? (some health research ideas will be more likely to be answerable than the others)
Attractiveness? (some health research ideas will be more likely to lead to publications in high–impact journals)
Novelty? (some health research ideas will be more likely to generate truly novel and non–existing knowledge)
Potential for translation? (some health research ideas will be more likely to generate knowledge that will be translated into health intervention)
Effectiveness? (some health research ideas will be more likely to generate/improve truly effective health interventions)
Affordability? (the translation or implementation of knowledge generated through some health research ideas will not be affordable within the context)
Deliverability? (some health research ideas will lead to / impact health interventions that will not be deliverable within the context)
Sustainability? (some health research ideas will lead to / impact health interventions that will not be sustainable within the context)
Public opinion? (some health research ideas will seem more justified and acceptable to general public than the others)
Ethical aspects? (some health research ideas will be more likely to raise ethical concerns than the others)
Maximum potential impact on the burden? (some health research ideas will have a theoretical potential to reduce much larger portions of the existing disease burden than the others)
Equity? (some health research ideas will lead to health interventions that will only be accessible to the privileged in the society/context, thus increasing inequity)
Community involvement? (some health research ideas will have more additional positive side–effects through community involvement)
Feasibility? (some health research ideas will be unlikely to lead to translation at the current stage of knowledge)
Relevance? (some health research ideas will be more relevant to the context than the others)
Fills key gap? (some health research ideas will be more likely to fill the key gap in knowledge that is required for translation and/or implementation than the others)
Cost? (some research ideas will require more funding than the others)
Fundability? (some research ideas will be more likely to receive funding support within the defined context than the others)
Alignment with political priorities? (some research ideas will be more likely to be aligned with contemporary political priorities than the others)
Likelihood of generating patents/lucrative products? (some research ideas will have greater likelihood of generating patents or other potentially lucrative products, thus promising greater financial return on investments, regardless of their impact on disease burden)

large part to its flexibility, as it can be readily tailored to many different contexts and purposes.

For example, in different contexts addressing of the "*answerability*" criterion may also require a separate assessment related to the ethics of the proposed research idea, an evaluation of the existing research capacity, or an assessment of the likely public acceptance of research results. The "*relevance*" criterion may need to be further refined into criteria that would separately assess effectiveness, deliverability, affordability, sustainability, and whether a critical gap in knowledge is being addressed. The "*maximum potential impact on the burden*" will occasionally not only assess the quantity of potential burden reduction, but also its quality – ie, whether this reduction is targeting those most heavily affected or underprivileged in the population [5]. **Table 4** lists a number of possible criteria that can be used for setting priorities between different research investment options and questions about each option that could address these criteria well.

The third key conceptual advance of the CHNRI method relates to the problem of consensus development and agreement on the priorities among many proposed research ideas. Before the introduction of the CHNRI method, a typical consensus development process would involve the so–called Delphi method [12,13]. This process would typically require background reading, followed by the first round of discussions among relatively small groups of experts. Expert interactions and the opportunities for the experts to influence one another defined the process of consensus development. There would usually be a step where a feedback would be provided to experts, which would further influence their independent opinion, followed by the second round of discussions. Eventually, the groups would reach a consensus on research priorities. The problem with this process was that it could not be considered transparent, replicable or democratic, because at each stage there was a large opportunity for the managers of the process, or individual participants with strong opinions, to influence all other participants.

At the time of the development of the CHNRI method, the rise of information technologies and online communication enabled the new approach to developing consensus among a larger group of people through so–called "crowdsourcing". It was proposed that simply reaching out to a large number of people and assessing their collective opinion (in this case, optimism toward a large number of research ideas to fulfil the specific priority–setting criteria) may result in surprisingly accurate predictions that would typically surpass any individual's expert judgement [14]. However, there were several requirements that needed to be met to ensure that the collective opinion would indeed be useful. Those included diversity of opinion (meaning that each participant should have his/her private information), independence of participants (meaning that participants' input wouldn't be influenced by the opinion of other participants), decentralization (meaning that participants would be diverse and able to draw on any local knowledge) and aggregation (meaning that a mechanism would be available for collecting many individual opinions and turning them into a collective opinion). The rise of information technologies–based communication allowed to collect information from a large number of international experts in global health quickly and efficiently, with all the above requirements met.

Thanks to this advance, the CHNRI method proposed a radically different approach to consensus development from the Delphi process. In both methods input from experts is required, and the invited experts have the same background characteristics. However, the CHNRI method collects opinion from many international experts through their e–mail input, no background reading is required, and no discussions or interactions would occur between many participating experts. Feedback on their collective opinion could still be returned to participants, but there would not be a need for a step where consensus would need to be developed, because a simple quantitative analysis of the received input would turn the information obtained from each expert into a "collective" result, which would belong to every participant, but no single participant would have a chance of influencing any substantial portion of it. Then, the areas of greater or smaller consensus could be identified through agreement statistics analysis of the input, without a need for a second round of discussion.

Thus, the CHNRI method innovated the process of consensus development through measuring collective optimism of a larger number of international experts on each research idea and each criterion. This was done through consulting a larger group of experts and using a very simple scoring system, where they only needed to say whether they thought that the research idea was likely, or not, to meet the priority–setting criterion within the specified context. This enabled the use of the knowledge of many experts in the field, "visualising" their collective opinion and presenting the list of many research ideas with their ranks, based on an intuitive score that ranges between 0% (absolutely no optimism) and 100% (where everyone is optimistic). In this way, the knowledge of a larger number of international experts is used, through "crowdsourcing", to discriminate between competing research options based on strictly defined criteria and the collective optimism toward compliance of each research option with each criterion. Such approach limits the potential of individual personal biases to substantially influence the outcome, which was identified as a major challenge that needed to be addressed.

The proposed conceptual advance in the CHNRI method ensured that the scoring experts provided their input independently of each other, and that the final scores for each competing research option were obtained and computed in a highly structured, transparent, systematic and replicable way. Through application of agreement statistics methods, the CHNRI method could then also identify and expose controversial issues (ie, responses with a large variation in scores among experts). Finally, the proposed method promised to generate a large amount of useful information for funders of research and research communities alike, by "visualising" the collective opinion of many leading experts on many research ideas and their key components.

The above characteristics of the CHNRI method also dealt with several other universal challenges. The flexibility in the choice of context and criteria ensured that the method would be adjustable to all contexts and areas of application [5]. It also envisioned a "feedback loop", because the process of priority setting could be repeated after certain periods of time, allowing the priorities to change with the changing context. Transparency and clarity of the proposed steps of the CHNRI method were intended to ensure that it is perceived as legitimate and fair by its users [15,16].

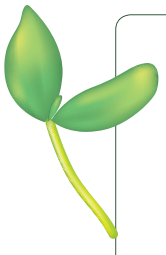
CONCLUSIONS

The CHNRI method measures collective optimism of a larger number of researchers toward various components of many proposed research ideas, within an agreed context and using the agreed criteria. Because of this process, a large number of health research ideas would receive their intermediate and overall "priority scores", which will be in a quantitative form, ranging from 0 to100%. This should provide a large amount of useful information to many funders, researchers and stakeholders alike. Advantages and disadvantages of each research idea should become transparent through this process, which would be based on a "democratic" assessment [5].

A common misconception in the early days of the CHNRI method development and implementation was that the CHNRI process would be telling the funders where to invest their resources. However, this is not what the CHNRI method does to any extent. It is designed to merely present a very large amount of information to the funders on many research ideas, including their strengths and weaknesses. In a way, this is not much different from the information available on the performance of various companies that are potential investment options in the stockmarket. The CHNRI process should simply allow the funders of health research to choose from many research ideas based on a lot of information that the expert group provided on each idea. This should protect funders from risky investments and allow them to develop their own investment style and portfolio [5].

There are further practical advantages of the CHNRI exercise, such as the ease of conducting the exercise over the internet, low cost of planning and conducting the exercise, ease of obtaining information from many experts online, and excellent prospects of publishing the results of each exercise. Modifications of the CHNRI method should also allow prioritization among investments in health care, emerging health technologies, and development assistance for health. The CHNRI method should be of possible use to research funding bodies, international organizations and for–profit companies in setting their own strategic priorities among many different ideas, based on collective knowledge of their most qualified employees or external experts [5].

The CHNRI method is not free from shortcomings and possible concerns over the validity of the process. I will address the most important among those concerns in the following papers of this series through a set of carefully designed experiments into quantitative properties of human collective knowledge and opinion. Those studies should bring more certainty over the components of the CHNRI process that are critical for its validity. This will be followed by definite guidelines for implementation, based on a review of more than 50 exercised conducted to date and their impact on research policy.



Acknowledgments: None.

Funding: This work was supported by the CHNRI Foundation award to the University of Edinburgh.

Ethical approval: Not required.

Authorship declaration: IR developed the idea for the paper and wrote the paper.

Declaration of interest: IR is an editor–in–chief of the Journal of Global Health. To ensure that any possible conflict of interest relevant to the journal has been addressed, this article was reviewed according to best practice guidelines of international editorial organizations. The author completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). The author declares no conflicting financial or other interest related to the work detailed in this manuscript.

REFERENCES

1 Black RE. The legacy of the Child Health and Nutrition Research Initiative (CHNRI). J Glob Health. 2016;6:010101. Medline:26955468

2 Rudan I, El Arifeen S, Black RE. A systematic methodology for setting priorities in child health research investments. Child Health and Nutrition Research Initiative (CHNRI), Dhaka, 2006

3 Rudan I, Gibson J, Kapiriri L, Lansang MA, Hyder AA, Lawn JE, et al. Child Health and Nutrition Research Initiative (CHNRI). Setting priorities in global child health research investments: assessment of principles and practice. Croat Med J. 2007;48:595-604. Medline:17948946

4 Kapiriri L, Tomlinson M, Gibson J, Chopra M, El Arifeen S, Black RE, et al. Child Health and Nutrition Research Initiative (CHNRI): setting priorities in global child health research investments: addressing the values of the stakeholders. Croat Med J. 2007;48:618-27. Medline:17948948

5 Rudan I, Chopra M, Kapiriri L, Gibson J, Lansang MA, Carneiro I, et al. Setting priorities in global child health research investments: universal challenges and conceptual framework. Croat Med J. 2008;49:307-17. Medline:18581609 doi:10.3325/cmj.2008.3.307

6 Rudan I, Gibson JL, Ameratunga S, El Arifeen S, Bhutta ZA, Black M, et al; Child Health and Nutrition Research Initiative (CHNRI). Setting priorities in global child health research investments: guidelines for implementation of the CHNRI Method. Croat Med J. 2008;49:720-33. Medline:19090596 doi:10.3325/cmj.2008.49.720

7 Tomlinson M, Chopra M, Hoosain N, Rudan I. A review of selected research priority setting processes at national level in low and middle income countries: towards fair and legitimate priority setting. Health Res Policy Syst. 2011;9:19. Medline:21575144 doi:10.1186/1478-4505-9-19

8 Rudan I. Global health research priorities: mobilizing the developing world. Public Health. 2012;126:237-40. Medline:22325672 doi:10.1016/j.puhe.2011.12.001

9 McGregor S, Henderson KJ, Kaldor JM. How are health research priorities set in low and middle income countries? A systematic review of published reports. PLoS ONE. 2014;9:e108787. Medline:25275315 doi:10.1371/journal.pone.0108787

10 Yoshida S. Approaches, tools and methods used for setting priorities in health research in the 21st century. J Glob Health. 2016;6:010507. Medline:26401271

11 Viergever RF, Olifson S, Ghaffar A, Terry RF. A checklist for health research priority setting: nine common themes of good practice. Health Res Policy Syst. 2010;8:36. Medline:21159163

12 Beretta R. A critical view of the Delphi technique. Nurse Res. 1996;3:79-89. Medline:27285530 doi:10.7748/nr.3.4.79.s8

13 Keeney S, Hasson F, McKenna HP. A critical review of the Delphi technique as a research methodology for nursing. Int J Nurs Stud. 2001;38:195-200. Medline:11223060 doi:10.1016/S0020-7489(00)00044-4

14 Surowiecki J. The wisdom of crowds. London: Abacus, 2013.

15 Gibson JL, Martin DK, Singer PA. Setting priorities in health care organizations: criteria, processes, and parameters of success. BMC Health Serv Res. 2004;4:25. Medline:15355544 doi:10.1186/1472-6963-4-25

16 Daniels N, Sabin JE. Setting limits fairly: can we learn to share medical resources? Oxford: Oxford University Press, 2002.

journal of global health

PAPERS

Setting health research priorities using the CHNRI method: V. Quantitative properties of human collective knowledge

Igor Rudan[1], Sachiyo Yoshida[2], Kerri Wazny[1], Kit Yee Chan[1,3], Simon Cousens[4]

[1] Centre for Global Health Research, The Usher Institute for Population Health Sciences and Informatics, the University of Edinburgh, Edinburgh, Scotland, UK

[2] Department for Maternal, Newborn, Child and Adolescent Health, World Health Organization, Geneva, Switzerland

[3] Nossal Institute for Global Health, University of Melbourne, Melbourne, Victoria, Australia

[4] Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK

Introduction The CHNRI method for setting health research priorities has crowdsourcing as the major component. It uses the collective opinion of a group of experts to generate, assess and prioritize between many competing health research ideas. It is difficult to compare the accuracy of human individual and collective opinions in predicting uncertain future outcomes before the outcomes are known. However, this limitation does not apply to existing knowledge, which is an important component underlying opinion. In this paper, we report several experiments to explore the quantitative properties of human collective knowledge and discuss their relevance to the CHNRI method.

Methods We conducted a series of experiments in groups of about 160 (range: 122–175) undergraduate Year 2 medical students to compare their collective knowledge to their individual knowledge. We asked them to answer 10 questions on each of the following: (i) an area in which they have a degree of expertise (undergraduate Year 1 medical curriculum); (ii) an area in which they likely have some knowledge (general knowledge); and (iii) an area in which they are not expected to have any knowledge (astronomy). We also presented them with 20 pairs of well–known celebrities and asked them to identify the older person of the pair. In all these experiments our goal was to examine how the collective answer compares to the distribution of students' individual answers.

Results When answering the questions in their own area of expertise, the collective answer (the median) was in the top 20.83% of the most accurate individual responses; in general knowledge, it was in the top 11.93%; and in an area with no expertise, the group answer was in the top 7.02%. However, the collective answer based on mean values fared much worse, ranging from top 75.60% to top 95.91%. Also, when confronted with guessing the older of the two celebrities, the collective response was correct in 18/20 cases (90%), while the 8 most successful individuals among the students had 19/20 correct answers (95%). However, when the system in which the students who were not sure of the correct answer were allowed to either choose an award of half of the point in all such instances, or withdraw from responding, in order to improve the score of the collective, the collective was correct in 19/20 cases (95%), while the 3 most successful individuals were correct in 17/20 cases (85%).

Conclusions Our experiments showed that the collective knowledge of a group with expertise in the subject should always be very close to the true value. In most cases and under most assumption, the collective knowledge will be more accurate than the knowledge of an "average" individual, but there always seems to be a small group of individuals who manage to out–perform the collective. The accuracy of collective prediction may be enhanced by allowing the individuals with low confidence in their answer to withdraw from answering.

Correspondence to:
Professor Igor Rudan
Centre for Global Health Research
The Usher Institute for Population Health Sciences and Informatics
The University of Edinburgh
Teviot Place
Edinburgh EH8 9AG
Scotland, UK
igor.rudan@ed.ac.uk

In 1906, Galton suggested that a group of individuals make better predictions as a collective than any individual expert [1]. Since then, our understanding of the "Wisdom of Crowds" has grown: in recent years, a widely appreciated example of this phenomenon has been evident to the audience of the quiz show "Who Wants To Be A Millionaire?" In this quiz show, a contestant needs to answer a series of increasingly difficult questions by picking from one of four possible responses, only one of which is correct – so that the probability that a random response is correct is 25%. In this show, an "*Ask the audience*" joker is available, whereby 100 persons in studio audience get to submit electronically their opinion on what the correct answer is, and the distribution of their individual opinions is then shown to the contestant. As an alternative, a "*Phone a friend*" joker allows contestants to phone one friend whom they consider the most knowledgeable, and then ask for his/her individual answer. Comparative analyses of the performance of the two jokers showed that the relative majority of the audience chose the correct answer about 91% of the time, while the most knowledgeable friend was right about 65% of the time. There are methodological concerns over the direct comparison between these two percentages, because these success rates were based on different questions, but the difference is still quite striking [1].

Crowdsourcing has become an increasingly popular human tool to address many problems–from government elections in democracies [2], formation of stock market prices [3], to modern online platforms such as TripAdvisor (to advise on the best hotels and restaurants) [4] or Internet Movie Database (IMDb) (to advise on the best movies, TV shows, etc.), all of which are based on the personal opinions of many hundreds or thousands of participants [5]. When crowdsourcing is used for gathering information, or in decision–making processes, there is probably a need to distinguish between at least three different scenarios in which collective knowledge might be used. The first is getting the right answer to a factual question, which we may consider "objective knowledge" and it represents the simplest case. The second is predicting the outcome of some future event, which can subsequently be verified with certainty and within a reasonable time frame. An example is betting on an outcome, eg, of football games or horse races. This is different from stock market predictions, where those who participate in predictions (investors) can also influence the outcomes through their actions. Finally, crowdsourcing could be used to gather information on subjective opinion on something that cannot be easily verified. This last scenario is the closest to how crowdsourcing is used in the CHNRI method (the acronym for: Child Health and Nutrition Research Initiative) [6,7], which seeks to gauge collective optimism with respect to different health research ideas and the benefits they might lead to at some point in the future.

The CHNRI method for setting health research priorities uses "crowds" of experts in global health – researchers, policy makers and programme implementers – to generate, assess and prioritize between many competing ideas in global health research. A CHNRI exercise produces a ranking of many research ideas according to the collective opinion of the expert group, but it is not possible to verify objectively how "valid" that ranking may be, not least because low ranked ideas are unlikely to be funded and therefore no outcomes are available for them. It is yet to be demonstrated that the collective opinion of an expert group should be regarded as more useful than the opinion of individual experts in the group [1,8]. However, the difficulties related to validating personal opinions do not apply to the validation of personal knowledge, and the accuracy of personal knowledge is an important component underlying the individual's opinion. Because of this, we should expect some parallels between the quantitative properties of human collective knowledge and human collective opinion. In this paper, we report several experiments to explore the quantitative properties of human collective knowledge and discuss their possible relevance to the validity of the CHNRI method. The aim of this paper is to examine the accuracy of collective compared to individual knowledge, using different approaches of assessment.

METHODS

We conducted a series of experiments among a group of undergraduate medical students. The number of participating students ranged from 122 to 175 in each exercise. Students who completed the second year lectures in Epidemiology and Statistics, as part of a practical application of epidemiological and statistical concepts were asked to answer 10 questions on each of the following: (i) an area in which they have a degree of expertise (subjects related to the medical curriculum for the first year undergraduate); (ii) an area in which they have some knowledge but do not have expertise (general knowledge); and (iii) an area in which they are not expected to have any knowledge (astronomy). The content of the lecture was entirely unrelated to the questions that were asked from the students. The ethics approval was obtained from a relevant research centre (Centre for Population Health Sciences at the University of Edinburgh).

The questions were chosen so that the answer to each question was numerical (an integer), and so that the answers ranged from a 1–digit number to a 10–digit number over the course of 10 questions in random order, with students unaware of this element of the design. This element was included to allow us to assess whether the students' answers were more accurate when the correct answer was a smaller or larger number (see **Online Supplementary Document**).

Table 1 shows the questions that were asked in each of the three areas, and the correct answers. The questions were asked at the end of 3 consecutive lectures spanning 10 days. Students were given 30 seconds to answer each question. The students were asked to record an answer for every question. For questions for which they were unsure of the answer they were asked to write down their best guess.

In addition, students were shown 20 pairs of well–known celebrities and asked them identify which was the older of the two. **Table 2** shows the pairs of celebrities in the order that the questions were asked. The questions were phrased as: "*Would you say that Celebrity X is older than Celebrity Y?*", and the possible answers were either "*Yes*" or "*No*", where they had to choose one of those two options. However, they were also given an option next to each answer to choose their "secondary" answer as either "*Not sure*" (when they were familiar of both celebrities, but it was too difficult to judge), or leaving the answer "*Blank*" deliberately, when not knowing one or both celebrities. Those two options would indicate their low confidence in their "Yes"/"No" answer. By adding "*Not sure*" (which would be coded with half a point) or "*Blank*" (which would remove them from the sample, leaving the others with more confidence in their answers), they could prevent a wrong answer and increase the chance of the collective answer to be close to the correct answer. This latter type of "scoring" is also used by the CHNRI method. In this way, the same group of students provided two different data sets with scores: one, where they all needed to provide a binary ("Yes"/"No") answer to each question, regardless of their confidence in answering the question correctly; and the other one, where they were able to use the answer "*Not sure*", or leave the answer blank, when they were not confident in their answer. Their input was then turned into a data sheet that was analogous to those produced in the CHNRI exercise, where "*Yes*" was

Table 1. Questions posed to a group of undergraduate Year 2 medical students*

Questions in an area of students' high expertise (undergraduate Year 1 medical curriculum)	
1. How many valence electrons does carbon have?	(4)
2. How many pairs of cranial nerves are there?	(12)
3. How many bones in the adult human body?	(206)
4. In which year did Freud publish "The interpretation of dreams"?	(1900)
5. How many genes does a human have?	(23 000)
6. What is an average salary of a GP in the UK?	(104 000)
7. How many erythrocytes in 1 mL of blood?	(5 000 000)
8. How many refugees are there in the world?	(15 400 000)
9. How many people in the world have diabetes?	(347 000 000)
10. How many bases (A, T, C or G letters) are in the haploid human genome?	(3 000 000 000)
Questions in an area of students' moderate expertise (general knowledge)	
1. How many marriages did Elizabeth Taylor have?	(8)
2. How old was Mozart when he died?	(35)
3. How many minutes does the movie "Casablanca" last?	(102)
4. In which year was Hamlet first published?	(1603)
5. How many diseases in ICD–10?	(14 400)
6. What is the average house price in the UK (in GBP)?	(238 976)
7. How many people live in Cape Town?	(3 740 000)
8. How much was Van Gogh's "sunflowers" painting sold for (in US$)?	(39 700 000)
9. What is the population size of Indonesia?	(246 900 000)
10. How many views did Psy's "Gangham Style" video have to date?	(1 764 039 000)
Questions in an area of student's low expertise (astronomy)	
1. How many light years from our Sun is Sirius?	(9)
2. How many moons does Saturn have?	(62)
3. How many times is Jupiter heavier than Earth?	(318)
4. In which year was Uranus first discovered?	(1781)
5. Distance between our Sun and the centre of Milky Way galaxy (in light–years)?	(27 000)
6. How many times is the Sun heavier than Earth?	(332 900)
7. What is the speed of the solar wind (in Km/h)?	(1 440 000)
8. How many years ago did the comet impact killed off dinosaurs?	(65 000 000)
9. Distance between the Sun and the Jupiter (in kilometres)?	(780 000 000)
10. How many years ago was our Solar System formed?	(4 568 000 000)

*The group was about 170 (range: 167–175) undergraduate Year 2 medical students from: (i) an area of their high expertise (ie, undergraduate Year 1 medical curriculum); (ii) an area where they have some expertise (general knowledge); and (iii) an area where they should have no expertise (astronomy). Correct answers are shown in brackets.

Table 2. Questions posed to a group of 122 undergraduate medical students to guess which well–known celebrity is older than the other*

Pair 1: Justin Bieber vs Miley Cyrus (19 vs 20)
Pair 2: George Clooney vs Brad Pitt (52 vs 49)
Pair 3: Madonna vs Susan Boyle (55 vs 52)
Pair 4: Beyonce vs Shakira (32 vs 36)
Pair 5: Dustin Hoffman vs Robert de Niro (76 vs 70)
Pair 6: Katy Perry vs Rihanna (28 vs 25)
Pair 7: Mick Jagger vs Paul McCartney (70 vs 71)
Pair 8: Lewis Hamilton vs Tiger Woods (28 vs 37)
Pair 9: Angela Merkel vs J. K. Rowling (59 vs 48)
Pair 10: Tony Blair vs George W. Bush (60 vs 67)
Pair 11: David Cameron vs Barack Obama (47 vs 52)
Pair 12: Ashton Kutcher vs Ben Affleck (35 vs 41)
Pair 13: Tom Cruise vs Nicole Kidman (51 vs 46)
Pair 14: Paris Hilton vs Jennifer Anniston (32 vs 44)
Pair 15: Jennifer Lopez vs Britney Spears (44 vs 31)
Pair 16: Eminem vs Jay–Z (40 vs 43)
Pair 17: Kim Kardashian vs Adele (33 vs 25)
Pair 18: Roger Federer vs Andy Murray (32 vs 26)
Pair 19: David Beckham vs Prince Harry (38 vs 29)
Pair 20: Elvis Presley vs Michael Jackson (42 vs 50)

*Correct answers (expressed in years of their age at the time of this exercise) are shown in brackets. The indicated age of individuals is relevant to October 17, 2013. For the last pair, the age at the time of death was being compared. The question was posed as: "Would you say that celebrity X is older than celebrity Y?" and possible answers were "Yes", "No", "Not sure" or "Blank" (see details in the text).

coded as "1", "*No*" as "0", "*Not sure*" as "0.5" and "*Blank*" responses were simply left as blank cells in the data sheet.

This design was carefully developed to allow us to study two questions: (i) how the students' collective opinion performs in comparison to that of individuals when the answers are no longer in a quantitative, but rather in a categorical format; and (ii) whether the *format* of categorical answer (with or without allowing for "*Not sure*" when students' confidence in their answer is low, or "*Blank*" when they simply don't have any knowledge on the question) altered the performance of the students' collective answer. Our hypothesis was that allowing students to answer "*Not sure*" or "*Blank*" would give better results, because it allows the participants within a team who are not sure of the correct answer to "withdraw" from providing their (possibly inaccurate) input, which would give more weight to the responses from students who were more confident in their individual knowledge.

Thus, four different experiments were conducted over the course of four consecutive lectures, which we label "Medical knowledge–quantitative" (MKQ), "General knowledge–quantitative" (GKQ), "Astronomy knowledge–quantitative" (AKQ) and "Celebrity knowledge–categorical" (CKC). In the MKQ, GKQ and AKQ exercises, we conducted the analyses in the following way: (i) we determined the median and the mean response for each of the 10 questions, based on all answers collected from the students (sample sizes were N = 167, N = 175 and N = 170, respectively); (ii) we also developed a parameter that we called "error size", to quantify the extent to which each student deviated from the correct answers over a series of 10 questions, and then we also applied it to the collective median and mean. Given that the responses could both over– or under–estimate the true value, we were interested in the ratio between the larger and the smaller of the two (ie, the correct answer and the answer provided by the student). As an example, this means that, if the correct answer was "10", and one student provided the answer "2" and the other "50", they would be making errors of the "same size": in our evaluation, it was equally wrong to over– or underestimate some value 5–fold. This also means that if the correct answer was provided for each question, then all the ratios contributing to "error size" parameter would be "1". Any deviation from the correct answer in either direction would increase the parameter from this theoretical minimum. (Note that this differs from other possible approaches, such a proportionally expressed increase or decrease, because the latter system would favour under–estimation as a smaller error than over–estimation, and under–estimation would be limited to 100% while overestimation would not be limited in any way). Once the individual errors, expressed as the ratio of the greater vs the smaller of the two values, was determined for each answer to each question, they were summarized for each individual student across all 10 questions and their sum was called "error size". In this way, each student was assigned his/her own "error size" in each of the three exercises (GKQ, MKQ and AKQ), and the students were then ranked by the error size parameter, from the smallest to the largest error made. This was then repeated for the entry of a collective (both using medians and means), and median and mean value rank within the entire student sample was then determined.

In the fourth exercise (CKC), which we designed as a series of 20 "*Yes* or *No*" questions, the task for the students was changed. In the first instance, the collective answer was taken to be the answer given by the majority of students–either "*Yes*" or "*No*". Then, there was an additional methodological caveat. First, those who were not confident about their answer could change some of their answers into the "*Not sure*" option, the effect of which contributed a certain 0.5 points to a total score, and minimised the risk of dropping a whole point for the collective for an incorrect answer. Second, those who had no knowledge of the question (eg, not recognising the names of celebrities) were allowed to change some of their responses to "*Blank*". This would have the effect of reducing the sample size of the collective, leaving all those with no knowledge out, and reducing the overall threshold of correct answers required from other students that the collective would need to answer correctly. Clearly, for those who are confident of their knowledge, this system would mean that they should answer "*Yes*" or "*No*" to all questions and not use either "*Not sure*" or "*Blank*" options at all.

The *correct* answer was then coded as "1", "*not sure*" as "0.5", the *incorrect* answer as "0", and "*blanks*" were excluded from the analysis, thus reducing sample size. The points assigned as described above were added ("1" for correct, "0.5" for "*not sure*", and "0" for incorrect) and then divided by the total number of "*non–blank*" responses received. The result was expressed as "*the percentage of correctness*" of the collective answer, and any value greater than 50% was considered a correct collective answer. This produced two data sheets–CKC1 (where everyone was required to submit either a *Yes* or a *No* answer) and CKC2 (with a *Yes–No–Not sure– Don't know scoring system*). The comparison between the two exercises was expected to reveal if "self–removal" through the use of "*Not sure*" or "*Blank*" improves the score of the collective considerably.

RESULTS

Students' collective answers (median and mean) to the 10 questions in three areas: (i) an area of their expertise, ie, Year 1 medical curriculum; (ii) the area of general knowledge; and (iii) the area outside of their expertise, ie, astronomy are shown in **Tables 3 to 5** (a total of 167, 175 and

Table 3. Year 2 undergraduate medical students' collective answers to the 10 questions in the area of their knowledge*

Question	Correct answer	Students' collective answer–median	Students' collective answer–mean
1. Valence electrons in carbon?	4	4	6
2. Number of cranial nerve pairs?	12	12	13
3. Number of bones in human body?	206	206	210
4. Freud's "Interpretation of dreams" published?	1900	1901	1890
5. Number of human genes?	23 000	38 000†	1 124 128 437
6. Average GP's salary in the UK?	104 100	76 001	85 568
7. Erythrocytes in 1 mL of blood?	5 000 000	8 679	12 124 582
8. Number of refugees in the world?	15 400 000	80 000 000	394 267 469
9. Number of people with diabetes?	347 000 000	100 000 000	444 785 232
10. Number of ATCGs in human genome?	3 000 000 000	23 500 327	178 090 845 668

*Number of responses N = 167.

†Question 5 was problematic because the number of human genes was revised down from about 40 000 to 23 000 only recently, ie, after the students learned of the former number; therefore, the median response from students was, in fact, very close to what they were likely to have learnt earlier in the course of their education).

Table 4. Year 2 undergraduate medical students' collective answers to the 10 questions in the area of general knowledge*

Question	Correct answer	Students'collective answer (median)	Students' collective answer–mean
1. Number of marriages of Elizabeth Taylor?	8	4	4
2. How old was Mozart when he died?	35	38	40
3. Minutes duration of "Casablanca"?	102	120	122
4. Year when "Hamlet" was published?	1603	1642	1637
5. Number of diseases in ICD–10?	14 400	48 132	76 480 054
6. Average house price in the UK?	238 976	193 271	369 819
7. Population size of Cape Town?	3 740 196	3 000 000	19 384 089
8. Price of van Gogh's "Sunflowers"?	39 700 000	15 000 000	3 875 825 789
9. Population size of Indonesia?	246 900 000	20 000 000	682 312 629
10. Number of views of "Gangnam Style"?	1 764 039 000	278 000 000	1 610 122 583

*Number of responses N = 175.

Table 5. Year 2 undergraduate medical students' collective answers to the 10 questions in the area outside of their expertise (astronomy)

Question	Correct answer	Students' collective answer (median)	Students' collective answer (mean)
1. Distance Earth–Sirius (in light–years)?	9	6900	5 800 659 084
2. Number of Saturn's moons?	62	12	20
3. How many times Jupiter heavier than Earth?	318	811	5 681 716 865
4. When was Uranus first discovered?	1781	1807	1720
5. Distance Sun–Milky Way Centre (in ly)?	27 000	5 000 000	22 584 267 640
6. How much Sun heavier than Earth?	332 900	8 000	8 561 716 703
7. Speed of Solar Wind (in km/h)?	1 440 000	43 027	7 948 573 823
8. Years since comet killed off dinosaurs?	65 000 000	24 564 456	1 396 252 256
9. Kilometres from Sun to Jupiter?	780 000 000	8 728 001	1 239 338 648 469
10. Years since solar system created?	4 568 000 000	7 119 851 052	721 049 090 361

*Number of responses N = 170.

171 responses received, respectively). **Table 6** shows the summary result of the three exercises, presenting both the rank and the percentile of the collective answer (based on either median or mode) among all individual answers provided by the students in three consecutive exercises where students had a decreasing level of expert knowledge. When answering the questions in their own area of expertise, the collective numerical median answer was 35/168 (21st centile) of the most accurate answers; in general knowledge, it was 21/176 (12th centile) most accurate answers; and in an area with no expertise, the group answer was the 12/171 (7th centile). However, the mean value of the collective didn't rank highly in any of the three exercises–in fact, it ranked near the bottom: 127/168 (76th centile) in Year 1 medical knowledge, 164/176 (93rd centile) in general knowledge and 164/171 (96th centile) in astronomy.

Table 7 shows the results of the exercise in recognizing the older of the two celebrities, based on the sample of 122 participating students. The age indicated in the table was relevant to October 17, 2013. All 20 questions were phrased as: "*Would you say that Celebrity X is older than Celebrity Y*?" The possible answers in the first round were

Table 6. The rank and the percentile of the collective answer (based on either median or mean) among all individual answers provided by the students in three consecutive exercises where students had a decreasing level of expert knowledge*

	Collective answer—median			Collective answer—mean		
Exercises on collective knowledge	Rank	Percentile (% top answers)	"Error size" parameter	Rank	Percentile (% top answers)	"Error size" parameter
Medical (Year 1) knowledge	35/168	20.83%	725	127/168	75.60%	48975
General knowledge	21/176	11.93%	38	164/176	93.18%	5430
Astronomy knowledge	12/171	7.02%	1132	164/171	95.91%	663265715

*Addition of the collective answer increased the total number of received answers by one, resulting in 168, 176 and 171 responses being ranked in each exercise, respectively; percentile of eg, 20.83% means that the collective response ranked among the 20.83% most accurate individual responses).

Table 7. Results of the exercise in recognizing the older of the two celebrities (N = 122)*

Older celebrity	Younger celebrity	Difference (years)	% correct (2—category system: yes/no)	% correct (4—category system: yes/no/ns/b)
Roger Federer (32)	Andy Murray (26)	6	97%	97%
George Clooney (52)	Brad Pitt (49)	3	95%	96%
David Beckham (38)	Prince Harry (29)	11	96%	96%
Tiger Woods (37)	Lewis Hamilton (28)	11	93%	95%
Jennifer Aniston (44)	Paris Hilton (32)	12	97%	94%
Miley Cyrus (20)	Justin Bieber (19)	1	93%	92%
Ben Affleck (41)	Ashton Kutcher (35)	6	85%	85%
George W. Bush (67)	Tony Blair (60)	7	85%	80%
Kim Kardashian (33)	Adele (25)	8	82%	79%
Jennifer Lopez (44)	Britney Spears (31)	13	83%	78%
Angela Merkel (59)	JK Rowling (48)	11	71%	73%
Michael Jackson (50)	Elvis Presley (42)	8	75%	67%
Barack Obama (52)	David Cameron (47)	5	66%	62%
Tom Cruise (51)	Nicole Kidman (46)	5	64%	60%
Katy Perry (28)	Rihanna (25)	3	63%	59%
Jay–Z (43)	Eminem (40)	3	56%	57%
Dustin Hoffman (76)	Robert de Niro (70)	6	44%	52%
Paul McCartney (71)	Mick Jagger (70)	1	59%	52%
Madonna (55)	Susan Boyle (52)	3	55%	51%
Shakira (36)	Beyonce (32)	4	43%	43%

*The questions were phrased as: "Would you say that Celebrity X is older than Celebrity Y?". The possible answers in the first round were "Yes" or "No" (2–category system); and in the second round the students were also allowed "Not sure" (when they were familiar of both celebrities, but it was too difficult to judge) and leaving the answer "Blank" deliberately (when not knowing one or both celebrities), in order to increase the chance of the entire collective of students to answer correctly. The latter type of "scoring" is used in the CHNRI method.

"Yes" or "No" (2–category system); and in the second round the students were also allowed "Not sure" (when they were familiar of both celebrities, but it was too difficult to judge) and leaving the answer "Blank" deliberately (when not knowing one or both celebrities), in order to increase the chance of the entire collective of students to answer correctly. The latter type of "scoring" is used in the CHNRI method.

The results show that, when everyone needed to provide a "Yes" or "No" answer, regardless of their confidence in their own answer, the collective was correct in 18/20 cases (90%), with 8 students outperforming the results of the collective–all of them with 19/20 correct answers (95%). This means that the collective answer based on this type of response ranked in the top 7.3% of individual answers. However, when the students were allowed to use the system of responses in which those who were not confident of their answer were allowed to ask for half a point, or withdraw from responding entirely, in order to improve the scores of the collective, the results changed somewhat. Looking at all specific celebrity pairs, they were not clearly better than when everyone gave an answer regardless of their confidence in being correct. However, with this type of scoring the collective was correct in 19/20 cases (95%), while the 3 most successful individuals among the 122 students now had 17/20 correct guesses (85%). This clearly shows that many students opted to only receive half a point, or withdrew, because the small group among them who gave best individual answers did not repeat the level of success from the first round of scoring in this second round–although they did manage to further improve the collective answer. A subsequent analysis showed that the median frequency of choosing the "*Not sure*" answer when this was possible was 44 (range: 3–59), or about one third of students, with very wide range–depending on the level

of difficulty of the question. The option "blank" was used much less frequently, with a median of 7 (range: 0–35).

The **Online Supplementary Document** presents several additional analyses. Figures S1–S3 show that the number of digits of the correct answer does not seem to be related to the likelihood that the group will identify the correct answer–this only seemed to possibly be the case in the exercise where students had expertise (ie, Year 1 medical curriculum questions), but was not replicated in the other two exercises. Figure S4, related to the fourth exercise, shows that the proportion of those guessing correctly in the group was associated with the age difference between the two celebrities, as might be expected.

DISCUSSION

The analyses conducted in this study tried to provide insights into quantitative properties of human collective knowledge, many of which are relevant to better understanding of the properties of the CHNRI method as originally proposed. First, the CHNRI method relies on the opinion of experts that is based on their knowledge of a specific subject, and asks them to express their optimism about research ideas through scores. Through this series of exercises we wanted to explore if this approach is likely to result in better predictions than if persons with limited knowledge of the subject are also invited to prioritize health research, or if persons with no knowledge at all are invited. In the student exercise in their own area of expertise (Year 1 medical curriculum, **Table 3**), the first 5 answers given by the students as a collective median value were all exactly right or extremely close (taking into account that the number of genes in the human genome was indeed close to 40 000 in their earlier textbooks, and it was only revised down to about 23 000 more recently). This level of precision was not observed in their responses to general knowledge questions (**Table 4**), or questions on astronomy (**Table 5**).

However, there are worrying signs that, when the majority of students don't know the correct answer to a question that should be covered by their expert knowledge, the collective median can be very wrong. The examples are the case of the number of erythrocytes in 1 mL of blood (where the collective median was 3 orders of magnitude smaller than the correct value) or the number of nucleotides in the human genome (where the underestimate was by 2 orders of magnitude) (**Table 3**). Because of those two questions, where most of the students didn't even know the right order of magnitude, the parameter "error size" of the collective median was even greater for the exercise on Year 1 medical knowledge, than it was for the exercise in general knowledge (**Table 6**). Although this may seem surprising at first, it can be easily explained. The parameter "error size" is very sensitive to the size of the departure from each of the 10 correct answers. In general knowledge questions, collective median answers were always reasonably close to the correct answers in terms of students' being able to guess the correct order of magnitude for the answer, as all the questions were related to topics in which they had at least some knowledge. However, a specific question in their own area of expertise in which they had no knowledge could quickly lead to very large departures from the correct answer. It would be difficult, given a small sample size, to reach a definite conclusion that there are some experts who do better than the crowd–"*the superforecasters*" [8], although this remains a possibility.

The exercise in the knowledge of astronomy (**Table 5**) was interesting because it clearly showed that humans do not possess a "cryptic" ability to collectively predict values on which they do not have any knowledge as individuals with any precision. This suggests that "wisdom of crowds" only works when the majority of participants in the group have at least some private knowledge of the quantity that is being predicted. As an example, the students had some intuition on the possible year when Uranus could have been discovered, the number of Saturn's moons, or even the number of years since the Solar system was created–they got the order of magnitude correct in those three questions. However, when asked about quantities of which they knew nothing, nor had any intuition, they were typically wrong by several orders of magnitude when their collective medians were compared to the correct answers.

Collective medians typically performed well across all three exercises: the collective median was among the 20.83% of the most accurate responses in the medical knowledge, 11.93% in the general knowledge, and 7.02% in the astronomy knowledge. We propose that the collective median is actually not among the top 10% scores in the area of expertise, because there is a smaller group of students among the entire cohort with excellent knowledge, and who would be seen as the top of their class. These students know the correct answers and the rest of the class simply dilutes their accuracy and moves the collective median away from the perfectly accurate response. We believe that this explains why the collective median in the area of expertise was only at the 21st percentile of the most accurate answers. However, as the collective moves towards answering the questions outside of the area of their expertise, the collective median begins to move up the ranks. Once there are no longer individuals who could easily answer all 10 questions with high accuracy, the collective median progresses to the 12th percentile (in the general knowledge exercise) and 8th percentile (in astronomy exercise).

We propose a mathematical explanation for this, which is relevant to the relationship between the correct answer and

the distribution of all responses in a series of questions. After each question, the collective median will be exactly at the 50th percentile of answers. When the distribution of answers is compared to the correct answer, the error size of the median will either be at the 50th percentile of the group or smaller. For individual students who don't have any knowledge on the subject and are simply guessing, they can expect to alternate between a position above and below the 50th percentile randomly, and occasionally making gross mistakes. After enough time and many iterations, the collective median of a group who are guessing entirely unknown quantities will always be either at the 50th percentile, or above, while the rest of individual answers will be above or below the 50th percentile half of the time. After a sufficient number of questions, this should ensure that the collective median acquires Rank 1, because median can sometimes be very close to a correct answer, but never worse than 50th percentile of all group's guesses. This protects it from gross errors that all other students will eventually experience over a large number of guesses. This may be a general mechanism that explains why collective median eventually outperforms individuals in a long time series of predictions of entirely unknown quantities.

All of the above is relevant to collective medians. Turning our attention to collective means, they did not fare well at all. They were at the 76th percentile of ranks in the area of medical knowledge, 94th in the area of general knowledge, and 96th in the area of astronomy. We found the explanation to this poor performance in a number of extremely wrong predictions made by several individuals, who made mistakes of such magnitude that they completely dominated the collective mean. Because of this, we suggest that – when the answers are being predicted in a quantitative form – medians will be more reliable than the means. One question that could be raised here is whether the entire cohort of medical students can be trusted to take this sort of exercise seriously, because if a small group deliberately put down extreme responses, this would certainly have an effect of skewing the mean.

The exercise in "guessing the older of the two celebrities" allowed us to establish that, in an area of "relative" expertise (because it has become difficult to avoid information on the celebrities that were chosen). There is considerable accuracy in collective prediction when "Yes"/"No" answers are allowed and the answer given by the majority is chosen as the correct one. The collective was correct in 90% of cases, and this translated to the rank 9/123 (8th percentile in the ranks), with 8 individuals who recorded 95% of correct answers and out–performed the collective. This exercise was analogous to a large extent to the "Ask the audience" joker that is used in the quiz show "Who wants to be a millionaire?", as mentioned earlier, and the accuracy of 90% is very similar to the one of about 91% observed in the quiz show.

The key question in this exercise was whether the collective response could be further improved by allowing some individuals, who were not confident in their answers, to minimise the "damage" to the collective by choosing "not sure" (which still gives them a guaranteed 50% of available points) or to drop out from the sample. When this option was given, the accuracy of the collective answer increased to 95%, while the three best individual answers only achieved 85%. A question–by–question comparison of 20 individual answers between the two types of scoring doesn't indicate that the collective answer with the 2nd type of scoring (4 options) is consistently better than the binary "Yes"/"No" type of scoring, so we cannot be sure that this finding is generalizable, rather than a chance effect, and we should continue to explore this with more questions and using larger sample sizes to confirm it.

We will now consider how the findings of this study are relevant to "validation" of the CHNRI method. This study shows that the collective knowledge in an area of expertise is likely to lead to more accurate responses than the collective knowledge in an area outside of the expertise. Moreover, the exercise shows that it may be better to only invite a reasonably small, highly selected group of experts and rely on their collective prediction, rather than trying to seek expertise from a large group, which may lead to deviations from the optimal collective prediction. This justifies the strategy that has been used in many early CHNRI exercises, where as few as 10–15 leading experts in a narrow research field were invited to conduct the exercise on setting research priorities in their field. Moreover, the type of response used in CHNRI exercises ("Yes" – "No" – "Not sure" – "Blank") seems to slightly improve the collective prediction in comparison to the alternative, where all scorers are forced to choose between only two binary options. However, the difference between the two types of scoring resulted in predictions that could be considered surprisingly similar, so further experiments will need to resolve whether there is a real difference between the two approaches or not. If there is no difference, then perhaps the "Yes"/"No" answer could be preferred as simpler and more discriminative in the process of prioritisation, because too many "not sure" answers lead to scores that show regression to the mean and the discriminatory power of the scoring process is gradually lost. This, therefore, remains an unresolved question that warrants further investigation.

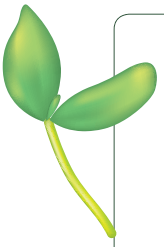
Applications of "crowdsourcing" are finding ways into many areas of human activity. In parallel, many interesting scientific experiments are being performed to improve our understanding of the principles underlying and governing crowdsourcing. Recent studies showed that sharing the in-

formation on confidence in their answers between the individuals in the group can substantially improve the prediction of the group, as we could see in our study (**Table 7**), but if those most confident are wrong, then it can also lead the collective opinion to dramatically wrong decisions [8,9]. Independence of the provided opinion, such as in the CHNRI exercise, is very important because studies have convincingly shown that interactions between participants in the group and social influence may both improve and undermine the "wisdom of crowds" effect [10,11]. We should also mention that this research was conducted in "artificial", well–controlled conditions, but in the real world every group will have its own unique dynamics. In many contexts, collective knowledge, opinion or intelligence may not be the main factor influencing the decisions, which is a limitation of this type of research and of its applications in complex real–world scenarios.

There seems to be agreement between researchers that select groups of "best–performing" experts can reach an optimal collective result with sample sizes as small as five, which cannot be easily improved by increasing sample size [12,13]. This observation has a potential practical application in the field of medical diagnostics [13]. However, it has also been shown that a well–designed mathematical or statistical model would still outperform any collective human opinion [13]. Two further interesting applications of crowdsourcing in the fields of medicine and health research have been proposed recently. One study proposed that, in the absence of clear guidelines on indications, stabilization of the prevalence of use of certain drugs–such as antidepressants–at the level of the whole population might indicate the optimal usage. This is because the stabilized frequency at the population level is likely to reflect hundreds of thousands of decisions on continued usage, made by treated individuals based on their personal experiences [14]. Finally, it has been proposed that complex, expensive and bureaucratic processes of research evaluations, such as the Research Excellence Framework (REF) that takes place every 6 years in the UK, could be replaced by crowd–sourced "prediction markets" [15]. Prediction markets enable individuals to trade "bets" on whether a specific outcome would occur or not, and they have been shown to be successful at predicting outcomes in different areas of human activity, such as sport, entertainment and politics. Given that they are based on expert judgements, which also form the basis of REF in the UK, there is no reason why prediction market could not theoretically offer an alternative to the REF that could be updated annually, or even track the performance in real time [15].

CONCLUSION

Our experiments showed that the collective knowledge of a group with expertise in the subject should always be very close to the true value. In most cases and under most assumptions, the collective knowledge will be more accurate than the knowledge of an "average" individual, but there always seems to be a small group of individuals who manage to out–perform the collective. The accuracy of collective prediction may be enhanced by allowing the individuals with low confidence in their answer to withdraw from answering. This study showed that the CHNRI method is based on the premises and designs that are likely to maximise the predictive value of the group: experts are being invited to score proposed research ideas (instead of persons with limited knowledge, or lay persons); experts are providing their answers independently (to protect the end result from social influences); and they are using the scoring system that is expected to maximise the accuracy of the collective answer over the individual ones.



Acknowledgements: The authors thank the Year 2 medical students of the University of Edinburgh for their participation in this study.

Funding: This work was supported by the CHNRI Foundation award to the University of Edinburgh.

Ethical approval: Obtained from the Centre for Population Health Sciences, the University of Edinburgh.

Authorship declaration: IR developed the idea for the paper. IR and KYC conducted the student exercise. IR and SY co–wrote the paper. IR, SY and SC designed and performed the analyses. SC, SY, KYC and KW provided important intellectual content.

Declaration of interest: IR is an editor–in–chief of the Journal of Global Health. To ensure that any possible conflict of interest relevant to the journal has been addressed, this article was reviewed according to best practice guidelines of international editorial organizations. The authors completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). Authors declare no conflicting financial or other interest related to the work detailed in this manuscript.

REFERENCES

1 Surowiecki J. The wisdom of crowds. London: Abacus, 2013.
2 Surowiecki J. Democracy: dreams of the common good. In: Surowiecki J. The wisdom of crowds. London: Abacus, 2013.
3 Graham B. The intelligent investor. New York: Harper Collins, 2003.
4 Tripadvisor. Available at: https://www.tripadvisor.co.uk/. Accessed: 15 April 2016.</eref>
5 Internet Movie Database. Available at: http://www.imdb.com/. Accessed: 15 April 2016.
6 Rudan I, Chopra M, Kapiriri L, Gibson J, Lansang MA, Carneiro I, et al. Setting priorities in global child health research investments: universal challenges and conceptual framework. Croat Med J. 2008;49:307-17. Medline:18581609 doi:10.3325/cmj.2008.3.307
7 Rudan I, Gibson JL, Ameratunga S, El Arifeen S, Bhutta ZA, Black M, et al. Setting priorities in Global child Health Research Investments: guidelines for implementation of the CHNRI Method. Croat Med J. 2008;49:720-33. Medline:19090596 doi:10.3325/cmj.2008.49.720
8 Tetlock PE, Gardner D. Superforecasting: the art and science of prediction. London: Random House Books 2015.
9 Koriat A. When are two heads better than one and why? Science. 2012;336:360-2. Medline:22517862 doi:10.1126/science.1216549
10 King AJ, Cheng L, Starke SD, Myatt JP. Is the true "wisdom of the crowd" to copy successful individuals? Biol Lett. 2012;8:197-200. Medline:21920956 doi:10.1098/rsbl.2011.0795
11 Lorenz J, Rauhut H, Schweitzer F, Helbing D. How social influence can undermine the wisdom of crowd effect. Proc Natl Acad Sci U S A. 2011;108:9020-5. Medline:21576485 doi:10.1073/pnas.1008636108
12 Mannes AE, Soll JB, Larrick RP. The wisdom of select crowds. J Pers Soc Psychol. 2014;107:276-99. Medline:25090129 doi:10.1037/a0036677
13 Kattan MW, O'Rourke C, Yu C, Chagin K. The wisdom of crowds of doctors: their average predictions outperform their individual ones. Med Decis Making. 2016;36:536-40. Medline:25878196 doi:10.1177/0272989X15581615
14 Patten SB. The wisdom of crowds (vox populi) and antidepressant use. Clin Pract Epidemiol Ment Health. 2015;11:1-3. Medline:25674154 doi:10.2174/1745017901511010001
15 Munafo MR, Pfeiffer T, Altmejd A, Heikensten E, Almenberg J, Bird A, et al. Using prediction markets to forecast research evaluations. R Soc Open Sci. 2015;2:150287. Medline:26587243 doi:10.1098/rsos.150287

journal of global health

Setting health research priorities using the CHNRI method: VI. Quantitative properties of human collective opinion

Sachiyo Yoshida[1], Igor Rudan[2], Simon Cousens[3]

[1] Department for Maternal, Newborn, Child and Adolescent Health, World Health Organization, Geneva, Switzerland

[2] Centre for Global Health Research, The Usher Institute for Population Health Sciences and Informatics, The University of Edinburgh, Edinburgh, Scotland, UK

[3] Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK

Introduction Crowdsourcing has become an increasingly important tool to address many problems – from government elections in democracies, stock market prices, to modern online tools such as TripAdvisor or Internet Movie Database (IMDB). The CHNRI method (the acronym for the Child Health and Nutrition Research Initiative) for setting health research priorities has crowdsourcing as the major component, which it uses to generate, assess and prioritize between many competing health research ideas.

Methods We conducted a series of analyses using data from a group of 91 scorers to explore the quantitative properties of their collective opinion. We were interested in the stability of their collective opinion as the sample size increases from 15 to 90. From a pool of 91 scorers who took part in a previous CHNRI exercise, we used sampling with replacement to generate multiple random samples of different size. First, for each sample generated, we identified the top 20 ranked research ideas, among 205 that were proposed and scored, and calculated the concordance with the ranking generated by the 91 original scorers. Second, we used rank correlation coefficients to compare the ranks assigned to all 205 proposed research ideas when samples of different size are used. We also analysed the original pool of 91 scorers to to look for evidence of scoring variations based on scorers' characteristics.

Results The sample sizes investigated ranged from 15 to 90. The concordance for the top 20 scored research ideas increased with sample sizes up to about 55 experts. At this point, the median level of concordance stabilized at 15/20 top ranked questions (75%), with the interquartile range also generally stable (14–16). There was little further increase in overlap when the sample size increased from 55 to 90. When analysing the ranking of all 205 ideas, the rank correlation coefficient increased as the sample size increased, with a median correlation of 0.95 reached at the sample size of 45 experts (median of the rank correlation coefficient = 0.95; IQR 0.94–0.96).

Conclusions Our analyses suggest that the collective opinion of an expert group on a large number of research ideas, expressed through categorical variables (Yes/No/Not Sure/Don't know), stabilises relatively quickly in terms of identifying the ideas that have most support. In the exercise we found a high degree of reproducibility of the identified research priorities was achieved with as few as 45–55 experts.

Correspondence to:
Sachiyo Yoshida
Department for Maternal, Newborn, Child and Adolescent Health
World Health Organization
Avenue Appia 20
CH–1211 Geneva 27
Switzerland
yoshidas@who.int

In 1906, Galton suggested that a group of individuals tend to make better predictions as a collective than any individual. Since then, our understanding of collective decision–making, termed by some as the "Wisdom of Crowds", has grown considerably [1]. Crowd–sourcing has become an increasingly important human tool to address many problems – from government elections in democracies [2], formation of stock market prices [3], to modern online platforms such as TripAdvisor (to advise on the best hotels and restaurants) [4] or Internet Movie Database (to advise on the best movies, TV shows, etc.) [5], all of which are based on personal opinions of many hundreds or thousands of participants. The CHNRI method (the acronym for: Child Health and Nutrition Research Initiative) also uses crowdsourcing as the major component of the process to set priorities among many competing health research ideas [6,7]. It relies on large groups of scientists who are invited to participate in each exercise. Within the CHNRI process, several dozens (or even hundreds) of scientists are typically invited first to generate, and then to assess many competing health research ideas using a pre–defined set of priority–setting criteria. Their collective optimism towards each research idea with respect to specific criteria is measured and the research ideas are then ranked according to the scores they achieve across all criteria.

However, researchers typically question several concepts in relation to the "validity" of the CHNRI exercises. The first question is fundamental to the entire process, asking the developers of the method to demonstrate convincingly that the opinion of a large expert group is more reliable and trustworthy than the opinion of only one, or a very small number of experts. This question has been addressed in a previous paper in this series [8], which demonstrated that the collective knowledge of a group (rather than opinion) generally outperforms the knowledge of any single individual. While for factual knowledge there is a "gold standard" against which we can compare the response of the collective to that of individuals, for opinions about future outcomes there is no such "gold standard". Nevertheless, given that individual knowledge, or lack of it, underlies a significant part of individual opinion, and that the same governing principles that make the collective knowledge superior to individual knowledge (described in our previous paper [8]) should also apply to opinion, we consider this question largely addressed. The substantial literature on so–called "prediction markets" provides further evidence of the reliability and effectiveness of collective opinion in comparison to individual opinion in predicting future events [9,10].

The second question concerns the "optimal" sample size of researchers to be invited to conduct a CHNRI exercise. Here, "optimal" refers to a minimum number of experts needed from a larger, global "pool" of experts, in order to reduce the cost and complexity of conducting the exercise while obtaining a replicable collective opinion. The question of the "sufficient" sample size can be investigated by exploring at which point addition of further experts from the larger, global "pool" of experts ceases to influence the outcomes of the CHNRI process. The third question is related to the composition of the sample of experts, and how this composition can potentially affect the final scores. Do the background characteristics of the experts invited to participate affect their collective opinion in such a way that one subgroup of experts would provide systematically different scores from another subgroup?

In this article, we address the latter two questions by exploring some of the quantitative properties of human collective opinion. We study the special case where the collective opinion is based on a set of individual opinions, all of which are expressed in the form of simple categorical variables. These variables relate to the optimism expressed by each participating expert regarding the extent to which each proposed research idea meets the different priority–setting criteria [6,7]. The opinion provided by the participating experts can be expressed as "Yes" (equals 1), "No" (equals 0), "Not sure" (equals 0.5) and "I don't know" (equals blank input), which is the typical input required in the CHNRI method. This special case is of particular interest, because in our previous paper [8] we demonstrated the effectiveness of this method of expressing individual opinion in comparison to other types. Finally, one of the concerns about this way of collecting opinion from groups of experts is the impact of low response rates and subsequent self–selection bias. We will mention this concern here because we find it potentially very important, although it will be difficult to study and we will not attempt to address it in this paper.

METHODS

In order to answer the latter two questions posed in the introduction, we conducted statistical analyses of the inputs provided by the group of experts who took part in a previous CHNRI exercise. These analyses focused on identifying whether there was a point of "saturation" in collective opinion. "Saturation" here refers to the idea that beyond a certain sample size of experts, adding further experts' opinions does not significantly change the results of the process. To study this, we used the data set with quantitative input from the experts who took part in a CHNRI exercise on newborn health in this series [11], which is freely available as a supplementary online material to the article in question [11]. All input was provided in the form of a simple categorical variable (ie, optimism towards each idea expressed as "Yes" (equals 1), "No" (equals 0), "Not sure" (equals 0.5) and "I don't know" (equals blank input)).

Our analysis strategy involved drawing many random sub–samples, with replacement, from the full sample of 91 expert participants in the CHNRI exercise on newborn health. The experts scored a set of 205 proposed research ideas [11]. Our aim was to identify the minimum sample size of experts required to produce stable results. We used two metrics to assess stability. First, we compared the 20 most highly ranked ideas for each resampled data set with the 20 most highly ranked ideas in the whole data set (ie, all 91 experts) and calculated how many ideas appeared in both top 20 lists. If all the opinions were assigned entirely at random, then we would only expect about 2 research ideas on average (out of the total of 205) to be in common across two samples. Given this reasonably low expected agreement by chance, we arbitrarily defined results as being stable when 15 (or more) of the 20 highest ranked ideas were concordant with those based on the opinion of the full sample of 91 experts. We believe that such an occurrence indicates a high level of stability/replicability compared with the 2 expected purely by chance.

Previous studies into the point of saturation in collective opinion

The question of the sample size at which the "saturation" of information occurs has been vigorously discussed over many years in relation to qualitative research, where interviews conducted with the participants are recorded and analysed to obtain insights into a wide variety of research topics. In qualitative research, saturation is typically described in the context of obtaining the "appropriate" sample size at which no new ideas, findings, or problems are found. Determining the "appropriate" sample size is critical, because a sample that is larger than needed would result in inefficient use of research funds, resources and time. On the other hand, too small a sample size may result in limited validity of the research findings.

The idea of "saturation" was first introduced in the late 1960s [12] through the notion that, though every research participant can have diverse ideas in principle, the majority of qualitative studies will inevitably reach a point of saturation. Since the work by Glaser and Strauss [12], researchers have attempted to provide sample size guidance for various research disciplines. Proposed sample sizes have ranged from fifteen in all qualitative research disciplines [13] to sixty [14] in the area of ethnographic interviews. These proposed sample sizes were rarely accompanied by a clear justification or description of how they were derived.

However, the idea of saturation does not necessarily translate to CHNRI exercises, where opinions are submitted in a form of quantitative categorical variables. This gives us perhaps a rare opportunity to perform an assessment of the quantitative properties of human collective opinion by analysing a data set underlying a typical CHNRI exercise. We found one study that attempted to analyse the stability of responses of the 23 health care and patient safety experts who participated in a Delphi survey using a categorical rating scale [15], which is the most similar case to the CHNRI process that we were able to find in the literature. In that study [15], the responses to each item were scored on a rating scale from 1 to 4, with "1" being unimportant to "4" being very important. The responses obtained in the first round of the survey were processed using sampling with replacement to produce hypothetical samples of 1000 and 2000 participants, from the initial sample size of 23 subjects. Then, means and 95% confidence intervals for the scores of the original 23 participants were compared with the hypothetical samples. Substantial similarity of inferential statistics between the actual and hypothetical samples was observed, from which the authors concluded that the "stability" of results was already achieved with only 23 actual study participants [15]. Clearly, this interpretation was limited by having an original sample as small as 23 individuals to generate large bootstrapping samples, and the result needs to be replicated using a larger initial sample of individuals to generate bootstrapping samples. In our study, the key improvement will be drawing sub–samples smaller than the original sample, while in the approach described in this study samples were created that were much larger than the original sample – which is an approach with major limitations.

Defining "saturation" in our study

In our study, we defined "saturation" in two ways. First, we defined it as the point where we observed replicability in the collective rankings of top 20 research ideas (among a total of 205 assessed) between two randomly generated sub–samples of a given sample size. In other words, involving further experts would no longer be expected to make any important difference to the 20 most highly ranked priorities. Given that randomness inherent to the process of sampling makes it unrealistic to expect all 20 priorities to always replicate at a certain sample size, and taking into account low "a priori" probability of replication (only 2 among the 20 most highly ranked research ideas would be expected to replicate by chance alone), we needed to define "saturation" arbitrarily. We considered the specific sample size as "saturation–reaching" when the same 15 (or more) research ideas in any two randomly generated samples of a specific size were expected to be found among the 20 most highly ranked research ideas in both samples.

Second, we used Spearman's correlation coefficient to compare the ranks assigned to all 205 proposed research ideas by the randomly generated sub–samples with the ranks derived from the full sample. We considered "saturation" to be achieved when the median rank correlation coeffi-

cient reached or exceeded 0.95 (which is an extremely high rank correlation coefficient). We believe that both definitions of saturation are stringent and conservative from the statistical point of view.

Database used in this analysis

We used anonymised raw scores provided by the participants in the CHNRI exercise on newborn health [11]. The database included all individual scores from 91 participating experts that were assigned to all 205 proposed research ideas using 5 pre–defined criteria. The criteria used in the exercise are summarized in **Box 1**, and they were posed in the form of simple "yes/no" questions. The requested input was provided in the form of numbers: 0 (meaning "no"), 0.5 ("informed, but undecided answer"), 1 ("yes"), and blank ("insufficiently informed"). "Blank" was used whenever the participants did not feel that they possessed enough technical knowledge to be able to answer, which is different from an "informed, but undecided" answer, where the expert could neither agree nor disagree although they felt that they had enough knowledge on the topic.

Statistical analysis

We used resampling with replacement, sometimes referred to as "bootstrapping", to simulate the diversity of samples drawn from a larger global pool of experts. All analyses were performed using the statistical program STATA 13.0 (www.stata.com). To study how the rankings assigned to proposed research ideas change and converge with increasing sample sizes of experts, we generated samples ranging in size from minimum 15 to a maximum of 90. For each selected sample size, 1000 random bootstrap samples were drawn.

Two statistical analyses were then performed to examine how the ranking list of research ideas changed as the number of experts contributing to the CHNRI exercise increased. In the first analysis, we examined the concordance in the top 20 research ideas between 1000 randomly generated subsamples of the same size that were developed using the bootstrap method. In the second analysis, we used Spearman's rank correlation coefficient to examine the concordance in the ranking order of all 205 research ideas between 1000 randomly generated subsamples of the same size that were developed using the bootstrap method.

Box 1. The five criteria used in the exercise.

Criterion 1. Answerability: Can the research question be answered ethically?

Criteria 2. Efficacy/Effectiveness: Can the new knowledge lead to an efficacious intervention or programme?

Criteria 3. Deliverability and acceptability: Is the proposed intervention or programme deliverable and acceptable?

Criteria 4. Maximum potential for disease burden reduction: Can the intervention or program improve newborn health substantially?

Criteria 5. Effect on equity: Can the interventions on program reach the most vulnerable groups?

Analysis of subgroups within the full sample

Research priority scores (RPS) were recalculated for each research question in sub–samples of scorers that were defined by participants' self–classified background and the country in which they were based. Participants originally classified themselves as researchers, policy makers, donor representatives, program managers or health practitioners (multiple choices were not allowed), and this information is available in the original paper [11]. In this exercise, we had combined all categories other than researcher into one category as "non–researcher", as the numbers of participants falling into each of the non–researcher categories were small. The country where the scorer was based was classified by the level of income as either a "high–income country" (HIC) or a "low– or middle–income country" (LMIC), using the World Bank's categorization [16]. We explored: (i) the differences in median scores that different sub–groups of scorers (ie, researchers vs non–researchers; and HIC–based vs LMIC–based) assigned to different criteria; the median scores were determined across all 205 research ideas to investigate whether subgroups of scorers systematically scored particular criteria differently; (ii) the overlap between the top 20 research ideas identified by different sub–groups of scorers (ie, researchers vs non–researchers; and HIC–based vs LMIC–based).

RESULTS

Figure 1 shows the how concordance with respect to the top 20 priorities increased as the number of sampled scorers increased. Note that when resampling 90 scorers with replacement, concordance with the top 20 priorities based on the original sample of 91 experts would not be expected to reach 100%. This reflects the fact scores derived from the original sample of 91 experts are themselves subject to sampling variation. The median concordance (across the 1000 sub–samples drawn for each sample size) increases from 12/20 (60%) with a sample size of 15 to 15/20 (75%) with a sample size of 55 experts. Thereafter there is no clear improvement in concordance with increasing sample size. The interquartile range for concordance with a sample size of 55 is 14/20 to 16/20 (70% to 80%) and this also appeared relatively stable as sample sizes were increased further. At a sample size of 90, the median concordance was 16/20 (85%) (IQR 15–16). Given that this gives an indication of the variability of the sample size we had available

to us for analysis, it appears that relatively stable results can be achieved with sample of 50 experts (median 14, IQR 13.5–15). There is little further increase in achieved overlap by increasing the pool of experts from 50 to 90 (**Figure 1**).

Figure 2 shows the relationship between the sample size of the scorers within the CHNRI newborn health exercise [11] and the median, IQR and range of Spearman's rank correlation for the ranks of all 205 proposed research ideas. As expected, the rank correlation coefficient increases as sample size becomes larger and a median correlation of 0.95 was reached at a sample size of 45 experts (median of the rank correlation coefficient = 0.95; IQR 0.94–0.96).

Among the 91 scorers in the newborn health exercise, 61 self–classified as "researchers" and 30 as "non–researchers"; 53 participants were based in HIC and 38 in LMIC. **Table 1** shows the differences in median scores (with inter–quartile range, IQR) that different subgroups of scorers (ie, researchers vs "non–researchers"; and high–income country (HIC)–based vs low– or middle–income country (LMIC)–based) assigned to different criteria. The differences between researchers and non–researchers were small, with non–researchers being slightly more optimistic about maximum potential impact, but all differences were well within the limits predicted by inter–quartile ranges. Larger differences were observed between HIC–based and LMIC–based researchers, with the latter tending to provide more optimistic scores, ranging from a 7 to a 24 point–difference on a scale from 0 to 100. The smallest difference was noted for answerability, followed by effectiveness and

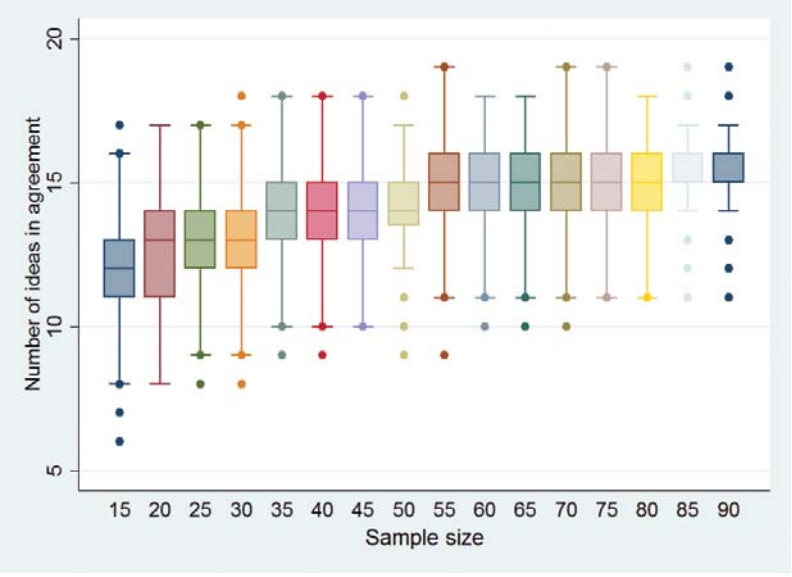


Figure 1. Level of overlap among the top 20 ranked research ideas (Y–axis) by the size of the sample of randomly selected experts (X–axis) from a total pool of 91 experts using a bootstrap method (simulation 1000 times with replacement of already selected experts, using bsampling function). The size of randomly generated samples ranged from 15 to 90 and it was based on the CHNRI exercise on newborn health research priorities [11].

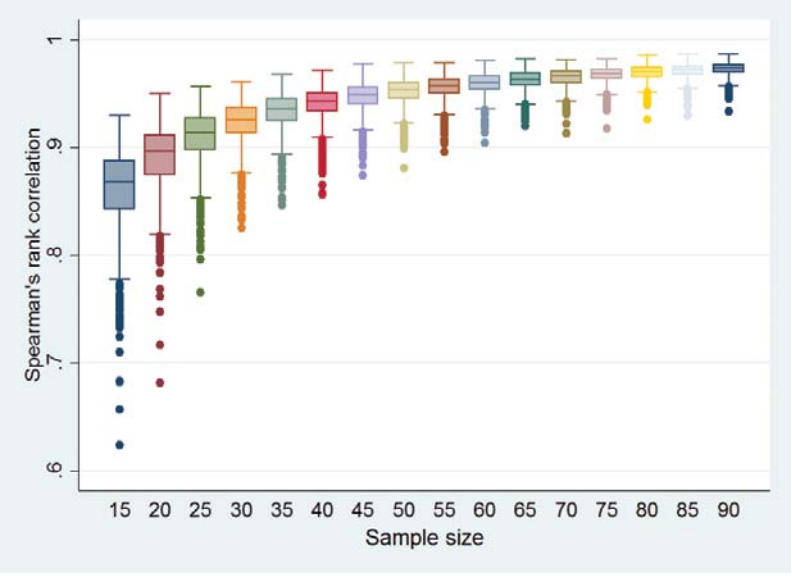


Figure 2. Spearman's rank correlation among all 205 ranked research ideas (Y–axis) by the size of the sample of randomly selected experts (X–axis) from a total pool of 91 experts using a bootstrap method (simulation 1000 times with replacement of already selected experts, using bsampling function). The size of randomly generated samples ranged from 15 to 90 and it was based on the CHNRI exercise on newborn health research priorities [11].

Table 1. The differences in median scores (with inter–quartile range, IQR) that different sub–groups of scorers (ie, researchers vs "non–researchers"; and high–income country (HIC)–based vs low– or middle–income country (LMIC–based) assigned to different criteria*

	All scorers (median, IQR) (N = 91)	Researchers (median, IQR) (N = 61)	"Non–researchers" (N = 30)	HIC–based (median, IQR) (N = 53)	LMIC–based (median, IQR) (N = 38)
Total score	63 (54–71)	62(54–70)	64 (53–73)	57 (47–66)	72 (61–80)
Answerability	76 (68–83)	76 (68–84)	77 (67–85)	74 (63–81)	81 (73–89)
Effectiveness	70 (61–77)	69 (61–78)	68 (59–78)	66 (54–74)	76 (66–84)
Deliverability	69 (58–77)	69 (59–78)	67 (57–78)	65 (54–72)	77 (65–84)
Maximum impact	42 (32–52)	39 (32–50)	44 (32–55)	32 (23–41)	54 (44–66)
Equity	57 (47–70)	57 (46–66)	60 (46–75)	48 (37–61)	72 (60–81)

IQR – interquartile range, HIC – high–income, LMIC – low– and middle–income

*The median scores were determined across all 205 research ideas in order to investigate if any sub–group of scorers deviated in their scoring of any particular criterion.

deliverability, while the largest differences were noted for maximum potential impact and equity.

Table 2 shows the overlap between the top 20 research ideas (RI–) identified by different sub–groups of scorers (ie, researchers vs "non–researchers"; and HIC–based vs LMIC–based). There was an overlap between researchers and "non–researchers" for 10 out of top 20 research ideas (50%). For HIC–based vs LMIC–based researchers, 8 of top 20 research ideas (40%) overlapped. We could judge this level of overlap against the expectation provided by the bootstrap analysis for comparable sample sizes. There is likely to be an effect of sub–stratification, which is smaller for the "researchers vs. non–researchers" comparison, but more considerable for the "HIC–based vs. LMIC–based" comparison.

Table 2. The overlap between the top 20 research ideas (RI–) identified by different sub–groups of scorers (ie, researchers vs "non–researchers"; and HIC–based vs LMIC–based)*

Rank	All scorers (N = 91)	Researchers (N = 60)	"Non–researchers" (N = 31)	HIC–based (N = 53)	LMIC–based (N = 38)
1	RI–30	**RI–30**	**RI–30**	**RI–30**	**RI–30**
2	RI–28	**RI–28**	**RI–28**	**RI–28**	**RI–23**
3	RI–15	**RI–15**	**RI–15**	RI–29	**RI–15**
4	RI–23	RI–29	RI–5	**RI–15**	RI–47
5	RI–33	**RI–23**	**RI–33**	**RI–33**	**RI–28**
6	RI–29	RI–36	RI–79	RI–7	RI–44
7	RI–149	RI–7	**RI–23**	RI–13	RI–18
8	RI–37	RI–13	RI–52	**RI–23**	RI–12
9	RI–5	**RI–33**	**RI–149**	RI–149	**RI–33**
10	RI–13	RI–58	RI–46	RI–36	RI–86
11	RI–79	**RI–149**	RI–47	RI–5	RI–58
12	RI–78	**RI–37**	RI–44	RI–37	**RI–46**
13	RI–36	RI–67	**RI–8**	RI–21	RI–60
14	RI–46	RI–75	**RI–78**	RI–55	RI–11
15	RI–8	**RI–78**	RI–129	**RI–79**	RI–8
16	RI–55	RI–86	RI–11	RI–22	RI–35
17	RI–52	**RI–55**	**RI–37**	RI–52	RI–67
18	RI–75	RI–12	**RI–55**	**RI–78**	RI–10
19	RI–58	**RI–8**	RI–127	RI–75	**RI–79**
20	RI–67	RI–158	RI–138	**RI–46**	**RI–78**

HIC – high–income, LMIC – low– and middle–income

*The research ideas that overlap between researchers vs "non–researchers", and HIC–based vs LMIC–based sub–samples, respectively, are in bold for easier recognition. Note: eg, RI–30 indicates research idea number 30 in the list of 205 ideas.

DISCUSSION

In this paper, we addressed two important questions relating to the quantitative properties of human collective opinion: (i) whether there is a point of "saturation" in the sample size, after which no significant changes in the collective opinion should be expected when more experts are brought into the exercise; and (ii) whether there is evidence that opinions differ between subgroups of experts defined by their professional background or their geographic location. We addressed both questions using data from a previous CHNRI exercise [11]. The data set based on the CHNRI exercise was useful in this regard, because it quantified a large number of expert opinions about 205 competing research ideas in a systematic and structured way, based on five pre–defined criteria, using simple categorical responses. We did not attempt to demonstrate that the collective would give more "useful" predictions than individual experts would, since this is examined in another paper on collective knowledge [8]. Perhaps the best support for the view that the opinion of a collective will prove more useful over time than that of individuals is provided in the literature on stock markets and prediction markets [3,9,10]. Over long periods of time, following the collective wisdom seems to be the most successful strategy. There are some important differences, though, because stock markets to a degree involve betting individual opinions against those of others, where investors are trying to identify stocks and shares that are undervalued by the collective opinion. Together, our previous paper from this series [8] and the large experience with stock markets and prediction markets [3,9,10] make a compelling case for collective decision–making.

Our analyses indicate that, in bootstrap samples that ranged in size from only 15 to 90, the level of overlap

among the top 20 scored research ideas increased with sample size up to about 50–55 experts. At this point, the median level of concordance stabilized at 15/20 top ranked questions (75%), with the interquartile range also generally stable (14–16). There was little further increase in overlap when the bootstrap sample of experts increased from 55 to 90. However, it should be noted that the overlap of 12/20 top ranked research ideas was achieved with sample sizes as small as 15 experts, as opposed to only 2 research ideas that would have been expected by chance. The conclusion from this analysis is that human collective opinion, when expressed in simple quantitative terms, tends to converge towards a similar outcome and saturate quickly. A sample size of 15 persons already shows an appreciable level of reproducibility, but with 50–55 experts the level of replicability becomes nearly equal to to that which is achievable with a sample size of 90.

It is important to note that the total sample of 91 experts, which is the maximum that we had available, represents only a sub–sample of a much larger global pool of experts. Therefore, it also carries a certain inherent random variation relative to the "total expert population". Sampling with replacement enables us to examine how variable the results for a given sample size will be, assuming that are full sample of 91 experts is representative of the diversity of the wider global pool. Thus two bootstrapped samples of size 91 participants would not be expected to have the top 20 research ideas fully replicated (although this is the entire original sample!). We used sampling with replacement to overcome, at least partly, the concern that the 91 experts are still only a reasonably small sample of the larger population and to produce a conservative estimate of the minimum sample size that produces replicable results in this particular CHNRI exercise.

We also tested the relationship between the sample size of the scorers and Spearman's rank correlation coefficient for the ranks of all 205 proposed research ideas. As expected, the rank correlation coefficient increased as the bootstrap sub–samples became larger. A median correlation of 0.95 was reached at the sample size of 45 experts (median of the rank correlation coefficient = 0.95; IQR 0.94–0.96), which again points to high reproducibility and relatively quick saturation.

Studying quantitative properties of human collective opinion, as opposed to collective knowledge verifiable against accepted facts, has the limitation that no gold standard is available against which the "accuracy" of the opinion can be judged. We therefore focused on the questions of saturation, reproducibility and subgroup stratification. Another limitation of this preliminary analysis is that it was based on a single data set from a previous CHNRI exercise. An analysis of multiple data sets with large numbers of experts and different numbers of research ideas being scored may offer further interesting insights into a nature of human collective opinion and results that are more generalizable than those based on the analysis of a single data set. Ideally, an analysis should involve as many experts as possible, because testing on exercises that only included reasonably small groups of experts will not be very useful. At this point, we should also declare that we can't predict the effects of low response rate and self–selection bias on the level of saturation achieved. The issue of missing responses of the experts who do not choose to participate should be explored separately and it remains an unresolved uncertainty related to the validity of the approach used in the CHNRI method.

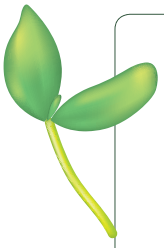
Any future work in this area could plan to acquire more data sets and replicate the analyses from this study. One emerging question that it would be interesting to answer is to examine the main determinants of the observed level of concordance in ranking lists. Examples of possible determinants are the composition and the nature of the proposed research ideas, the composition and sample size of scorers, and the criteria used for discrimination. Answering this question would require a study into how an increasing number of experts participating in the CHNRI exercise introduces variation in the data set across different exercises; then, how does the number of research questions in the data set introduce variation; how does the substance (ie, content, plausibility) of research ideas introduce variation; and how does the level of agreement between all experts participating in the CHNRI exercise introduce further variation. It would be important to understand whether the key determinant of variation in the data set is the number of experts, the diversity of experts, the number of research ideas, or the content and diversity of research ideas. This could be understood if the number of research ideas and the number of experts are standardized (ie, made equal) across several different CHNRI exercises and then the rank correlation analysis and a comparison of the concordance of the top 20 research priorities are repeated using the methodology in this paper.

An important question is whether by increasing the sample size of scorers we would obtain a wider spectrum of opinions, and therefore greater variation between responses, or whether we would simply continue to observe the same level of variation. One way of addressing this would be to look at a CHNRI exercise where we could separate those who responded to the initial request and those who only responded after reminders, and study whether there was evidence that the late responders differed from the early responders in their opinions.

A search for the presence of sub–stratification in this study could only examine the two characteristics that were known

for each invited scorer: a background in research vs "non–researchers", and affiliation to HIC vs LMIC. When the analysis of concordance was conducted, a reduced level of agreement was detectable when HIC–based vs LMIC–based samples were compared. This observation lends support to the recommendation that an inclusive approach to the sample selection in the CHNRI method should be preferred, so that the result of the exercise reflects the opinion of a wide group of experts. This should help to prevent any particular sub–group among the scorers, with particular views, having undue influence on the results. An analysis of a much larger set of data set from the CHNRI exercises might help to suggest how best to manage the problem of sub–stratification within the sample of invited experts and whether there were examples of exercises in which this concern was reduced to a minimum, or even avoided [17].

Finally, it is of interest to the field of qualitative research to draw analogies between the observations on "saturation" of quantitatively expressed human collective opinion, which we observed in this study, and the long–term notion of quick saturation of information content obtained through interviews with human subjects. Researchers studying the question of the "saturation of ideas" in qualitative research often conclude that 15 interviews may be all it takes to reach a very high degree of "saturation", with 20–30 interviews being sufficient [18]. The numbers as small as those proposed are often counter–intuitive to researchers who conduct quantitative research in the fields such as epidemiology, public health and/or clinical trials, where new information is still discovered even after hundreds or thousands of participants have been enrolled, and having larger sample sizes often leads to a better study with more statistical power to demonstrate convincing results. We conclude that the results of our study seem to support the notion that human collective opinion tends to saturate surprisingly quickly and there does seem to be a point at which adding further experts is unlikely to significantly affect the results that were derived from the initial 45–55 experts. This interesting finding warrants further exploration to understand why this seems to be the case and whether there is a wider significance of this finding, or perhaps any immediate opportunities to implement it in solving practical problems in different areas of human activity.



Acknowledgements: The authors thank the participants of the CHNRI exercise on newborn health for providing scores that were used as a basis for analyses in this study. This work was supported by the CHNRI Foundation award to the University of Edinburgh.

Funding: None.

Ethical approval: Not required.

Authorship declaration: SY led the work on database development and co–wrote the paper. IR participated in study design, co-wrote the paper and provided important intellectual content. SC led the design of the analyses, wrote the script for analyses, mentored SY in running the analyses and provided important intellectual content.

Declaration of interest: IR is an editor–in–chief of the Journal of Global Health. To ensure that any possible conflict of interest relevant to the journal has been addressed, this article was reviewed according to best practice guidelines of international editorial organizations. The authors completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). Authors declare no conflicting financial or other interest related to the work detailed in this manuscript.

REFERENCES

1 Surowiecki J. The wisdom of crowds. London: Abacus, 2013.
2 Surowiecki J. Democracy: dreams of the common good. In: Surowiecki J. The wisdom of crowds. London: Abacus, 2013.
3 Graham B. The intelligent investor. New York: Harper Collins, 2003.
4 Tripadvisor. Available at: https://www.tripadvisor.co.uk/. Accessed: 15 April 2016.
5 Internet Movie Database. Available at: http://www.imdb.com/. Accessed: 15 April 2016.
6 Rudan I, Chopra M, Kapiriri L, Gibson J, Lansang MA, Carneiro I, et al. Setting priorities in global child health research investments: universal challenges and conceptual framework. Croat Med J. 2008;49:307-17. Medline:18581609 doi:10.3325/cmj.2008.3.307
7 Rudan I, Gibson JL, Ameratunga S, El Arifeen S, Bhutta ZA, Black M, et al. Setting priorities in global child health research investments: guidelines for implementation of the CHNRI Method. Croat Med J. 2008;49:720-33. Medline:19090596 doi:10.3325/cmj.2008.49.720
8 Rudan I, Yoshida S, Wazny K, Chan KY, Cousens S. Setting health research priorities using the CHNRI method: V. Quantitative properties of human collective knowledge. J Glob Health. 2016;6:010502 doi:10.7189/jogh.06.010502.

REFERENCES

9 Prediction market. Available: https://en.wikipedia.org/wiki/Prediction_market. Accessed: 15 April 2016.

10 Pennock DM, Lawrence S, Giles CL, Nielsen FA. The real power of artificial markets. Science. 2001;291:987-8. Medline:11232583 doi:10.1126/science.291.5506.987

11 Yoshida S, Martines J, Lawn JE, Wall S, Souza JP, Rudan I, et al. Setting research priorities to improve global newborn health and prevent stillbirths by 2025. J Glob Health. 2016;6:010508. Medline:26401272 doi:10.7189/jogh.06.010508

12 Glaser BG, Strauss AL. The discovery of grounded theory: Strategies for qualitative research. Chicago: Aldine Publishing Co, 1967.

13 Bertaux D. From the life-history approach to the transformation of sociological practice. In: Bertaux D, ed.: Biography and society: The life history approach in the social sciences. London: Sage, 1981.

14 Bernard HR. Social research methods. Thousand Oaks: Sage, 2000.

15 Akins RB, Tolson H, Cole BR. Stability of response characteristics of a Delphi panel: application of bootstrap data expansion. BMC Med Res Methodol. 2005;5:37. Medline:16321161 doi:10.1186/1471-2288-5-37

16 World Bank country data. Available: http://data.worldbank.org/country. Accessed: 15 April 2016.

17 Rudan I. Global health research priorities: mobilizing the developing world. Public Health. 2012;126:237-40. Medline:22325672 doi:10.1016/j.puhe.2011.12.001

18 Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. Field Methods. 2006;18:59-82. doi:10.1177/1525822X05279903

Electronic supplementary material:
The online version of this article contains supplementary material.

journal of global health

Assessing global, regional, national and sub–national capacity for public health research: a bibliometric analysis of the Web of Science™ in 1996–2010

Anna Badenhorst[1], Parisa Mansoori[1], Kit Yee Chan[1,2]

[1] Centre for Global Health Research and WHO Collaborating Centre for Population Health Research and Training, The Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Scotland, UK

[2] Nossal Institute for Global Health, University of Melbourne, Victoria, Australia

Background The past two decades have seen a large increase in investment in global public health research. There is a need for increased coordination and accountability, particularly in understanding where funding is being allocated and who has capacity to perform research. In this paper, we aim to assess global, regional, national and sub–national capacity for public health research and how it is changing over time in different parts of the world.

Methods To allow comparisons of regions, countries and universities/research institutes over time, we relied on Web of Science™ database and used Hirsch (h) index based on 5–year–periods (h5). We defined articles relevant to public health research with 98% specificity using the combination of search terms relevant to public health, epidemiology or meta–analysis. Based on those selected papers, we computed h5 for each country of the world and their main universities/research institutes for these 5–year time periods: 1996–2000, 2001–2005 and 2006–2010. We computed h5 with a 3–year–window after each time period, to allow citations from more recent years to accumulate. Among the papers contributing to h5–core, we explored a topic/disease under investigation, "instrument" of health research used (eg, descriptive, discovery, development or delivery research); and universities/research institutes contributing to h5–core.

Results Globally, the majority of public health research has been conducted in North America and Europe, but other regions (particularly Eastern Mediterranean and South–East Asia) are showing greater improvement rate and are rapidly gaining capacity. Moreover, several African nations performed particularly well when their research output is adjusted by their gross domestic product (GDP). In the regions gaining capacity, universities are contributing more substantially to the h–core publications than other research institutions. In all regions of the world, the topics of articles in h–core are shifting from communicable to non–communicable diseases (NCDs). There is also a trend of reduction in "discovery" research and increase in "delivery" research.

Conclusion Funding agencies and research policy makers should recognise nations where public health research capacity is increasing. These countries are worthy of increased investment in order to further increase the production of high quality local research and continue to develop their research capacity. Similarly, universities that contribute substantially to national research capacity should be recognised and supported. Biomedical journals should also take notice to ensure equity in peer–review process and provide researchers from all countries an equal opportunity to publish high–quality research and reduce financial barriers to accessing these journals.

Correspondence to:
Dr Kit Yee Chan
Centre for Global Health Research
The University of Edinburgh Medical School
Teviot Place, Edinburgh EH8 9A
Scotland, UK
k.chan@ed.ac.uk

Investment in global public health research and development has seen a huge increase in recent years. Funding for health research increased from US$ 50 billion in 1993 to US$ 240 billion in 2009 [1], whilst financial contributions to international Development Assistance for Health (DAH) increased from US$ 5.6 billion to US$ 28.1 billion between 1990 and 2012 [1]. These substantial increases in funding have coincided with a "paradigm shift" from "*International Health*" to "*Global Health*", which occurred over the past two decades. "*International Health*" had its focus on national public health efforts to assist poorer countries [2]. However, "*Global Health*" centres its attention on "*collaborative transnational research and action for promoting health for all*" [3]. This shift provoked recognition that collaborative global action was required to tackle new and evolving health issues, such as SARS, pandemic flu, Ebola, re–emergence of tuberculosis or increase in antibiotic resistance. Additional concerns were raised over the rapidly increasing burden of non–communicable diseases (NCDs) and the need to address health inequities within and between countries [4].

The landscape of global health changed, too, with the World Health Organization and specific countries no longer being seen as the only relevant actors in global health, and with hundreds of organisations now funding global health in an increasingly complex and fragmented manner [5,6]. Whilst the increase in available funding opens up new realms of possibility within global public health research, there is a demand for increased coordination. There were a number of attempts to track and monitor the funding for health research [1,7–10], yet their estimates are strikingly varied, revealing methodological challenges in categorising how the money is spent. To ensure that funding for global health research is being efficiently used, it is necessary not only to understand what is being supported, but also how the funding allocation relates to national and institutional capacity for global health research. Locations with improved capacity for research that are being under–utilised should be identified. As an example, it has been shown that the BRICS nations (Brazil, Russia, India, China and South Africa) made a considerable academic progress in the 21st century: between 2002 and 2007, India doubled the number of original health research papers they produced from 4494 to 9066 [11]; whilst Elsevier (2013) reported that these emerging nations, particularly China, were beginning to overpower the traditional stalwarts such as the UK and USA through the volume of research they are producing. However, it is not only the quantity but also the quality of research, which is improving [12]. To our knowledge, no comprehensive evaluation of the capacity for global public health research has been conducted and the changes in this capacity explored.

In trying to map the capacity, several tools may be utilised. Bibliometric tools allow an evaluation of research productivity, quality, visibility and/or impact at an individual to global level, and therefore can provide a measure of capacity for research. They present objective evidence to describe current research trends and development. The most used bibliometric tools, their advantages and limitations are outlined below. The aim of this study is to assess global capacity for public health research and progression of changes in this capacity over time. In order to achieve this aim, the following objectives must be met:

1. To develop a new scientometric approach, based on *h*–index, which allows an assessment of research characteristics of institutions, countries and regions and their comparison over time;
2. To perform a bibliometric analysis of global public health research based on *h*–index, which is calculated by the Web of Science™;
3. To identify countries and Universities that are improving their capacity for public health research, and those that are stagnating or lagging behind;
4. To identify the research topics of interest within global public health, and their trends over time.

METHODS

Definition of geographic regions and countries included in this study

The countries within each region were defined using the six World Health Organisation's regions [13]. Two of the WHO regions were further subdivided, resulting in a total of 8 separate regions. This was done in order to allow a more comprehensive representation of LMIC and the BRICS nations. The additional regional groupings were created by further dividing the Americas and West Pacific WHO regions into Americas I and II, and West Pacific I and II [14]. A total of 193 countries were included in the analysis. The countries included are shown by region in **Online Supplementary Document**. As the country list was taken from the WHO, disputed countries or territories were not analysed, including Kosovo and Taiwan. The countries that had merged, separated or changed their status or names between 1996 and 2010 were only analysed using their current name (in 2015). Wherever possible, countries with names that have different formats, spelling or abbreviations were identified and all formats of the name used in the search. Due to address restrictions on WoS, publication and citation data from Sudan and South Sudan was aggregated and presented as Sudan and considered in the Eastern Mediterranean Region (EMR). The UK was presented as a single statistical entity, combining England, Scotland, Wales and Northern Ireland.

Definition of time periods

The h–indices, calculated by the Web of Science™, were investigated over three time periods, each of five years:

1996–2000, 2001–2005 and 2006–2010. Five–year periods were chosen to reduce year–to–year stochastic variation within countries. To accommodate for the expected lag between publications and citations, a "citation window" of an additional 3 years following each 5–year period was allowed. This means that, eg, when calculating the h–index for the 5–year time period spanning between 1996–2000, publications with dates 1996–2000 were included, but all citations attributed to those publications in the period 1996–2003 were taken into account in calculation of h–index. This also attenuated the concern related to the temporal nature of the h–index, where older publications would have had a longer time period within which they would have attracted citations.

Search of the literature

After considering the information obtained through the literature review using several available databases (eg, Scopus, Google Scholar and Web of Science), and examining the strengths and weaknesses of each database, Web of Science™ (WoS) was chosen as the database used for this bibliometric analysis. The WoS "Core Collection" was used to ensure that only the publications in the journals with regularly assessed quality are considered.

Given that "public health" is not available as a specific category of articles within WoS, and given that alternative pre–defined categories available in the WoS have serious limitations, it was necessary to devise a search strategy that would efficiently identify public health research to enable an assessment of global, regional, national and sub–national capacity for such research. The search strategy needed to allow an evaluation that would be fair to all countries and allow their meaningful comparison. Public health research can include a multitude of topics, but we chose three search terms as highly specific "indicators" of public health research, as opposed to other types of health research. Those were "epidemiology", "public health" or "meta–analysis". The first two are clear indicators of public health research, whilst meta–analyses are increasingly being performed in response to a growing need to generate evidence for health policy. Although we could have arguably included the term "systematic reviews", we felt that the more rigorous methodology that underlies a meta–analysis process would be a better indicator of research capacity. The search was automated so that all the papers that had any of the words "public health OR epidemiology OR meta–analysis" anywhere in the article were identified by Web of Science™. There were no follow–up steps to this search and all subsequent analyses were then performed on the identified sample of studies.

We validated this approach through studying all 2654 articles that contributed to any of the regional h–indices throughout any of the three time periods (and formed a sub–sample of about 1% of all retrieved articles). One researcher (AB) read the title and the abstract to verify if the article was indeed related to public health. Among those, 58 articles were not related to public health topics, and there were no ambiguities – most of them were meta–analyses related to environmental sciences. This meant that our chosen approach showed about 98% specificity in finding the articles in h–core that are relevant to public health. Whereas the sensitivity of our approach would be very difficult to estimate, the high level of specificity was very encouraging.

Categorisation of papers by type of research and topics of research

To analyse the types of research and the topics of interest that were studied globally over the three time periods, the abstract of each publication contributing to the h–core was reviewed and the publication was categorized using a number of criteria. In terms of topics, papers were characterized as being mainly related to the study of non–communicable diseases (NCDs), infectious diseases (ID), other diseases, or predominantly methodological papers. According to instruments (domains) of the research that were used, a conceptual framework proposed by Rudan et al. was used [15], with the 4 categories and the criteria for categorization shown in **Table 1**.

Database development

Once the search was completed, we used the citation report function on WoS to calculate h–indices for each time period and each geographic region and country. To compute h–index as described in our methods above, it was necessary to download all citation data into a Microsoft Excel format and extract the citation data for each individual paper during the chosen time period, while adding the three–

Table 1. Research instruments (domains) in global public health research*

Research domain	Research avenue
"Description": Epidemiological research	Measuring the burden
	Understanding risk factors
	Evaluating the existing interventions
"Delivery": Health policy and systems research	Studying capacity to reduce exposure to proven health risks
	Studying capacity to deliver efficacious interventions
"Development": Improving existing interventions	Research to improve deliverability
	Research to improve affordability
	Research to improve sustainability
"Discovery": Developing novel interventions	Basic research
	Clinical research
	Public health research

* Source: Rudan et al. [15].

year citation window. The sum of the number of citations per year would then be calculated for each publication.

These totals would then be ranked from highest to lowest and numbered accordingly. This allows the h–index to be calculated by reviewing where the highest rank number is greater than or equal to the corresponding number of citations. This process was repeated for each country for each of the three time periods and the results collated.

As WoS only allows data for 500 papers to be downloaded at once, this was a very time–consuming process. For countries producing more than 500 papers in the area of public health during the 5–year period, the citation information needed to be downloaded 500 papers at a time and then collated into a single data set. Furthermore, it was not possible to download the citation data for searches that produce more than 10 000 results. In all such cases, searches were split into years, and the results were further subdivided using marked lists to enable the citation data to be accessed.

Data analysis

Once we collected the relevant citation data from WoS, we recorded the h–index, total number of publications and total number of citations for each country in a separate database. To perform all the planned analyses, the Gross Domestic Product (GDP) for each country was also recorded, using the World Bank's national–level estimates for the year 2010, or as close as possible, and recorded in US$ [16]. If this data was not available from the World Bank, alternative sources with best estimates were used, typically national estimates generated by the countries themselves and reported at the websites of their national governments.

We used the databases described above to rank all countries by their absolute number of publications and h–indices in each time period, to calculate and rank the absolute rate of increase in h–index between the first and last time period (which was only computed for those countries with an h–index in the first time period of ≥10), to rank all countries by their absolute number of publications per GDP for the most recent time period (for those countries whose number of publications ≥30), and by their h–index per GDP for the most recent time period (for those with an h–index of ≥10).

All papers that formed the h–core had the author's address information reviewed and manually recorded for all contributing authors. The papers with multiple contributing authors were counted more than once when the co–authorship was cross–regional and inter–institutional. This was done through manual data extraction. The institutions to which the authors of h–core papers were affiliated to were recorded in a separate Microsoft Excel data set. Institutions were only verified as universities after a Google search was performed to investigate the institution type. The Universities that contributed more than 2 papers to the h–core were considered as making a notable contribution to global public health in their specific research environment.

RESULTS

The results shall initially focus on describing the characteristics of public health research on a global scale, before focusing on the impact, measured using h–indices, within the 8 geographic regions, individual countries and at specific universities. This will be followed by the analysis of the distribution of papers in h–core by research topics and types of research used.

Global level

The total number of papers that could be considered public health research has dramatically increased over the three time periods, from 63 571 (in 1996–2000) to 89 992 (in 2001–2005) and 158 938 (in 2006–2010). This is a 2.5–fold increase (**Figure 1**). Note that these values will slightly differ from the sum of each country's publications because some papers were allocated to more than one country based on authors' affiliations.

Regional level

As the eight regions differ with regard to their productivity in public health research and impact of their research, they shall be considered separately through an in–depth analysis to identify the hubs of research within those regions, as well as the topics of interest. **Figures 2 to 9** provide summary results in the form of a "fact sheet" for each region.

In the first time–period (1996–2000), the most productive region was Europe with 27 688 publications, closely followed by Americas I with 25 951 publications. This pattern

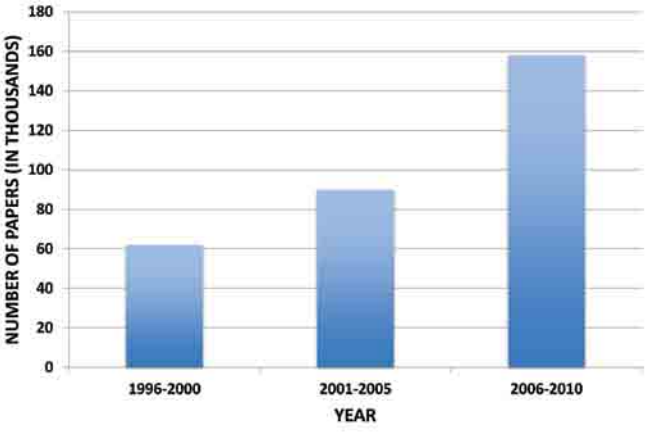


Figure 1. Total number of public health–related publications worldwide over 3 time periods.

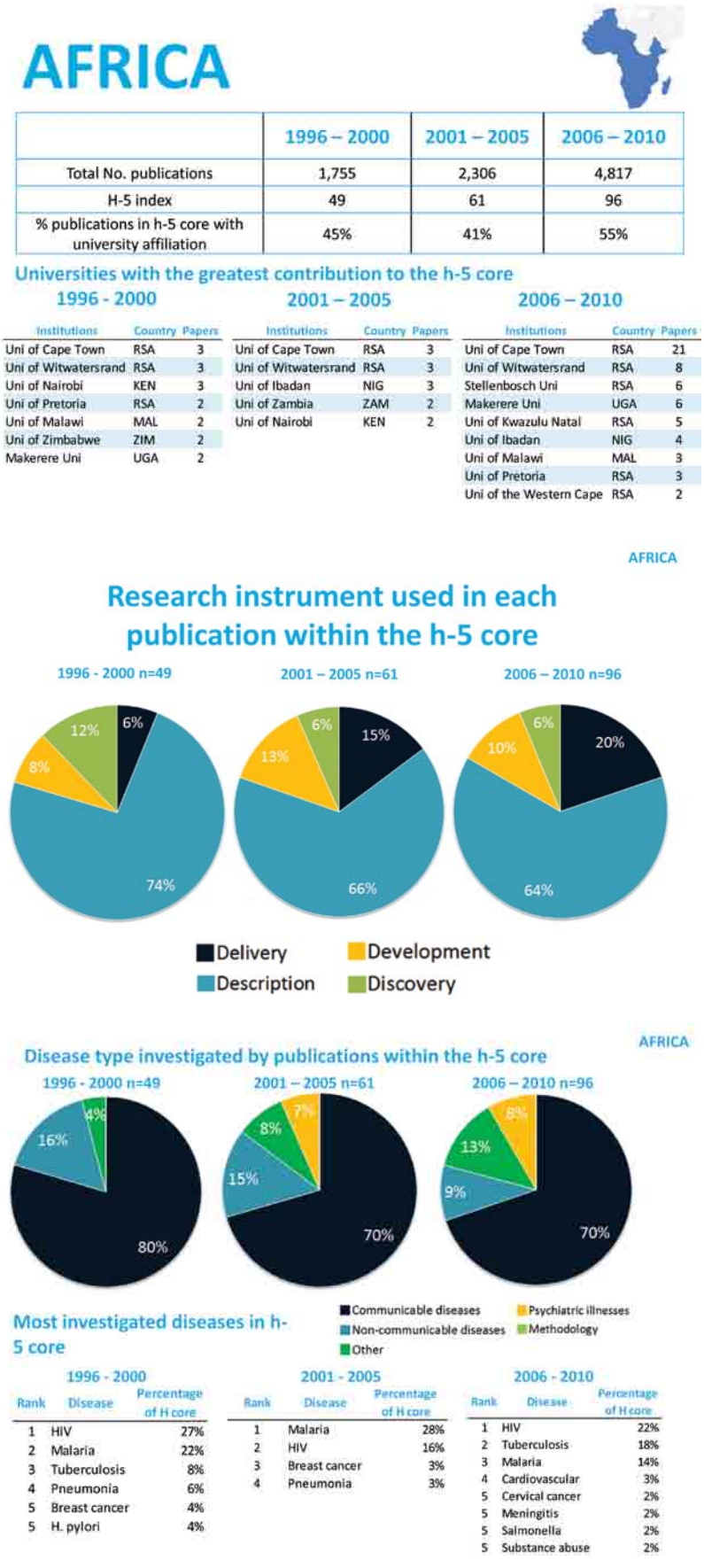


Figure 2. An assessment of capacity to conduct public health research for African region.

is followed in the further two time periods, with Europe and Americas producing 68 260 and 66 933 publications in 2006–2010, respectively. The least productive region in 1996–2000 is the Eastern Mediterranean with 820 publications and the region remain lowest–ranked in 2006–2010 with 3962 publications. However, these regions do not have similar population sizes or number of countries. Therefore, the absolute rate of increase should also be considered in cross–regional comparisons. The region with the largest absolute increase in productivity is West Pacific II. The number of publications in that region increased from 1137 in 1996–2000 to 8837 in 2006–2010, representing an absolute increase of 677%. Europe had the lowest increase in publications during the same period, of 146% (**Table 2**).

The region with the highest h–index throughout all three time periods was Americas I. Their h–index increased from 174 to 300. However, they had the lowest absolute rate of increase in h–index, of 72% (**Table 2**). The Eastern Mediterranean Region (EMR) had the lowest h–index in all three periods – 23 (in 1996–2000), 36 (in 2001–2005) and 70 (in 2006–2010). However, they were also the region with the greatest increase in h–index, by 204%. In every region, the absolute increase in number of publications (productivity) was greater than the increase in h–index (**Table 2**).

National level

The countries were ranked by total number of publications over the three investigated five–year periods. **Figure 10** ranks the top 25 most productive countries over the three time periods. The complete set of results can be seen in **Online Supplementary Document**. The USA dominates by a wide margin, with the UK, Canada, Germany and France consistently ranking in the top five. Of note is the overall improvement in productivity and well as ranking of some the BRICS nations, specifically Brazil and China, with South Africa making an entrance into the top 25 in 2006–2010.

Considering the h–index of individual countries, it can be noted that the overall trend is an increase in h–index over the three time periods. **Figure 11** ranks the top 25 countries with the highest h–indices over the three time

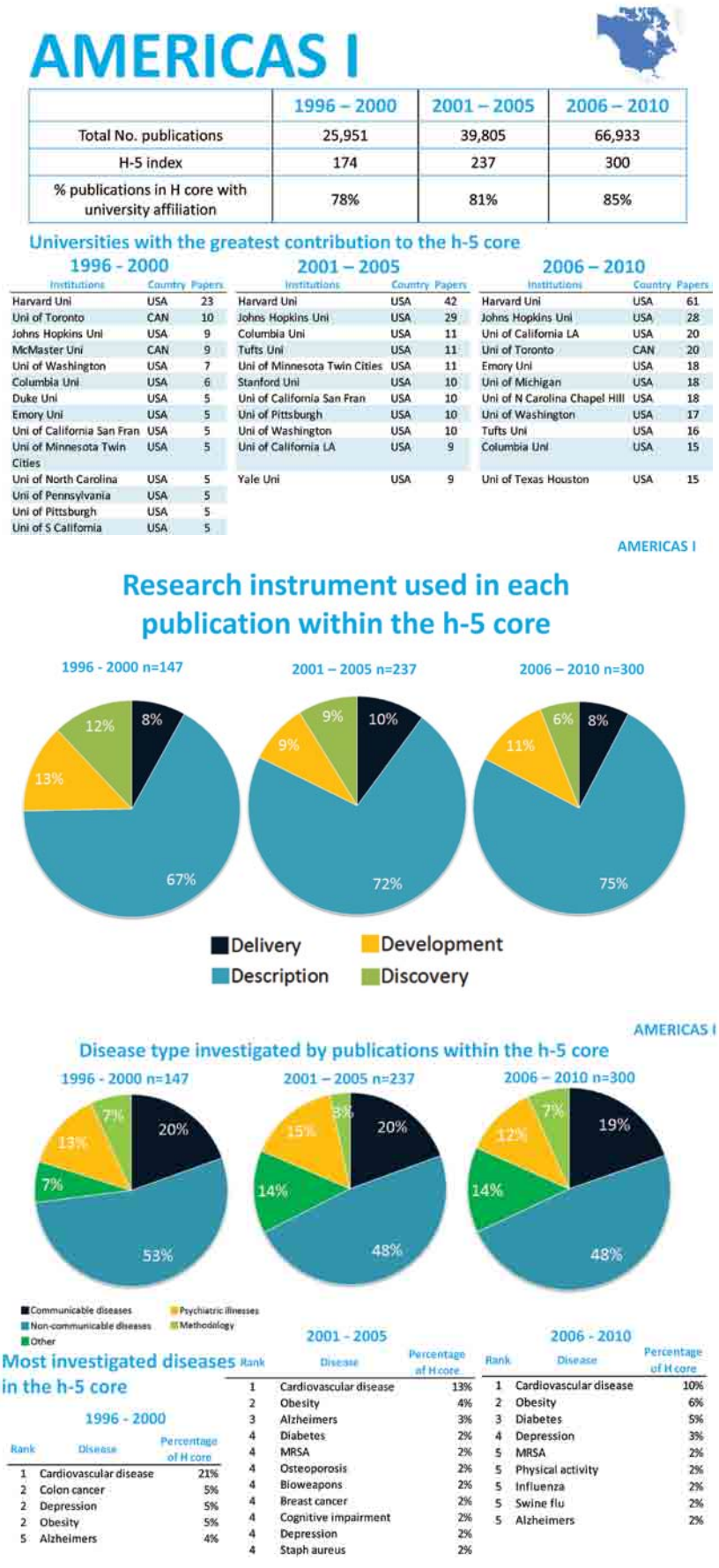


Figure 3. An assessment of capacity to conduct public health research for North–American region.

periods. The complete set of results can be seen in **Online Supplementary Document**. The USA is the leading country on this list, but not to such a degree as in total publication number. Smaller European nations, such as Sweden, Finland and Switzerland, have risen up the ranks based on their h–index although they did not feature as highly in total publication number. The BRICS nations continue to improve, particularly China, Brazil and India, with both an increase in quantity of papers and h–index.

To explore which nations were the most successful throughout the entire study period in improving their capacity for research, the absolute rate of increase for nations with an original h–index greater than 10 between the 1995–2000 hours–index and the 2006–2010 hours–index was calculated. Estonia and Pakistan are at the top of the rankings, with an absolute rate of increase of 230%. In comparison, the USA's rate of increase was 74% and the UK's 106%. The only countries found to have a negative rate of change between the two time periods were Jamaica (with a decrease of 15%) and Guinea–Bissau (with a decline of 20%).

The total number of publications in relation to GDP was considered for the 2006–2010 period. To avoid spurious results, only countries with more than 30 publications were included. The 25 countries that were most productive in relation to their GDP are ranked in **Table 3**, and the full results can be found in **Online Supplementary Document**. African Nations dominate the top 25 ranks, indicating that some of them are being very productive with limited resources, particularly the Gambia – whose GDP is amongst the lowest worldwide.

The h–index was then reviewed in relation to GDP and the list of top 25 countries is shown in **Table 4**. Full results are available in **Online Supplementary Document**. As with the absolute rate of increase, only countries with an h–index greater than 10 in 2006–2010 were considered. The upper ranks are again dominated by African nations, whilst the USA now ranks second to last. The Gambia has again performed particularly well, indicating that they are producing high quality research with limited re-

Table 2. Absolute increase in number of publications and h–indices for each region

Region	Absolute increase in number of publications between 1996–2000 and 2006–2010 (%)	Absolute increase in h–index between 1996–2000 and 2006–2010 (%)
Africa	174	95
Americas I	158	72
Americas II	334	110
East Med	383	204
Europe	146	96
South–East Asia	327	152
West Pacific I	209	108
West Pacific II	677	148

sources. This is clearly a result of the research activity of a well–known international research centre, supported largely by the Medical Research Council in the UK, that was established in the Gambia in the 20th century.

In the period 2006–2010, it was noted that there was a considerable gap between the country at the top of the rankings and all others from the same region in the number of publications and h–index. In the African region, South Africa was at the top (1579 publications and h–index of 77); in Americas I, the USA (59416 publications and h–index of 294); in Americas II, Brazil (6540 publications and h–index of 78); in East Mediterranean, Iran (1326 publications and h–index of 42); in Europe, the UK (publications 18918, h–index 223); in SE. Asia, India (2843 publications; h–index of 72); in West Pacific I, Australia (8025 publications; h–index 143); and in West Pacific II, China (6049 publications; h–index 100). When h–index is considered in relation to GDP, the only country that remains at the top within its own region is India – which has the highest h–index per GDP in SE. Asia. The remaining countries all moved down their regional rankings, because other nations with lower total publications and h–indices perform better in relation to their GDP. Countries which are particularly successful in relation to their GDP are the Gambia, Malawi, Barbados, Nicaragua, Jordan, Lebanon, Iceland, Estonia, Thailand, Laos and Mongolia.

Sub–national level

In general, the percentage of papers in regional h–cores that were originated at a regional university increased throughout the three time periods. The exceptions were SE. Asia and West Pacific II, where the percentages in the first and the last time period were very similar. The region with the greatest university contribution to the regional h–core was Europe, where 89% of h–core publications had authorship from a European university. This was similar in other regions with high–income countries, such as Americas I (with 85%) and West Pacific I (with 82%). However, in

Table 3. Top 25 countries ranked by total number of publications/gross domestic product (GDP) in 2006–2010

Rank	Country	GDP (2010, in US$ billion)	Papers in 2006–2010	Papers per GDP
1	Gambia	0.78	73	93.1
2	Malawi	3.29	143	43.5
3	Uganda	13.36	326	24.4
4	Zimbabwe	5.20	122	23.5
5	Kenya	23.53	482	20.5
6	Burkina Faso	7.11	127	17.9
7	Tanzania	19.72	341	17.3
8	Iceland	16.39	259	15.8
9	Senegal	10.37	159	15.3
10	Nepal	10.10	148	14.6
11	Mongolia	3.45	47	13.6
12	Zambia	9.80	131	13.4
13	Ghana	14.80	197	13.3
14	Estonia	13.90	181	13.0
15	Lao	4.02	52	12.9
16	Benin	5.23	61	11.7
17	Ethiopia	20.40	235	11.5
18	Cambodia	8.69	100	11.5
19	New Zealand	120.04	1361	11.3
20	Croatia	45.87	506	11.0
21	Cameroon	19.21	210	10.9
22	Rwanda	3.79	41	10.8
23	Denmark	256.82	2757	10.7
24	Mozambique	9.13	96	10.5
25	Madagascar	5.76	60	10.4

Table 4. Top 25 countries ranked by h–index/ gross domestic product (GDP) in 2006–2010

Rank	Country	GDP (2010, in US$ billion)	H–index for 2006–2010	H–index per GDP
1	Gambia	0.78	22	28.1
2	Malawi	3.29	29	8.8
3	Lao	4.02	18	4.5
4	Zimbabwe	5.20	23	4.4
5	Rwanda	3.79	15	4.0
6	Fiji	3.03	12	4.0
7	Niger	4.38	17	3.9
8	Burkina Faso	7.11	25	3.5
9	Mali	6.97	23	3.3
10	Iceland	16.39	54	3.3
11	Mongolia	3.45	11	3.2
12	Madagascar	5.76	18	3.1
13	Uganda	13.36	39	2.9
14	Benin	5.23	14	2.7
15	Cambodia	8.69	23	2.6
16	Papua New Guinea	6.55	17	2.6
17	Malta	6.65	17	2.6
18	Barbados	4.03	10	2.5
19	Mozambique	9.13	22	2.4
20	Estonia	13.90	33	2.4
21	Senegal	10.37	24	2.3
22	Nepal	10.10	23	2.3
23	Zambia	9.80	22	2.2
24	Kenya	23.53	52	2.2
25	Gabon	9.68	19	2.0

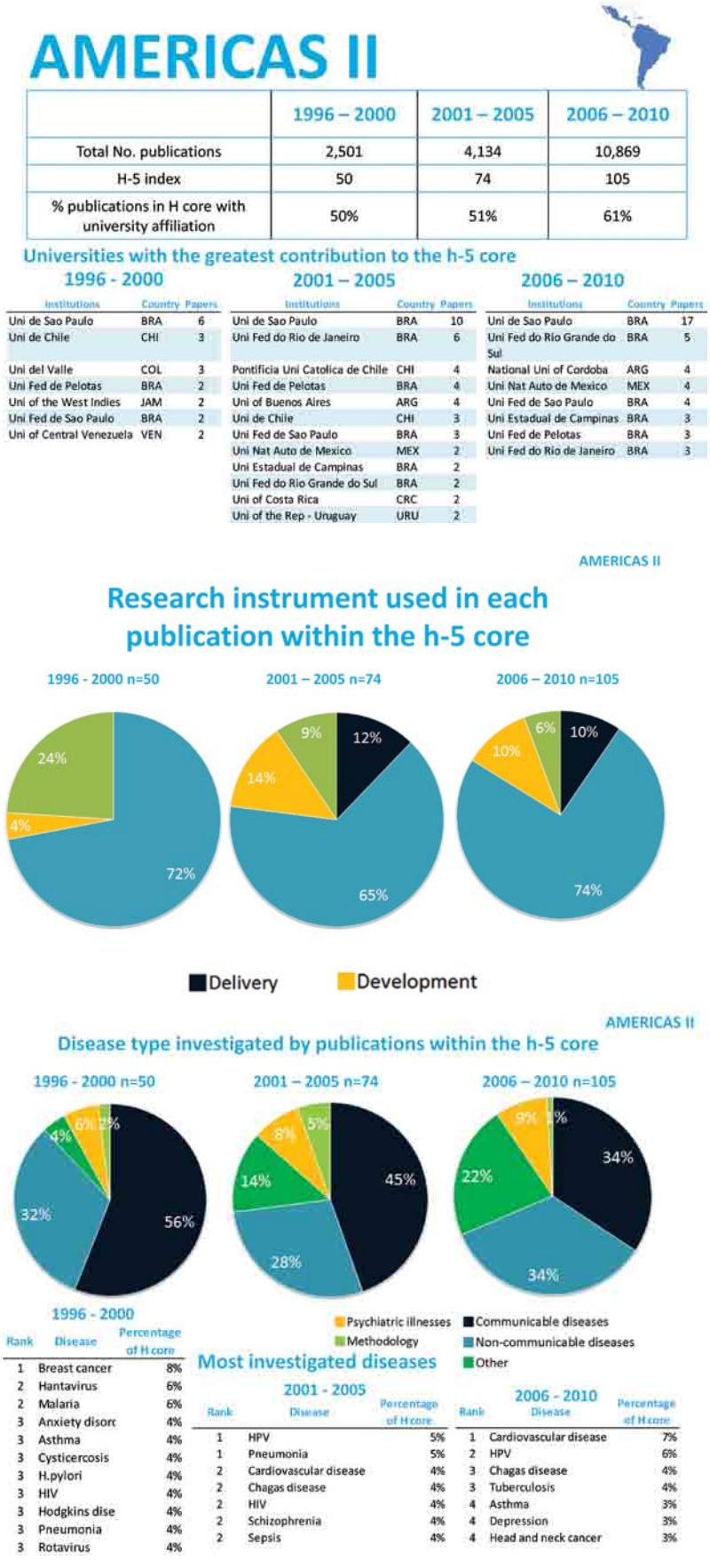


Figure 4. An assessment of capacity to conduct public health research for Latin–American region.

poorer regions, the percentage of papers in the h–core originated from a regional university was lower. In SE. Asia, only 49% of papers were university–based in 2006–2010, and in Africa they contributed to 55%. Some of the leading regional universities are Cape Town's, Harvard, Universidade de Sao Paulo, Oxford, Madihol and Sydney's.

Types of research

The four instruments (or "domains") of health research, as described by Rudan [15], could be summarized as "the four D's": "description", "delivery", "development" and "discovery". We categorized each paper that contributed to the regional h–core in each time period into one of those four domains. The results for each individual region can be seen in **Figures 2 to 9**. In each region, the majority of papers in the h–core were "descriptive" papers – ranging from 64% (in West Pacific I) to 79% (in South–East Asia). In all regions, the proportion of research in the h–core relating to "discovery" research decreased, with the exception of Eastern Mediterranean region (EMR) where it remained stable. There was little change in the proportion of research that related to "development", but in the majority of regions, research on "delivery" in public health increased (the only exceptions being Americas I and West Pacific I).

Topics of research

Each publication that related to a disease in a region's h–core throughout the three time periods was classified into non–communicable diseases (NCDs), infectious diseases (ID), other diseases, or a predominantly methodological papers. In three regions, NCDs were the topic of most interest in the h–core throughout all three time periods: Americas I, Europe and West Pacific I. In two regions, the research interest was mainly focused on infectious diseases throughout all three periods: Africa and South–East Asia. In the remaining three regions (Americas II, Eastern Mediterranean and West Pacific II), a similar pattern can be seen – the proportion of papers relating to communicable diseases is decreasing, and the proportion relating to NCDs is increasing (**Figures 2 to 9**).

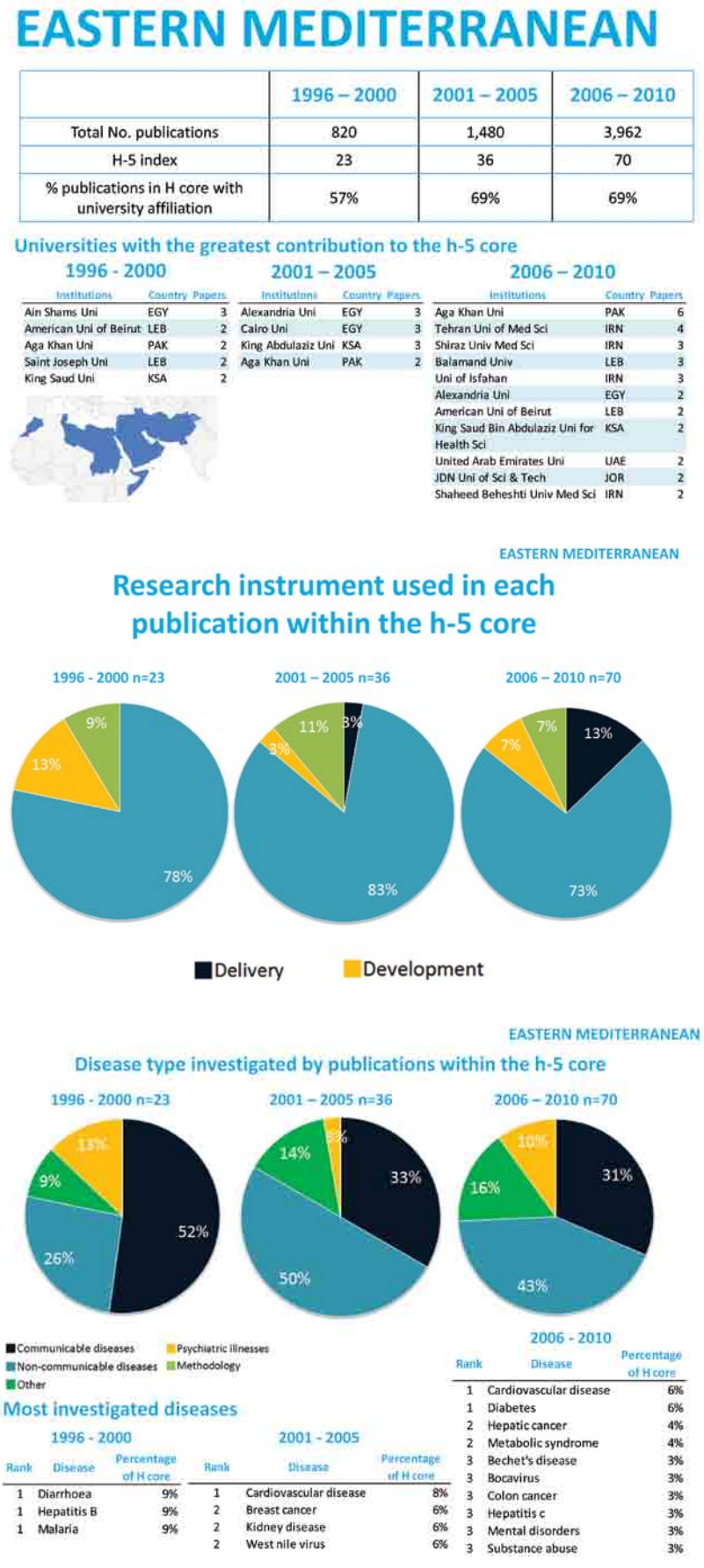


Figure 5. An assessment of capacity to conduct public health research for Eastern Mediterranean region.

The specific diseases under investigation by the publications in the h–core follow a similar pattern to the proportion of topics. Cardiovascular diseases were most frequently studied in high–income regions, and increasing in importance in regions with lower income. Moreover, in high–income regions, diabetes, obesity and depression are increasing in importance. Overall, there is a slight increase in the proportion of papers relating to psychiatric illnesses, with the greatest increase in the West Pacific I. In Europe, papers relating to the methodology of performing public health research are increasing.

DISCUSSION

Increasing investment in global public health research has resulted in a need to understand where capacity to perform research lies. Currently, some areas of the world may not be seen as "worthy" of research investment by some funders. However, there is a lack of an established and effective methodology that can be used to identify the nations and institutions that are demonstrating an improved capacity for public health research globally. This study was successful in developing a new bibliometric approach to address this question, by adapting the h–index to allow research capacity in public health worldwide to be assessed over time. The results clearly highlight countries that improved their capacity for public health research and the institutions that are contributing substantially to public health research. In addition, this study has been successful in providing an understanding of the trends in research instruments ("domains") used and topics that were investigated through public health research.

This study has, therefore, not only established a methodology to assess public health research capacity worldwide, but also provided a baseline to which future evaluations can be compared. In addition, the methodology developed here could be adapted to any other topic of scientific research in order to assess global, regional, national and sub–national capacity for research.

On viewing the total number of publications and h–indices over the three time periods for

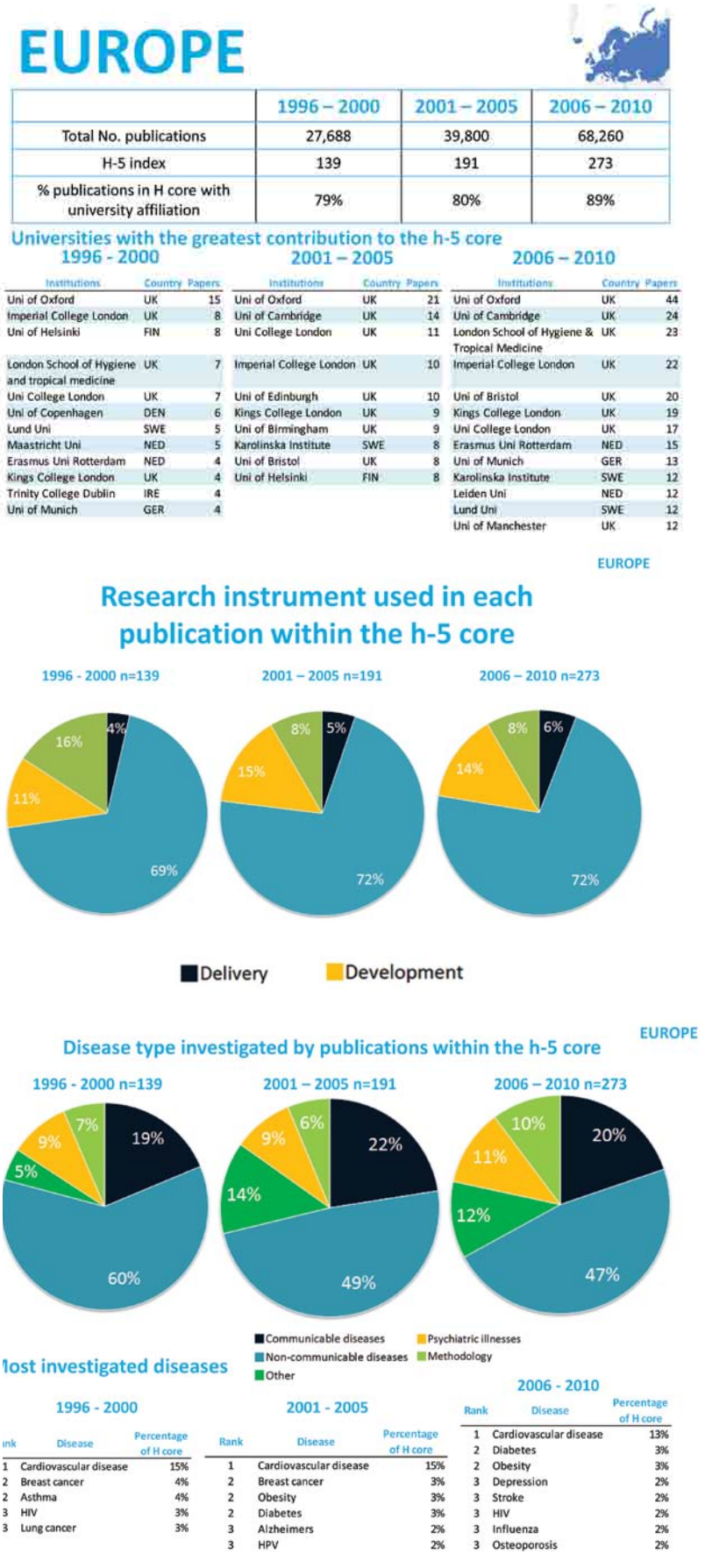


Figure 6. An assessment of capacity to conduct public health research for European region.

each region, it can be seen that the large majority of papers come from European and America I regions. However, despite the large numbers of publications, their absolute increase over time in both number of publications and h–index is relatively low. Other regions, such as Eastern Mediterranean and South–East Asia, are showing a considerable improvement in both publication number and h–index. At the same time, Western Pacific II region has seen a huge absolute increase in publications, but the increase in h–index is not correspondingly high. Africa has a fairly low absolute increase in both publication number and h–index, with low values to start from, too.

The USA clearly dominates in terms of productivity and h–index. However, when GDP is taken into account, the USA actually ranks rather low. In comparison to the UK, which consistently ranks second in terms of both quantity and quality, the USA is producing a huge amount of research, yet their h–index is not correspondingly high. At the same time, the BRICS nations have been making substantial improvements, all of them ranked in the top 25 countries for productivity and h–index in 2006–2010 period, with the exception of Russia. They all had absolute rates of increase in h–index greater than 140% except Russia, whose rate of increase was only 44%. It is possible this could be explained by the frequency at which countries publish in the English language. As reported in the literature review, non–English language journals are less frequently indexed in WoS. Some nations may appear not to be performing well, when in fact it is simply that their research is predominantly published in non–English language journals. This has been reported to be the case for Russia in stroke–related research [17]. However, this could also be the case for many other countries, whose research capacity is being under–represented in this analysis. Furthermore, as non–English language papers are less likely to be cited [18], they may incorrectly appear to be of lower quality whenever citations are used as a partial indicator of research quality.

Despite low numbers of publications and low h–indices in general, African nations can be seen to be performing well, considering

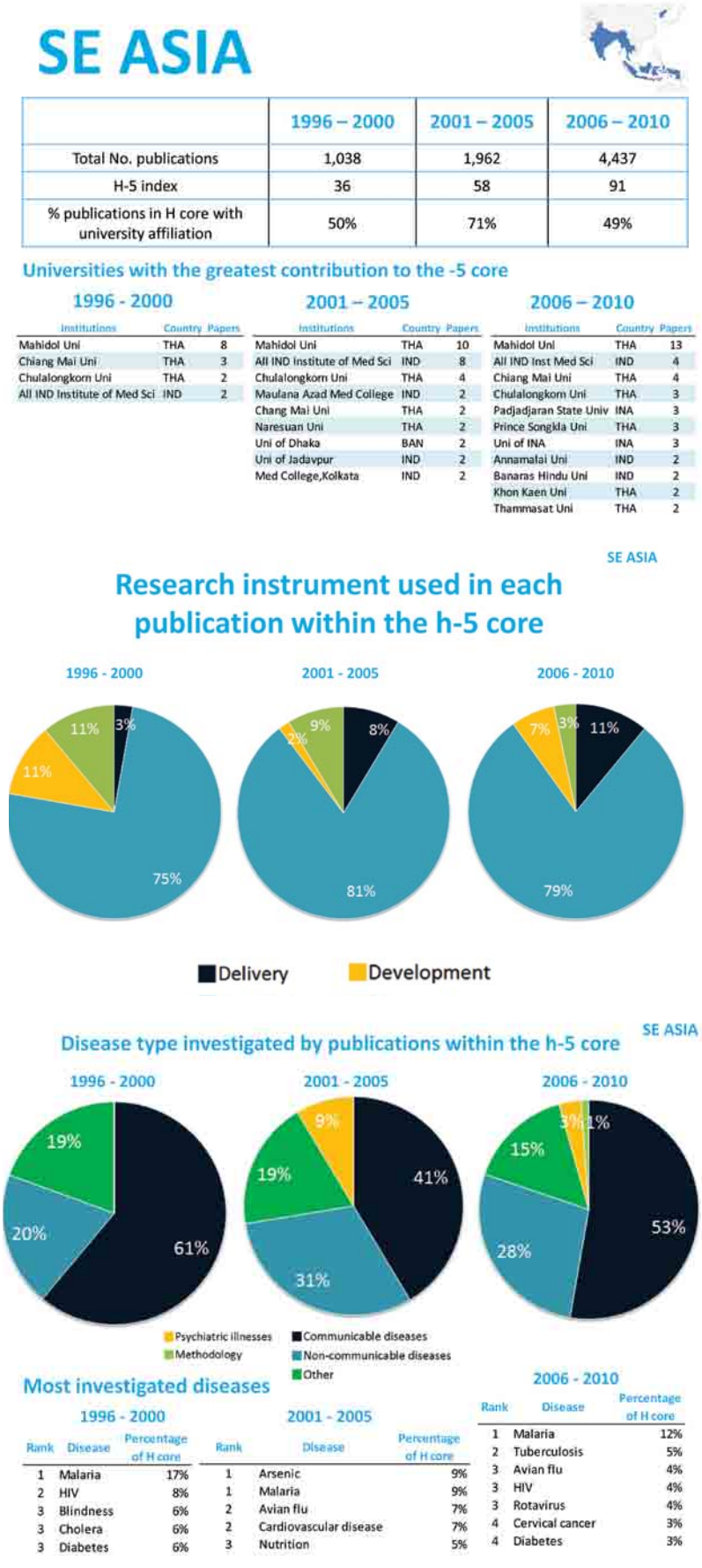


Figure 7. An assessment of capacity to conduct public health research for South–East Asian region.

the resources available (measured by GDP). However, similar to South–East Asia, about half of the papers in the h–core for the region have been produced by non–university institutions. It is, therefore, likely that international research organisations are performing large portion of this regional research, which may inhibit the progress of local universities. For example, in Egypt, the US Navy performed much of the research in the h–core. However, in the majority of regions, the proportion of non–university authored publications in the h–core is declining, suggesting that university–based research is improving in quality almost universally.

On reviewing the research topics that occur in the h–core of the regions, it can be noted that Africa and South–East Asia are the only two regions where communicable diseases remain proportionally the most studied in the 2006–2010 time–period. The Eastern Mediterranean, America II and West Pacific II regions can be seen as transitioning from their historic focus on communicable diseases to NCDs, whilst Europe and America I have a very similar distribution of research throughout. Regarding research instruments ("domains"), both Rudan et al. and Leroy et al. proposed that too much research funding may be allocated to the development of new interventions, which could not be as effective in reducing child mortality as implementing the existing interventions effectively [15,19]. It is, therefore, pleasing to see an increase in research related to delivery of interventions, whilst research relating to novel discoveries is decreasing, thus achieving a more desirable balance. In Africa in particular, research on delivery of public health interventions is increasing in both quality and quantity, demonstrating the capacity in this region to improve implementation of available interventions.

The key strength of our study lies in the methodology developed, which allowed not only an assessment of global public health research capacity, but also the trends over time. This was the first application of this novel methodology, using existing large data sets on WoS in a novel way, allowing the emerging research hubs to be identified and the current research trends to be visualised. The use of the three–year citation window

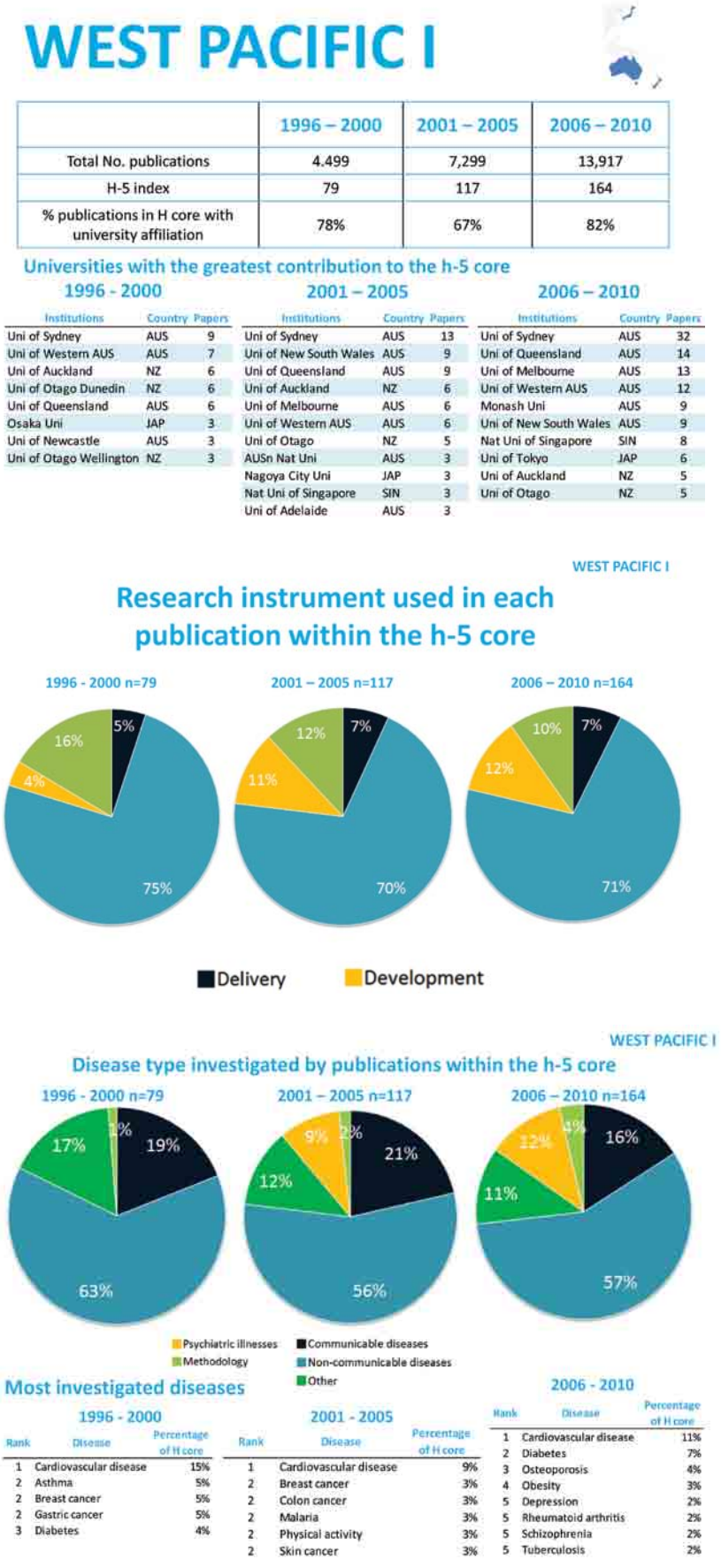


Figure 8. An assessment of capacity to conduct public health research for West Pacific I region.

following each 5–year period ensured that studies towards the end of the time–period had adequate time to be cited. Furthermore, in the validation of 2654 articles that contributed to the regional h–indices throughout the 3 time–periods, 2% were found to not be relevant to public health. This was felt to be an acceptable level of specificity. When considering the possible biases related to sensitivity of the proposed approach to literature search, whilst there are undoubtedly public health papers that remained unidentified using our search strategy, we find it unlikely that this problem could affect the overall results or rankings of nations that we reported here, and which seem plausible to a large extent.

The novel use of the h–index proposed in this study has provided a single measure with which the quality and quantity of research produced by regions, nations and institutions can be compared over time. Whilst the h–index is superior to citations per paper and IF, it does have its limitations. As an example, it does not provide an understanding of the proportion of low quality studies produced by a country or region. In the case of the USA, this could be particularly interesting, as their h–index is very high, yet they have a vast number of publications which do not contribute to it, which is proportionally much greater than other nations. There is also a possible concern about the phenomenon known as the "Matthew effect", where more recognised and established researchers may have their work cited more, simply due to name recognition rather than the true quality of the publication [20]. This would falsely inflate the apparent gap between more established research nations and those that are emerging. In addition, it has been shown that the h–index is higher when there is more international collaboration between nations [21]. As a metric, it therefore disadvantages those LMIC who do not have as much opportunity for collaboration as North America and Europe. This would, again, act to increase apparent inequalities between established and emerging research nations.

As with many bibliometric–type studies, this study has limitations that are inherent in using an online database to access citation

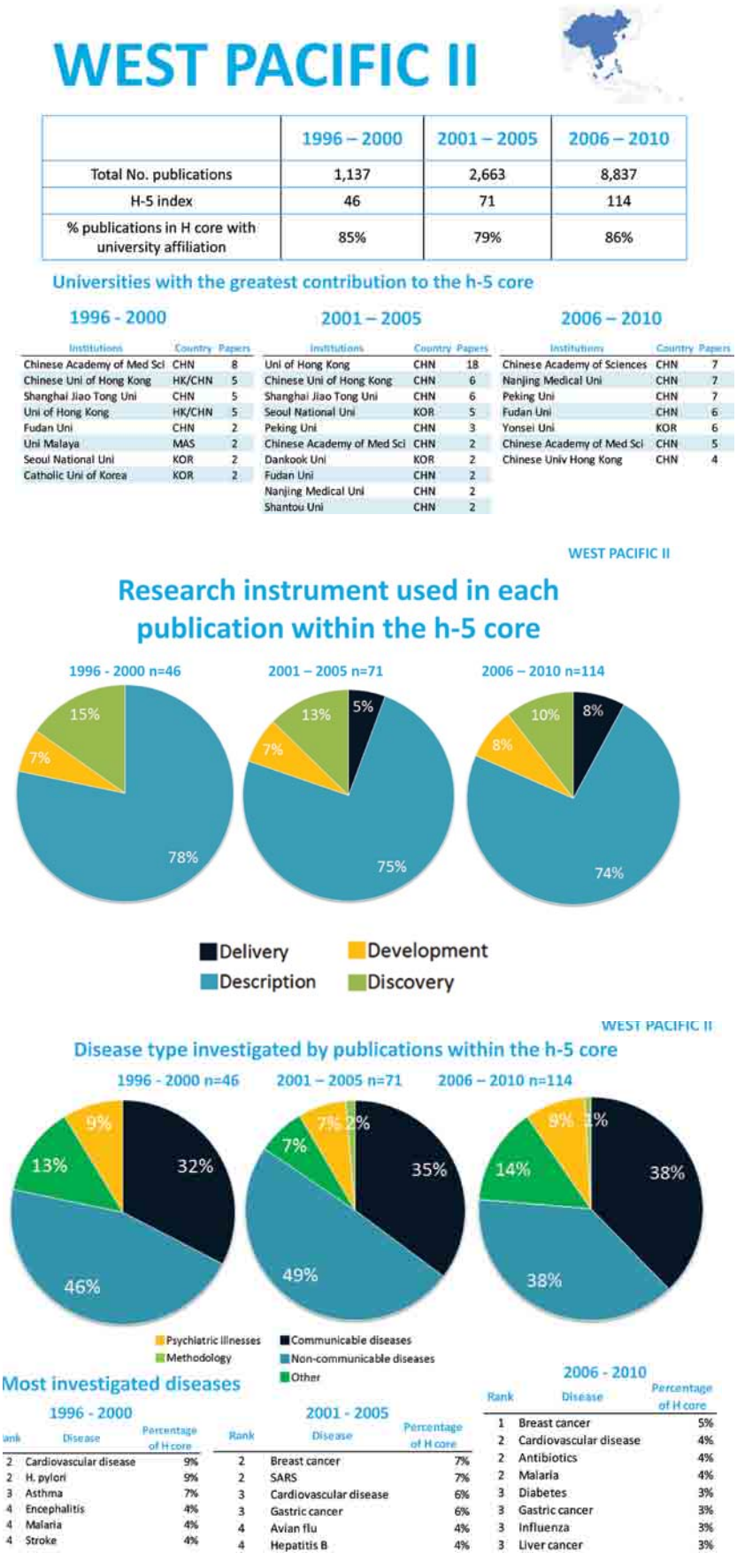
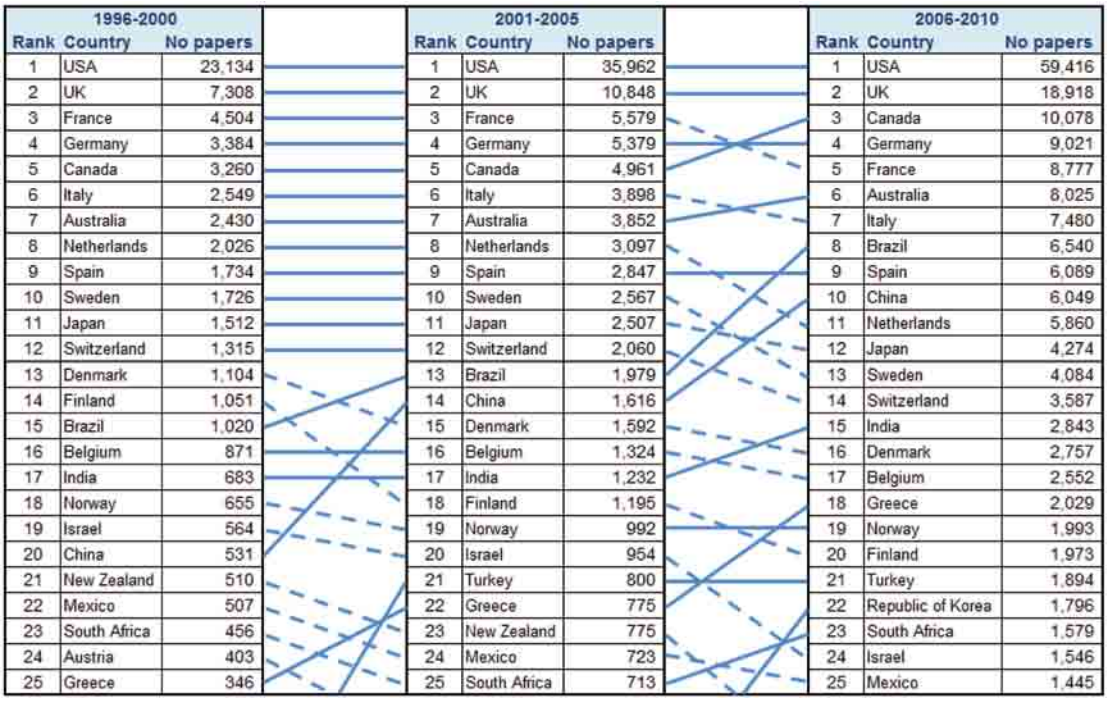


Figure 9. An assessment of capacity to conduct public health research for West Pacific II region.

data. These databases have language bias, with papers and journals not writing in English less likely to be indexed. The result would be fewer publications from emerging research countries, where research is more likely not to be published in English. Another problem was studied by Gingras, who noted that some wealthy institutions from middle–income countries may be able to manipulate their citation numbers by offering highly cited researchers attractive contracts for minimal work if they would agree to affiliate themselves with the paying university as the secondary affiliation. Gingras describes these as "dummy affiliations", with no real impact on teaching and research in universities, allow marginal institutions to boost their position in the rankings of universities without having to develop any real scientific activities [22].

There are also many academics who view the use of citation metrics to measure quality of research as "a terrible idea". Sabaratnam and Kirby wrote a response to the Higher Education Funding Council for England, who were considering using citation metrics when assessing research quality, and received over 200 signatories objecting to the idea [23]. They quite rightly pointed out that a citation is not necessarily an endorsement of quality. They state that all methods currently available to assess quality are flawed. Whilst the h–index is certainly not a perfect measure of research quality or capacity, it seems that it may be the best currently available. The fact that there is not a perfect measurement technique does not mean that no attempt should be made to understand public health research capacity, and identify those who are improving.

Hirsch himself believed that "a single number can never give more than a rough approximation to an individual's multifaceted profile, and many other factors should be considered in combination in evaluating an individual" [24]. It is certainly not possible that a single metric, such as h–index, can truly describe an institution or country's contribution to global public health research. However, this study provides a bibliometric profile of regions, countries and institutions which, when viewed together, can characterise their



1996-2000			2001-2005			2006-2010		
Rank	Country	No papers	Rank	Country	No papers	Rank	Country	No papers
1	USA	23,134	1	USA	35,962	1	USA	59,416
2	UK	7,308	2	UK	10,848	2	UK	18,918
3	France	4,504	3	France	5,579	3	Canada	10,078
4	Germany	3,384	4	Germany	5,379	4	Germany	9,021
5	Canada	3,260	5	Canada	4,961	5	France	8,777
6	Italy	2,549	6	Italy	3,898	6	Australia	8,025
7	Australia	2,430	7	Australia	3,852	7	Italy	7,480
8	Netherlands	2,026	8	Netherlands	3,097	8	Brazil	6,540
9	Spain	1,734	9	Spain	2,847	9	Spain	6,089
10	Sweden	1,726	10	Sweden	2,567	10	China	6,049
11	Japan	1,512	11	Japan	2,507	11	Netherlands	5,860
12	Switzerland	1,315	12	Switzerland	2,060	12	Japan	4,274
13	Denmark	1,104	13	Brazil	1,979	13	Sweden	4,084
14	Finland	1,051	14	China	1,616	14	Switzerland	3,587
15	Brazil	1,020	15	Denmark	1,592	15	India	2,843
16	Belgium	871	16	Belgium	1,324	16	Denmark	2,757
17	India	683	17	India	1,232	17	Belgium	2,552
18	Norway	655	18	Finland	1,195	18	Greece	2,029
19	Israel	564	19	Norway	992	19	Norway	1,993
20	China	531	20	Israel	954	20	Finland	1,973
21	New Zealand	510	21	Turkey	800	21	Turkey	1,894
22	Mexico	507	22	Greece	775	22	Republic of Korea	1,796
23	South Africa	456	23	New Zealand	775	23	South Africa	1,579
24	Austria	403	24	Mexico	723	24	Israel	1,546
25	Greece	346	25	South Africa	713	25	Mexico	1,445

Figure 10. Top 25 countries ranked by total number of publications in each time period. Continuous blue lines indicate improvement in rank between the two periods or no change in rank. Dashed blue lines indicate decrease in rank between the two periods.

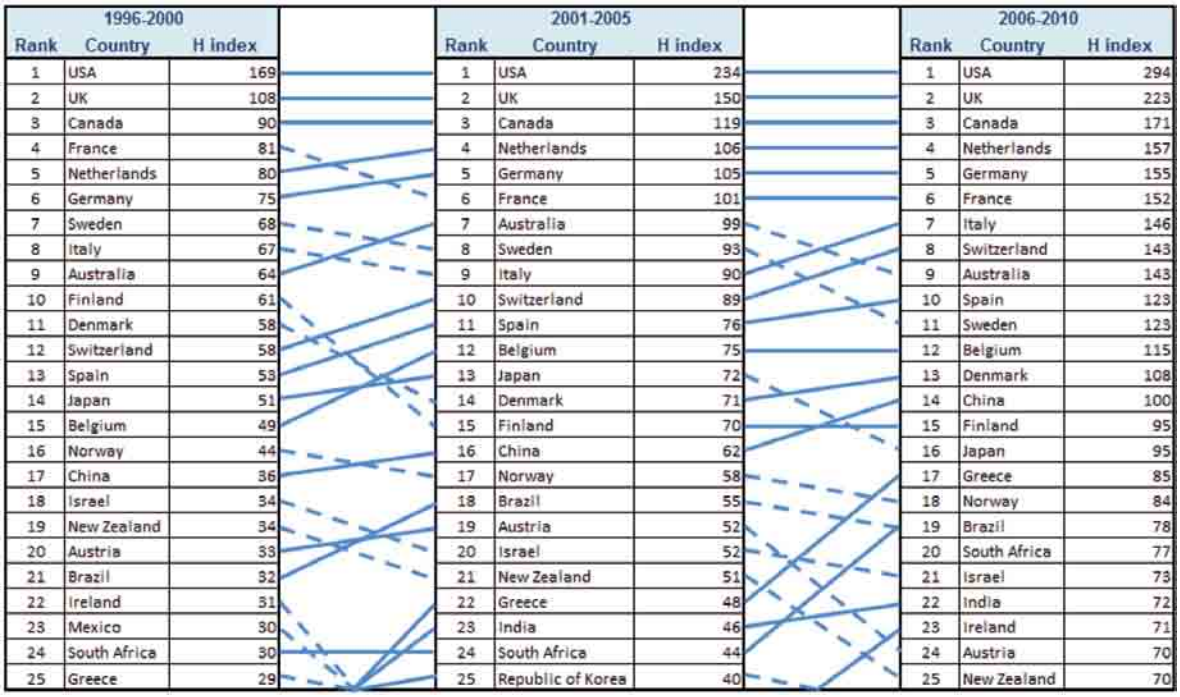
publication and research efforts and provide an indication of their capacity to perform public health research. Despite the limitations of bibliometric research, this study has been successful in identifying nations in each region which have capacity for public health research, which are improving and which are performing well despite limited resources.

Many of the nations seem to be improving both the quality and the quantity of their public health research with comparatively limited resources. Whilst some of these countries were expected to be making improvements, based on their rapid economic development (such as Brazil, South Africa, China and India), there have also been other unexpected nations demonstrating great capacity for public health research. Some, like Estonia and Pakistan, have made huge strides in improving their research quality and quantity. Others, like the Gambia, Malawi and Laos are producing high quality research despite extremely limited domestic resources. In addition, those universities which are contributing substantially to national research capacity should be recognised and supported.

We mentioned in the introduction section that the use of a country's GDP for expenditure on health research is a proxy, as there is no other reliable method to track such expenditures. In light of this knowledge, social and political differences (such as war, conflict, or financial instability) between countries or regions might also make it a challenge in figuring out how governments spend money on health research [25].

In the future, public health research shall likely become increasingly specialized, which may result in cutting–edge research becoming more expensive and based on large–scale "biobanks". Therefore, identifying universities that perform well in all regions and increasing international communication and cooperation will be beneficial to the global public health research community. In many of the low–income countries, there is also a discrepancy between their current disease burden and the ability to perform public health research. Their universities should further focus on studying delivery of the existing public health interventions, to allow evidence–based decisions to be made based on locally relevant research. Increasing collaboration between LMICs and forming so–called "South–South partnerships" to address common health problems would also be beneficial, with a focus on those diseases that contribute significantly to national disease burdens, such as diabetes and cardiovascular disease. Ranasinghe argued that researchers in LMIC face additional challenges when attempting to publish their research, which is largely due to language and funding issues [26]. Therefore, medical journals should be encouraged to provide researchers throughout the world with equal opportunity to publish their research, and offer guidance how to improve its quality.

In the future, this study should be repeated at five–yearly intervals to identify new and emerging hubs of public health research. In order for future studies to be completed more efficiently, there are a number of steps that Web of



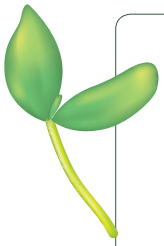
1996-2000			2001-2005			2006-2010		
Rank	Country	H index	Rank	Country	H index	Rank	Country	H index
1	USA	169	1	USA	234	1	USA	294
2	UK	108	2	UK	150	2	UK	223
3	Canada	90	3	Canada	119	3	Canada	171
4	France	81	4	Netherlands	106	4	Netherlands	157
5	Netherlands	80	5	Germany	105	5	Germany	155
6	Germany	75	6	France	101	6	France	152
7	Sweden	68	7	Australia	99	7	Italy	146
8	Italy	67	8	Sweden	93	8	Switzerland	143
9	Australia	64	9	Italy	90	9	Australia	143
10	Finland	61	10	Switzerland	89	10	Spain	123
11	Denmark	58	11	Spain	76	11	Sweden	123
12	Switzerland	58	12	Belgium	75	12	Belgium	115
13	Spain	53	13	Japan	72	13	Denmark	108
14	Japan	51	14	Denmark	71	14	China	100
15	Belgium	49	15	Finland	70	15	Finland	95
16	Norway	44	16	China	62	16	Japan	95
17	China	36	17	Norway	58	17	Greece	85
18	Israel	34	18	Brazil	55	18	Norway	84
19	New Zealand	34	19	Austria	52	19	Brazil	78
20	Austria	33	20	Israel	52	20	South Africa	77
21	Brazil	32	21	New Zealand	51	21	Israel	73
22	Ireland	31	22	Greece	48	22	India	72
23	Mexico	30	23	India	46	23	Ireland	71
24	South Africa	30	24	South Africa	44	24	Austria	70
25	Greece	29	25	Republic of Korea	40	25	New Zealand	70

Figure 11. Top 25 countries ranked by h–index in each time period. Continuous blue lines indicate improvement in rank between the two periods or no change in rank. Dashed blue lines indicate decrease in rank between the two periods.

Science™ (WoS) itself could take to make the process more streamlined. It would be very beneficial to allow citation data to be collected for those searches which have >10 000 results. As the quality of research continues to grow, there will soon be many countries who produce >10 000 public health publications in a 5–year period. In addition, removing the cap, which only allows the citation data of 500 publications to be downloaded at a time, would be helpful. As some countries have over 50 000 publications to be analysed, collating all these results is extremely time consuming and could easily be avoiding by some simple adjustments by WoS. This methodology could also be extended to other fields of science, allowing them to assess the development of research capacity worldwide. However, it should be remembered that the evaluations of different fields based on h–indices are often not comparable, primarily due to large differences in the number of participating researchers and an overall number of citations.

CONCLUSION

This is an exciting time for public health research. The potential funding available for research is larger than ever, allowing the quantity of research to increase, and the quality to improve. However, there is a danger that funding will continue to be allocated mainly to established and traditional "hubs" of research. In recent years, many nations, particularly LMIC, have been improving their research quantity and quality – thereby gaining capacity for public health research. This study was successful in developing a methodology, based on the h–index, which provides an assessment of capacity for public health research from 1996–2010. As expected, the USA and UK dominated public health research globally. However, there were a number of countries with limited resources, demonstrating improved capacity for public health research. In addition, university contributions to high quality research were increasing. There has been a shift in research domains – with more research on improving deliverability of existing interventions. The research being performed is also more representative of the burden of disease worldwide, with a shift towards NCDs. In order to improve the overall quality of public health research, international collaborations should be encouraged, while medical journals should seek to ensure that publication is a fair and equitable process.



Acknowledgements: The authors thank Professor Igor Rudan on providing the idea for this paper.

Funding: None.

Ethical approval: Not required.

Authorship declaration: AB conducted the review of the relevant literature, developed the framework and wrote the paper. PM and KYC double–checked all results, provided important intellectual content and prepared the final version of the paper for publication.

Declaration of Competing Interest: The authors completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). Authors declare no conflicting financial or other interest related to the work detailed in this manuscript.

REFERENCES

1 Evaluation IfHMa. Financing Global Health 2012: The End of the Golden Age? Seattle, WA: IHME; 2012. Available: http://www.healthdata.org/policy-report/financing-global-health-2012-end-golden-age. Accessed: 15 January 2016.

2 Ooms RH, Decoster K, Van Damme W. Global health: What it has been so far, what it should be, and what it could become. 2011. Available: http://www.itg.be/itg/Uploads/Volksgezondheid/wpshsop/SHSOP%20WP%202%20Ooms%20Global%20Health.pdf. Accessed: 15 January 2016.

3 Beaglehole R. What is global health? Global Health Action. 2010;3:5142. Medline:20386617 doi:10.3402/gha.v3i0.5142

4 Fidler DP. After the revolution: Global health politics in a time of economic crisis and threatening future trends. Mauer School of Law: Faculty publication, 2009; Paper 145. Available: http://www.repository.law.indiana.edu/facpub/145/. Accessed: 15 January 2016.

5 Bloom BR. WHO needs change. Nature. 2011;473:143-5. Medline:21562538 doi:10.1038/473143a

6 Sridhar D. Who sets the global health research agenda? The challenge of multi-bi financing. PLoS Med. 2012;9:e1001312. Medline:23049485 doi:10.1371/journal.pmed.1001312

7 Røttingen JA, Regmi S, Eide M, Young AJ, Viergever RF, Ardal C, et al. Mapping of available health research and development data: what's there, what's missing, and what role is there for a global observatory? Lancet. 2013;382:1286-307. Medline:23697824 doi:10.1016/S0140-6736(13)61046-6

8 Burke AM, Matlin SA, editors. Monitoring financial flows for health research in 2008. Prioritizing research for health equity. Geneva: Global Forum for Health Research, 2008. Available: http://announcementsfiles.cohred.org/gfhr_pub/assoc/s14888e/s14888e.pdf. Accessed: 15 January 2016.

9 Moran M, Guzman J, Ropars A-L, McDonald A, Jameson N, Omune B, et al. Neglected disease research and development: How much are we really spending? PLoS Med. 2009;6:e30. Medline:19192946 doi:10.1371/journal.pmed.1000030

10 Sridhar D, Batniji R. Misfinancing global health: a case for transparency in disbursements and decision making. Lancet. 2008;372:1185-91. Medline:18926279 doi:10.1016/S0140-6736(08)61485-3

11 Dandona L, Raban MZ, Guggilla RK, Bhatnagar A, Dandona R. Trends of public health research output from India during 2001-2008. BMC Med. 2009;7:59. Medline:19828017 doi:10.1186/1741-7015-7-59

12 Elsevier. International comparative performance of the UK research base. 2013. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/263729/bis-13-1297-international-comparative-performance-of-the-UK-research-base-2013.pdf. Accessed: 15 January 2016.

13 World Health Organization. WHO regional offices 2014. Available: http://www.who.int/about/regions/en/. Accessed: 15 January 2016.

14 World health Organization. Country groupings 2014. Available: http://www.who.int/quantifying_ehimpacts/global/ebdcountgroup/en/. Accessed: 15 January 2016.

15 Rudan I, Chopra M, Kapiriri L, Gibson J, Lansang MA, Carneiro I, et al. Setting priorities in global child health research investments: Universal challenges and conceptual framework. Croat Med J. 2008;49:307-17. Medline:18581609 doi:10.3325/cmj.2008.3.307

16 The World Bank. World development indicators 2014. Available: http://data.worldbank.org/indicator/NY.GDP.MKTP.KD. Accessed: 15 January 2016.

17 Asplund K, Eriksson M, Persson O. Country comparisons of human stroke research since 2001: a bibliometric study. Stroke. 2012;43:830-7. Medline:22282892 doi:10.1161/STROKEAHA.111.637249

18 Man JP, Weinkauf JG, Tsang M, Sin DD. Why do some countries publish more than others? An international comparison of research funding, English proficiency and publication output in highly ranked general medical journals. Eur J Epidemiol. 2004;19:811-7. Medline:15469039 doi:10.1023/B:EJEP.0000036571.00320.b8

19 Leroy JL, Habicht J-P, Pelto G, Bertozzi SM. Current priorities in health research funding and lack of impact on the number of child deaths per year. Am J Public Health. 2007;97:219-23. Medline:17194855 doi:10.2105/AJPH.2005.083287

REFERENCES

20 Merton RK. The Matthew effect in science: The reward and communication systems of science are considered.. Science. 1968;159:56-63. Medline:17737466 doi:10.1126/science.159.3810.56

21 Kelly CD, Jennions MD. The h-index and career assessment by numbers. Trends Ecol Evol. 2006;21:167-70. Medline:16701079 doi:10.1016/j.tree.2006.01.005

22 Gingras Y. The abuses of research evaluation. 2014. Available: http://www.universityworldnews.com/article.php?story=20140204141307557. Accessed: 15 January 2016.

23 Meera Sabaratnam PK. Why metrics cannot measure research quality: A response to the HEFCE consultation. Available: http://thedisorderofthings.com/2014/06/16/why-metrics-cannot-measure-research-quality-a-response-to-the-hefce-consultation/. Accessed: 15 January 2016.

24 Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci U S A. 2005;102:16569-72. Medline:16275915 doi:10.1073/pnas.0507655102

25 Labonte R. Setting global health research priorities: burden of disease and inherently global health issues should both be considered. BMJ. 2003;326:722-3. Medline:12676822 doi:10.1136/bmj.326.7392.722

26 Ranasinghe P, Jayawardena R, Katulanda P. Diabetes mellitus in South Asia: Scientific evaluation of the research output. J Diabetes. 2013;5:34-42. Medline:23379533 doi:10.1111/1753-0407.12003

journal of global health

PAPERS

Structure, function and five basic needs of the global health research system

Igor Rudan, Devi Sridhar

Centre for Global Health Research and WHO Collaborating Centre for Population Health Research and Training, The Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Scotland, UK

Background Two major initiatives that were set up to support and co–ordinate global health research efforts have been largely discontinued in recent years: the Global Forum for Health Research and World Health Organization's Department for Research Policy and Co-operation. These developments provide an interesting case study into the factors that contribute to the sustainability of initiatives to support and co–ordinate global health research in the 21st century.

Methods We reviewed the history of attempts to govern, support or co–ordinate research in global health. Moreover, we studied the changes and shifts in funding flows attributed to global health research. This allowed us to map the structure of the global health research system, as it has evolved under the increased funding contributions of the past decade. Bearing in mind its structure, core functions and dynamic nature, we proposed a framework on how to effectively support the system to increase its efficiency.

Results Based on our framework, which charted the structure and function of the global health research system and exposed places and roles for many stakeholders within the system, five basic needs emerged: (i) to co–ordinate funding among donors more effectively; (ii) to prioritize among many research ideas; (iii) to quickly recognize results of successful research; (iv) to ensure broad and rapid dissemination of results and their accessibility; and (v) to evaluate return on investments in health research.

Conclusion The global health research system has evolved rapidly and spontaneously. It has not been optimally efficient, but it is possible to identify solutions that could improve this. There are already examples of effective responses for the need of prioritization of research questions (eg, the CHNRI method), quick recognition of important research (eg, systems used by editors of the leading journals) and rapid and broadly accessible publication of the new knowledge (eg, *PLoS One* journal as an example). It is still necessary to develop tools that could assist donors to co–ordinate funding and ensure more equity between areas in the provided support, and to evaluate the value for money invested in health research.

Correspondence to:

Professor Igor Rudan, MD, DSc, PhD, MPH, HonMFPH
Chair in International Health and Molecular Medicine
Joint Director, Centre for Global Health Research and WHO Collaborating Centre for Population Health Research and Training
Director of Research, The Usher Institute for Population Health Sciences and Informatics
University of Edinburgh
Teviot Place, Edinburgh EH8 9AG
Scotland, UK
Igor.Rudan@ed.ac.uk

In the past four years, two major initiatives that were set up with the aim to support and co–ordinate global health research efforts have been largely discontinued. The first is the Global Forum for Health Research, which was established in Geneva in 1998 to support WHO's focus on health re-

search [1]. The second is WHO's Department for Research Policy and Cooperation (WHO RPC), which ceased its operations in 2012 during the WHO's internal reform. Almost ironically, the annual WHO World Health Report for 2012 announced its theme as: "*No health without research*" and was to be coordinated by the WHO RPC [2]. The journal *PLoS Medicine* agreed to publish a special series on health research in parallel to the release of the World Health Report, as discussed in the journal's editorial to the series, entitled: "*The World Health Report 2012 that Wasn't*" [3]. Eventually, the report was retitled "*Research for Universal Health Coverage*" and published in 2013 [4].

These developments provide an interesting case study into the factors that contribute to the sustainability of initiatives to govern, support and co–ordinate global health research in the 21st century. A timeline of key events that set the current context is shown **Figure 1**. In this viewpoint, we will map the structure of the global health research system as it has evolved under the funding increases of the past decade. Bearing in mind its structure, core functions and dynamic nature, we will propose a framework on how to effectively support the system to increase its efficiency.

THE EVOLVING STRUCTURE OF THE GLOBAL HEALTH RESEARCH SYSTEM

Over the past two decades, the funding available for health research has increased rather dramatically from US$ 50 billion in 1993 to US$ 240 billion in 2009 [5], but this did not happen in any planned or coordinated way. Those who tried tracking this funding – such as the Global Forum for Health Research in its annual reports, G–FINDER, the Institute for Health Metrics and Evaluation and other academics, provided rather different figures [5–9]. This discrepancy is largely due to the difficulty in distinguishing research funding from broader development assistance for health. There is also lack of consensus on whether the funding invested in high–income countries to study health challenges that may be relevant to low and middle–income countries should also be included. Still, under any assumption, the interest in funding global health research is growing, and the structure of this system is rapidly evolving.

In **Figure 2**, we show the simplified representation of the key stakeholders and processes, based on how the funds flow through the system. At the beginning of the system is the source of the funding – with donors being either public, private, or the emerging "class" of donors – the large philanthropies, such as the Bill and Melinda Gates Foundation (BMGF), the Carlos Slim Foundation, and the Rockefeller Foundation. They all provide financial support for the projects of researchers employed in universities, research institutes, international organizations, biotech companies and small and medium enterprises (SME are a growing "class" of recipients). They also fund stakeholders with research capacity in low and middle–income countries that can help carry out the research projects as equal partners. Eventually, the responsibility for spending the funds is passed down to research teams and their international consortia, which conduct research to generate new knowledge in several generic areas: measuring a problem; understanding its cause(s); elaborating solutions; translating the solutions or evidence into policy, practice and products; and/or evaluating the effectiveness of solutions [10].

The decision over the channel of dissemination of this knowledge is made by a new set of stakeholders (**Figure 2**), which may involve research committees of public institutions, journal editors, reviewers, donor representatives, company managers or owners. The bulk of work will end up published by research journals, where editors and reviewers, and sometimes even private publishers, influence decisions on the shape and form of publication. The funders increasingly require researchers to publish in open–access journals. Some of the findings do not get published because placing the knowledge in the public domain would invalidate patent applications and subsequent financial profits. This new knowledge can also be presented at conferences, published as a re-

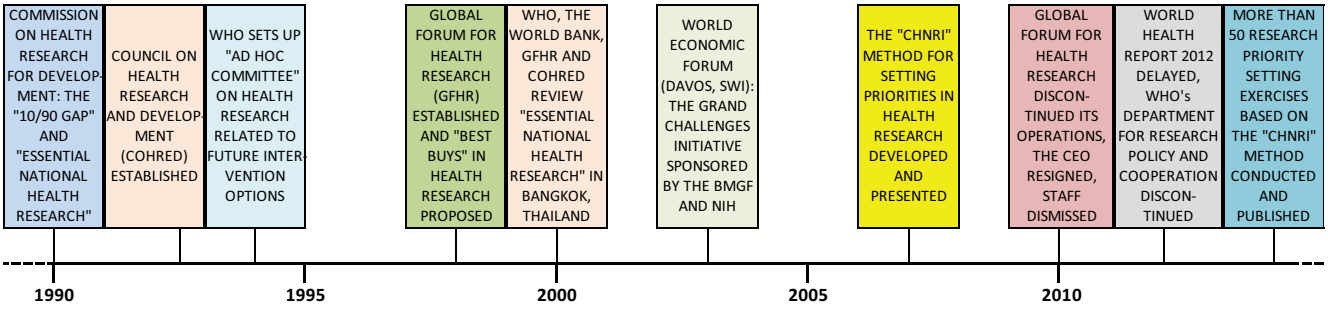


Figure 1. A timeline of several important events relevant to governance, support and co–ordination of global health research that determined the current context.

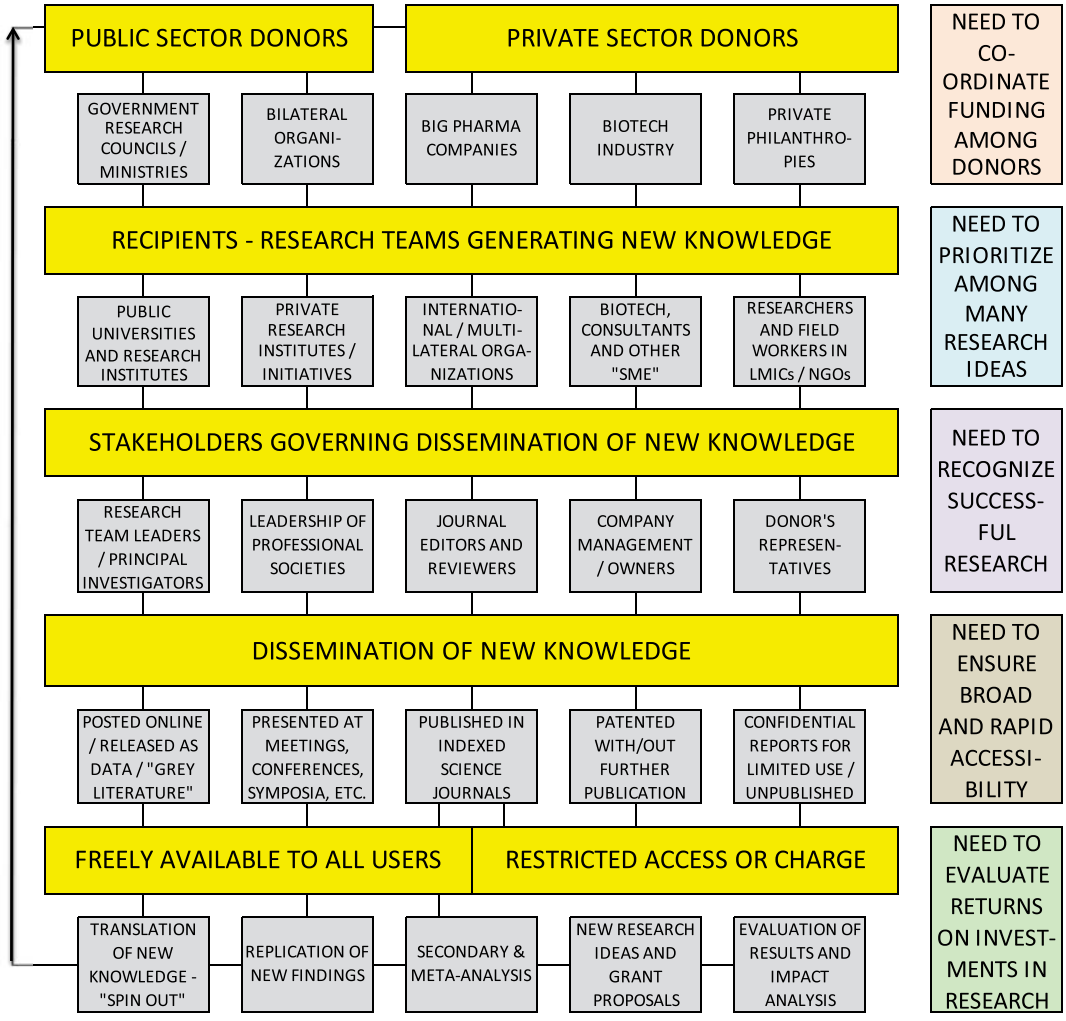


Figure 2. The structure of the global health research system and the five basic needs to ensure its efficient performance.

port to the funder, as "grey literature", or simply posted on the internet. Finally, in many cases, new knowledge does not get published in any way – perhaps due to insufficient relevance or novelty, concerns over its quality, or simply a lack of a positive result. In the end, the published knowledge can be professionally evaluated and replicated, with a growing industry of companies offering those services. Moreover, universities have set up structures to help researchers to commercialize their work and set up spin–out companies.

THE CORE FUNCTIONS OF THE GLOBAL HEALTH RESEARCH SYSTEM

There should not be much controversy over the main function of the global health research system: it is there to use donors' funding to support experiments that address pertinent health research questions. In this way, through answering those questions, new knowledge is continuously being generated. This knowledge is then translated into both clinical and public health practice in order to reduce the burden of disease in the population and improve health–related outcomes.

The effectiveness of the global health research system to perform its main function will depend on the efficiency of several of its sub–components (**Figure 2**). First, donors need to be *motivated* to continue investing; *informed* to understand the targets; and *coordinated* to avoid over– and under–funding certain areas. This, in turn, ensures efficiency of their investments. Second, researchers need to prioritize research ideas well, to balance those that could benefit the public relatively soon with more speculative and downstream ones. They need to design and conduct the experiments carefully to ensure that their efforts are useful even when the result is negative. Third, managers, journal editors and media need to recognize important progress accurately to ensure efficiency in selection of work that re-

ceives attention. Fourth, publishers need to ensure broad open access to all new knowledge that results from health research and rapid accessibility of information without exception. Fifth, the effectiveness of translation of the new knowledge into practice needs to be evaluated. This is important because it could help recognizing the most promising research projects and ideas earlier in the process. It would also allow comparisons of returns on investments in health research with other competing investments that could also improve health, such as development assistance, infrastructure projects, or simply increased purchase and coverage of existing interventions.

After the relatively stagnant nature of the global health research system throughout most of the second half of the 20th century, the system evolved rapidly over the past decade and took a life of its own in all of its segments. Attempts to support and co–ordinate such a dynamic and unpredictably evolving system using a 'top–down' approach may have seemed a feasible and sustainable mission from the perspective of the post–World War II world, when the UN was established. However, the 21st century global health research system has developed in a "bottom–up", "laissez–faire" manner, in which the stakeholders themselves are continuously inventing improved practices and introducing changes in the models that worked well in previous decades. This is happening at all levels – with emerging big donors, innovative finance mechanisms, creative organization of large international consortia of research teams and their collaborations on "big science". There are now many web–based routes to publication, new tools and measures of assessment of research output (like *Google Scholar, Scopus, Research Gate* and *H–index* metric), and increased support mechanisms for rapid translation, commercialization and implementation of research results. In such a dynamic system, any attempt to influence the relevant stakeholders and processes from the "outside" by a group of experts who drive their legitimacy exclusively from a fact that they are employees or affiliates of the UN is largely unrealistic and outdated.

FIVE BASIC NEEDS OF THE GLOBAL HEALTH RESEARCH SYSTEM AND PROPOSED SOLUTIONS TO IMPROVE ITS EFFICIENCY

We now propose an alternative route to improved efficiency of the global health research system that would be primarily needs–based, and therefore likely welcomed by the stakeholders in the system. At the top of **Figure 2**, it is clear that the emergence of new donors is certainly a positive development, but it requires their sustained motivation and also carries a large risk of becoming un–coordinated and unbalanced, with high preference towards certain topics and neglect of others. This is a real risk that has already been exposed in even the most basic analysis of funding flows [7]. To help the system develop and grow in an equitable way at this level, there is a need to continuously track funding using an internationally agreed methodology, preferably by more than one agency/institute. Beyond simply tracking funding, a tool is needed to ensure that no areas are neglected in comparison to areas of strong donor preference, thus assisting policy–makers and donor representatives. As a possible solution, we are working to propose a "Stock Market for Global Health Research Investment Options" – a tool that would use analogy to real–time stock markets to compare the burden of different health problems with the investments being committed to those problems, using the most recent available information.

The main need at the level of the recipients in the system – the communities of researchers (**Figure 2**) – is to find ways to communicate and agree on their own field's research priorities, so that a more balanced and unified case on funding priorities could be presented to donors from the "cutting edge" of research. As a possible solution, "*the CHNRI method*" developed by the Child Health and Nutrition Research Initiative (CHNRI) of the Global Forum for Health Research seems to be an example of this need being met rather effectively [11]. This "crowd–sourcing" approach to generating and managing research ideas, while balancing short–term and long–term vision and different instruments of health research, has been validated through many applications [12–15]. The results from 50 conducted research prioritization exercises have been published by mid–2015, and many further exercises are being conducted presently [15]. A recent independent review showed that 18% of prioritization exercises in global health research in recent years used the CHNRI methodology, which made it the most frequently used specific priority–setting method [16].

Then, at the level of stakeholders who govern dissemination of research results (**Figure 2**), there is a need for a tool, process or a system that would recognize important research, promote and reward it appropriately [17]. Interestingly, journal editors operate such systems already while reaching their decisions on which papers to publish. Given that many of them select less than 10% of submissions for publication, the journals that manage to maintain high quality and substantial impact over time have clearly developed well–performing systems. We propose to learn more of their decision–making systems and processes and review the results of their work – both at the level of journal's impact, and of individual papers – over long periods of time. This should allow development of a system that would be highly sensitive to important research and ensure its publication, but also quite specific, reducing the amount of published work that is not relevant.

Clearly, it is difficult to predict the impact that research articles may have in the future at the point at which they are being evaluated. However, in the new world of "big data", it is possible to conduct massive exercises in available databases of research papers and their received citations to search for common patterns that are shared among those papers that have most impact. Recently, the journal *Nature* devoted a special news feature to analysis of the 100 most cited papers of all time [18]. In a related feature, titled "Is your most cited work your best?", Ioannidis et al. tried to capture the key dimensions that need to be addressed to make any biomedical research "exceptional". They asked about 400 most cited biomedical scientists in the world (123 of whom responded) to score their 10 most cited papers from 0–100 for each of the six criteria that they hypothesized may be inherent to truly exceptional work. They termed these six criteria "Continuous Progress, Broader Interest, Greater Synthesis, Disruptive Innovativeness, Surprise and Publication Difficulty" [19]. Their exercise made some of the first steps towards a more systematic and transparent framework that could allow capturing the exceptional nature of biomedical research articles at the time they are evaluated, rather than having to wait for many years to determine their importance through impact they generated and citations they received [19].

At the next level – dissemination of new knowledge (**Figure 2**) – the need for broad and rapid access to new knowledge is presently being addressed through the "open access" movement, world wide web development, IT–based solutions for publication, dissemination and search engines, social networks and internet–based media [20]. The success of *PLoS One* journal can be used as an excellent example. We believe that the journal succeeded in a very short time, and well beyond expectations, precisely because it provided an effective solution to this particular need of the global health research system. It is enough to state that in the year of its inception, in 2006, it published 137 papers; in 2007 it already published 1230 papers, and in 2013 a staggering 31 498 papers, with the number per year still growing strongly. At the same time, given an unprecedentedly large denominator, it still manages to keep a very decent impact factor of around 4.0 in the past several years. Clearly, many participants in the global health research system have recognized *PLoS One* as a solution that addresses one of the system's major needs.

Finally, at the level of research outputs, a tool is needed that could evaluate returns on investments in global health research, and what is seen as the value for money gained through those investments [21]. The tool should also monitor success rates in translation and implementation of the outcomes into products and programmes, all the way to measurable benefits for global public health. Such a tool would allow a proper understanding of the actual value of investing in health research, in comparison to alternative forms of investments that can also benefit health – eg, community infrastructure projects, improved education, safety, social welfare, and transportation. It is perhaps time to get some understanding on whether the many trillions invested in health research have been a reasonable investment – especially in the wake of Big Pharma largely closing down their R&D departments, which may provide an indication that they are concerned about the feasibility of those investments in comparison to alternatives. This need will be the most difficult to address, but we aim to propose a draft solution and keep improving it over time.

CONCLUSION

The global health research system has evolved rapidly and spontaneously. It has not been optimally efficient, but it is possible to identify solutions that could improve this. There are already examples of effective responses for the need of prioritization of research questions (eg, the CHNRI method), rapid recognition of important research (eg, systems used by editors of the leading journals) and quick and broadly accessible publication of the new knowledge (eg, *PLoS One* journal as an example). It is still necessary to develop tools that could assist donors to co–ordinate funding and ensure more equity between areas in the provided support, and to evaluate the value for money invested in health research.

Funding: None.

Ethical approval: Not required.

Authorship declaration: IR and DS conducted the review of the relevant literature, developed the framework and wrote the paper.

Declaration of interest: The authors completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). Authors declare no conflicting financial or other interest related to the work detailed in this manuscript.

REFERENCES

1 Peterson–Stearns B. Why has the global forum for health research collapsed? SciDev.Net, 15 November 2010. Available: http://www.scidev.net/global/health/opinion/why–has–the–global–forum–for–health–research–collapsed–.html. Accessed: 15 August 2015.

2 Pang T, Terry RF, PLoS Medicine Editors. WHO/PLoS Collection "No health without research": A call for papers. PLoS Med. 2011;8: e1001008. doi:10.1371/journal.pmed.1001008

3 PLoS Medicine Editors, Barbour V, Clark J,Connell L, Simpson P, Winker M, Veitch E. The World Health Report 2012 that wasn't. PLoS Med. 2012;9:e1001317. Medline:23049486 doi:10.1371/journal.pmed.1001317

4 World Health Organization. Research for universal health coverage: World Health Report 2013. Geneva: WHO, 2013.

5 Rřttingen JA, Regmi S, Eide M, Young AJ, Viergever RF, Ardal C. Mapping of available health research and development data: what's there, what's missing, and what role is there for a global observatory? Lancet. 2013;382:1286-307. Medline:23697824 doi:10.1016/S0140-6736(13)61046-6

6 Burke MA, Matlin SA. Monitoring financial flows for health research 2008: Prioritizing research for health equity. Global Forum for Health Research, Geneva, 2008. Available: http://announcementsfiles.cohred.org/gfhr_pub/assoc/s14888e/s14888e.pdf. Accessed: 15 August 2015.

7 Moran M, Guzman J, Ropars AL, McDonald A, Jameson N, Omune B, et al. Neglected disease research and development: How much are we really spending? PLoS Med. 2009;6:e1000030. Medline:19192946 doi:10.1371/journal.pmed.1000030

8 Institute of Health Metrics and Evaluation. Financing global health 2012: The end of the golden age? Seattle: Institute of Health Metrics and Evaluation, 2012. Available: http://www.healthmetricsandevaluation.org/publications/policy–report/financing–global–health–2012–end–golden–age. Accessed: 15 August 2015.

9 Sridhar D, Batniji R. Misfinancing global health: a case for transparency in disbursements and decision–making. Lancet. 2008;372:1185-91. Medline:18926279 doi:10.1016/S0140-6736(08)61485-3

10 World Health Organization. WHO's role and responsibilities in health research: Draft WHO strategy on research for health. World Health Assembly A63/22. Geneva: World Health Organization, 2010. Available: http://www.wpro.who.int/health_research/policy_documents/draft_who_strategy_on_research_for_health.pdf. Accessed: 15 August 2015.

11 Rudan I, Gibson JL, Ameratunga S, El Arifeen S, Bhutta ZA, Black M, et al; Child Health and Nutrition Research Initiative (CHNRI). Setting priorities in global child health research investments: guidelines for implementation of the CHNRI Method. Croat Med J. 2008;49:720-33. Medline:19090596

12 Fontaine O, Kosek M, Bhatnagar S, Boschi–Pinto C, Chan KY, Duggan C, et al. Setting research priorities to reduce global mortality from childhood diarrhoea by 2015. PLoS Med. 2009;6:e41. Medline:19278292 doi:10.1371/journal.pmed.1000041

13 Rudan I, El Arifeen S, Bhutta ZA, Black RE, Brooks A, Chan KY, et al. the WHO/CHNRI Expert Group on Childhood Pneumonia. Setting research priorities to reduce global mortality from childhood pneumonia by 2015. PLoS Med. 2011;8:e1001099. Medline:21980266 doi:10.1371/journal.pmed.1001099

14 Tomlinson M, Swartz L, Officer A, Chan KY, Rudan I, Saxena S. Research priorities for health of people with disabilities: an expert opinion exercise. Lancet. 2009;374:1857-62. Medline:19944866 doi:10.1016/S0140-6736(09)61910-3

15 Rudan I. Global health research priorities: mobilizing the developing world. Public Health. 2012;126:237-40. Medline:22325672 doi:10.1016/j.puhe.2011.12.001

16 McGregor S, Henderson KJ, Kaldor JM. How are health research priorities set in low and middle income countries? A systematic review of published reports. PLoS ONE. 2014;9:e108787. Medline:25275315 doi:10.1371/journal.pone.0108787

17 The maze of impact metrics. Nature. 2013;502:271. Medline:24137834 doi:10.1038/502271a

18 Van Noorden R, Maher B, Nuzzo R. The top 100 papers. Nature. 2014;514:550-3. Medline:25355343 doi:10.1038/514550a

19 Ioannidis JP, Boyack KW, Small H, Sorensen AA, Klavans R. Bibliometrics: Is your most cited work your best? Nature. 2014;514:561-2. Medline:25355346 doi:10.1038/514561a

20 Wolpert AJ. For the sake of inquiry and knowledge — the inevitability of open access. N Engl J Med. 2013;368:785-7. Medline:23445088 doi:10.1056/NEJMp1211410

21 Lane J, Bertuzzi S. Measuring the results of science investments. Science. 2011;331:678-80. Medline:21310987 doi:10.1126/science.1201865

journal of global health

PAPERS

Improved measurement for mothers, newborns and children in the era of the Sustainable Development Goals

Tanya Marchant[1], Jennifer Bryce[2], Cesar Victora[3], Allisyn C Moran[4], Mariam Claeson[5], Jennifer Requejo[2], Agbessi Amouzou[6], Neff Walker[2], Ties Boerma[7], John Grove[5]

[1] Department of Disease Control, London School of Hygiene and Tropical Medicine, London, UK

[2] Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, USA

[3] International Center for Equity in Health, Post–Graduate Programme in Epidemiology, Federal University of Pelotas, Pelotas, Brazil

[4] Global Health Fellows Program II, Bureau for Global Health, US Agency for International Development, Washington, USA

[5] Bill & Melinda Gates Foundation, Seattle, WA, USA

[6] UNICEF, New York, NY, USA

[7] WHO, Health Systems and Innovation, Geneva, Switzerland

Background An urgent priority in maternal, newborn and child health is to accelerate the scale–up of cost–effective essential interventions, especially during labor, the immediate postnatal period and for the treatment of serious infectious diseases and acute malnutrition. Tracking intervention coverage is a key activity to support scale–up and in this paper we examine priorities in coverage measurement, distinguishing between essential interventions that can be measured now and those that require methodological development.

Methods We conceptualized a typology of indicators related to intervention coverage that distinguishes access to care from receipt of an intervention by the population in need. We then built on documented evidence on coverage measurement to determine the status of indicators for essential interventions and to identify areas for development.

Results Contact indicators from pregnancy to childhood were identified as current indicators for immediate use, but indicators reflecting the quality of care provided during these contacts need development. At each contact point, some essential interventions can be measured now, but the need for development of indicators predominates around interventions at the time of birth and interventions to treat infections. Addressing this need requires improvements in routine facility based data capture, methods for linking provider and community–based data, and improved guidance for effective coverage measurement that reflects the provision of high–quality care.

Conclusion Coverage indicators for some essential interventions can be measured accurately through household surveys and be used to track progress in maternal, newborn and child health. Other essential interventions currently rely on contact indicators as proxies for coverage but urgent attention is needed to identify new measurement approaches that directly and reliably measure their effective coverage.

Correspondence to:
Tanya Marchant
Department of Disease Control
London School of Hygiene and Tropical Medicine
Keppel Street
London WC1E 7HT
UK
tanya.marchant@lshtm.ac.uk

Within the 17 Sustainable Development Goals (SDGs) a total of 169 targets and over 230 indicators have been defined [1]. In alignment with the SDGs, the Global Strategy for Women's, Children's and Adolescents' Health (the Global Strategy) has described an ambitious action and measurement agenda around the three pillars "Survive, Thrive and Transform" [2]. In the immediate future many countries have an unfinished agenda to accelerate the scale–up of cost–effective essential maternal, newborn and child health (MNCH) interventions that save lives as well as help families to thrive [3]. Tracking intervention coverage is a top priority to

assist this scale–up so that countries know the extent to which populations in need are benefiting, and delivery strategies are refined as a result [4]. In this paper we examine priorities in coverage measurement of essential MNCH interventions, distinguishing between those that can be measured now and those that require methodological development.

Of particular importance is to explicitly acknowledge known measurement challenges across the continuum from pregnancy to childhood [5–7], and categorise indicators that can be measured now using existing methods and tools ("*indicators for immediate use*"), and those that are high priority in the context of life–saving, quality care but require further methodological development and validation ("*priority indicators for development*"). Once validated using feasible methods, these priority indicators for development can be further described in global guidance and integrated within existing data collection systems.

The remainder of this paper proposes a transparent set of evidence–based considerations for the global MNCH measurement improvement agenda. We draw on evidence supporting cost–effective investments in MNCH [3], recommendations by the Global Strategy [8], and the priorities identified by other initiatives including the Global Reference List of 100 core indicators [9], the World Health Organization's consultation on quality MNCH [10], the Every Newborn Action Plan (ENAP) [11], and Ending Preventable Maternal Mortality (EPMM) [12].

METHODS AND CONSIDERATIONS IN SELECTING INDICATORS FOR IMMEDIATE USE

Figure 1 presents a typology of indicators related to intervention coverage. Level A encompasses all women and children who can benefit from receiving care, including preventive and curative services. From this group, only some will access care and have the opportunity to benefit from the services they need (level B). But making contact with services does not ensure receipt of a specific intervention (level C), irrespective of whether the population making contact needs a preventive or curative intervention. Currently, coverage measurement for any given intervention is defined as C/A, or the proportion of women and children who need an intervention who actually receive it. The innermost element of the framework (level D) highlights the importance of incorporating dimensions of quality within coverage, often referred to as "effective coverage", for example including measures of appropriate diagnosis, drug dosage, or counselling. The need for development of globally standardised measures of effective coverage is described in more detail below.

Our considerations for determining the measurement status of indicators builds on the experience and evidence

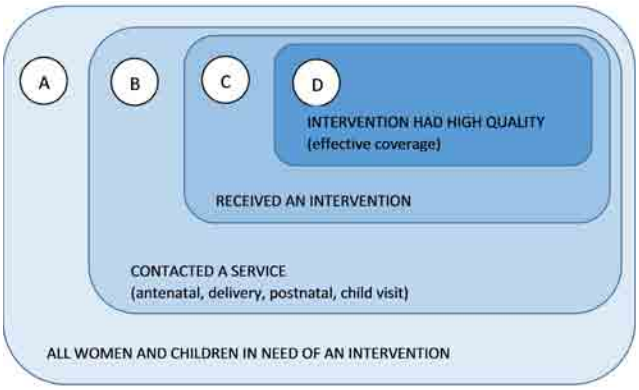


Figure 1. Typology of indicators for maternal, newborn and child health.

base generated by others, including household survey programs such as Demographic and Health Surveys (DHS) [13] and the Multiple Indicator Cluster Surveys (MICS) [14], the Countdown to 2015 for Maternal, Newborn and Child Survival (Countdown) initiative [15], and the investment and visibility promoted by the Commission on Information and Accountability for Women's and Children's Health (CoIA) and its independent Expert Review Group (iERG) [16].

We took five characteristics into account in selecting priority indicators.

1) **Public health importance**. Priority indicators should measure progress in coverage for an intervention that has the potential to save a large number of women's and children's lives, because it is linked through known channels to changes in health status. We estimate this potential using the Lives Saved Tool (LiST) [17], calculating the number of maternal, newborn and child lives that could be saved by 2030 based on the underlying assumptions within the model, and if universal coverage was achieved for the intervention in the 75 countries that accounted for 99% of deaths among those groups in 2014, assuming coverage trajectories for all other interventions remain the same (**Table 1**). We have included indicators for malaria and HIV because of their importance in some high burden countries, even though they do not account for large numbers of deaths in all countries.

2) **Feasibility and affordability**. Indicators for immediate use must be affordable and feasible for accurate measurement in the majority of high–MNCH mortality countries to inform immediate actions. But high–impact interventions for which feasible and cost–effective measurement strategies are not currently available must not be lost and are the target of an urgent developmental research agenda, described below under *priority indicators for development*.

3) **Accuracy.** Measurement approaches that do not produce valid results are a waste of scarce resources, and

Table 1. LiST analysis of lives saved by labor and delivery management, and life–saving interventions for mothers, newborns and children*

	Estimated number of deaths averted					
Intervention	**Stillbirths**	**Neonatal**	**Child**	**Maternal**	**Total**	**Rank**
Labor & delivery management	689758	549031		76850	1315639	1
Full supportive care for prematurity		544458			544458	2
Full supportive care for sepsis/pneumonia		409877			409877	3
Oral Rehydration Solution		12653	369423		382076	4
Water connection in the home			368313		368313	5
Treatment with antimalarials			303653		303653	6
Oral antibiotics for pneumonia			300682		300682	7
Promotion of breastfeeding		74699	191976		266675	8
Hand washing with soap			235898		235898	9
Neonatal resuscitation		212439			212439	10
Therapeutic feeding for severe wasting			209442		209442	11
Injectable antibiotics for neonatal sepsis/pneumonia		181512			181512	12
Kangaroo Mother Care		158853			158853	13
Syphilis detection and treatment	149597	7060			156657	14
Pneumococcal vaccine			139779		139779	15
Improved sanitation			136256		136256	16
Clean postnatal practices		131782			131782	17
Clean birth practices		101266		20148	121414	18
Treatment for moderate acute nutrition of children			110671		110671	19
Immediate assessment and stimulation of newborns		109585			109585	20
Hib vaccination			106998		106998	21
Zinc–for treatment of diarrhea			106481		106481	22
Zinc supplementation			104426		104426	23
Magnesium sulphate for pre–eclampsia	64939			23681	88620	24
Homes protected from malaria by ownership of insecticide treated nets or indoor residual spraying			87733		87733	25
Chlorhexidine for cord care		82283			82283	26
Appropriate complementary feeding			80081		80081	27
Intermittent presumptive treatment for malaria in pregnancy	59942	16111	1539	1404	78996	28
Oral antibiotics for neonatal sepsis or pneumonia		74462			74462	29
Thermal care for newborns		72391			72391	30
Hygienic disposal of stools			64653		64653	31
Periconceptual Folic Acid / Ferrous Sulfate	17711	43296			61007	32
Antibiotics for premature preterm rupture of membranes		49257		7903	57160	33
Rotavirus vaccine			56788		56788	34
Induction of labor for pregnancies beyond 42 weeks	47230				47230	35
Balanced energy protein supplementation for pregnant women		41268	3309		44577	36
Multiple micronutrients for pregnant women		39615	2788		42403	37
Active management of third stage of labor				33782	33782	38
Case management of maternal sepsis				23528	23528	39
Iron supplementation for pregnant women		21964	1555		23519	40
Diabetes case management for pregnant women	22585				22585	41
Magnesium sulphate for treatment of eclampsia				22572	22572	42
Improved water			21470		21470	43
Case management of hypertensive disorders in pregnant women				20025	20025	44
Safe abortion services				15529	15529	45
DPT3 vaccination			15428		15428	46
Tetanus toxoid vaccination		14940		161	15101	47
Vitamin A supplementation			14967		14967	48
Vitamin A–for treatment of measles			14574		14574	49
Post abortion case management				13391	13391	50
Calcium supplementation				8124	8124	51
Ectopic pregnancy case management				2980	2980	52
Case management of malaria in pregnant women				2347	2347	53
Antibiotics for dysentery			1017		1017	54

Hib – *Haemophilus influenzae* type B, DPT3 – diphtheria–tetanus–pertussis

*The potential number of lives saved by 54 evidence based interventions by 2030, estimated using the Lives Saved Tool if universal coverage was achieved for each intervention in the 75 countries that accounted for 99% of maternal, newborn and child deaths in 2014, assuming coverage trajectories for all other interventions remain the same.

can misdirect policy and program decisions. There is a growing body of research demonstrating that mothers interviewed during household surveys (as in DHS or MICS) can report accurately on whether they and their children received some interventions, but not others. Particularly problematic are high impact interventions around the time of birth and curative interventions for episodes of illness such as antibiotics for pneumonia [18–20]. New and innovative approaches for measuring coverage for these interventions are needed urgently, while maintaining support for household surveys able to produce highly–accurate estimates of coverage for most MNCH interventions. Surveys are also essential for assessing equity through disaggregated analyses, as required by SDG target 17.18 on the measurement of inequalities.

4) **Production of timely results with clear action implications.** Indicator levels should change in response to increases or decreases in program inputs and outputs and improvements in program processes, within a time frame of one to three years, to provide information useful to program managers. Experience has demonstrated that monitoring systems work best and are more likely to be sustained if the data they contain are used first at the level at which they are collected, and also at each higher level throughout the reporting system. Of importance is to encourage reporting and use of individual indicator components from the point of data collection through national level, but combining the components for global monitoring.

5) **Consistency with historical indicators, to permit tracking of trends.** Lists of indicators evolve over time. New interventions are scaled up that require new indicators, but also the validity of existing indicators may be challenged by new evidence. For example, the indicator for diarrhea management used in most surveys since the 1990s was oral rehydration therapy (ORT), but more recently there has been a shift towards reporting on oral rehydration salts (ORS) plus zinc [21]. For the purpose of assessing time trends as we transition from the Millennium Development Goals (MDGs) to the SDGs it is useful to continue to report on ORT as well as ORS for a period of time, while also designing measurement methods so that adjustments to indicator definition can be made.

GETTING STARTED: INDICATORS FOR IMMEDIATE USE

In **Figure 2** we present the contacts and interventions prioritized by different global groups in MNCH (for example ENAP, EPMM, the Global Strategy), and include those supported by evidence of impact from LiST analysis (**Table 1**). After consideration of the five characteristics above these have been categorized as "*current*" or "*priority for development*".

Contacts are included in order to measure the proportion of individuals accessing care, and thus potential to receive interventions, corresponding to level B in **Figure 1**. In addition to the contacts for antenatal care, skilled attendant at birth, and postnatal care, we also include care seeking for sick children (specifically fever and symptoms of childhood pneumonia), consistent with the typology that distinguishes accessing care from actual receipt of a life–saving intervention. Correct treatment of these two conditions are among the highest–impact interventions, but cannot be measured accurately through household surveys. We also indicate the need to develop, agree on and validate indicators that reflect quality care at these contact points to enable tracking of effective coverage measures [22]. The remainder of **Figure 2** presents intervention indicators. High impact interventions are represented across the continuum from pregnancy to childhood and measurement development needs are identified at each stage. Addressing these needs requires immediate action, as described in the next section.

DOING BETTER: AN ACTION AGENDA FOR IMPROVED MEASUREMENT

Priority indicators for development predominate around interventions at the time of birth, interventions to treat infections, and quality of care. Some of these represent relatively rare events (for example antibiotics for preterm premature rupture of membranes) and may never be suitable for population level tracking at national level, but nonetheless require advances in measurement in order to report accurately to country programs. For many, service contact indicators have been used to represent imperfect proxy measures of care but the need for measures of quality care means that we have to do better. For example, the service contact indicator "skilled attendant at birth" is the most widely used proxy indicator for care at birth, but the evidence linking increases in skilled attendant coverage with reductions in mortality has not been consistent [23–25], probably reflecting the fact that only a subset of locally–defined skilled attendants actually have the skills, commodities and facilities needed to deliver essential interventions at birth.

We propose that four specific types of measurement innovations are required.

First, a measurement improvement agenda is needed for routine data capture, so that the accuracy of reporting clinical interventions for women, newborns and children is improved at different levels of the health system. This will allow delivery of high impact interventions to be tracked at local, national and global levels. It will require improved routine data systems, review and consolidation of facility assessment tools and methods, and engagement with health system strengthening efforts more broadly.

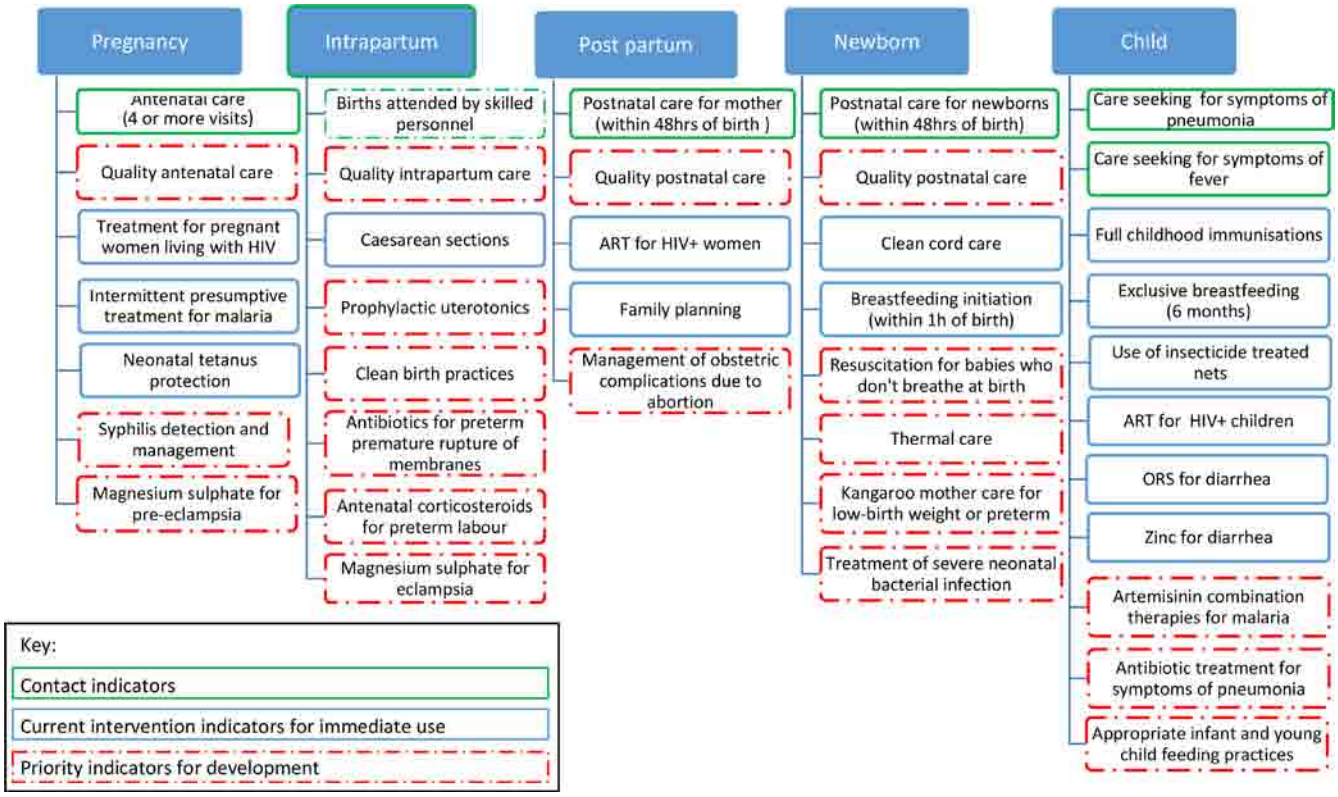


Figure 2. Measurement status of priority contacts and evidence based interventions across the continuum from pregnancy to childhood.

Second, to realize the potential of these improved data sources, methods for linking population and provider–based data sources are needed [5,17]. Household survey methods provide population level data and permit equity analysis but can be limited by poor recall and infrequent reporting. Facility data can be continuous and timely, has potential to improve reporting on clinical events, and can be stratified by level and type. However, present reporting tools cannot provide accurate equity breakdowns or population level estimates. Combining these two data streams has the potential to be transformative for monitoring the delivery of essential interventions that cannot currently be measured reliably, and for measuring effective coverage so that coverage indicators are defined as level D/A in **Figure 1**.

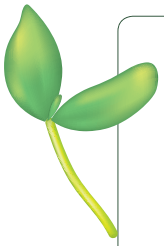
Third, further advances in implementation science are needed in order to place indicator development in the context of research on the design, implementation and impact of large scale programs.

And fourth, as new measures and approaches are tested and proven ready for wider adoption, global resources and guidance should be developed. Resources would include access to questionnaires, forms, and protocols; perhaps in one accessible system. Global guidance would include definitions, strengths and limitations of potential data sources, and interpretation notes.

TAKING THE AGENDA FORWARD

This paper adds to other recent calls for improved measurement that can enhance accountability and refine strategies to save lives [26]. At this time of transition from the MDGs to the SDGs it is essential that baselines are established, ambition is maintained, guidance and resources are shared, and momentum is not lost. Clarity about which essential interventions can be measured directly, reliably and feasibly using existing methods is an integral part of that plan. But here we also identify the need for focused, intensive commitment to advance the coverage measurement agenda for all essential interventions–especially those that save lives during and immediately after childbirth, and for sick children–so that we progress from reliance on measuring contacts with health care providers to measuring the effective coverage of clinical high–impact interventions.

As we enter the SDG era, several key partners are stepping forward to join this global measurement agenda for maternal, newborn and child health to agree on priorities, to coordinate actions and learning, and to work together with countries so that ownership of and capacity for an improved measurement agenda sits where the ability to act on evidence is greatest.



Funding: None.

Disclaimer: The views expressed in this manuscript are those of the authors and do not represent the US Government.

Authorship declarations: The concept for the article was defined by TM, JB, CV, and JG. All authors contributed to and approved the final draft.

Competing interests: All authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/coi_disclosure.pdf (available upon request from the corresponding author) and declare no conflict of interest.

REFERENCES

1 The Sustainable Development Goals. Available at: http://www.sustainabledevelopment2015.org/. Accessed on: 31 January 2016.

2 World Health Organisation. Global strategy for women's children's and adolescents health. Available: http://www.who.int/life–course/partners/global–strategy/en/. Accessed: 13 January 2016.

3 Black RE, Levin C, Walker N, Chou D, Liu L, Temmerman M; DCP3 RMNCH Authors Group. Reproductive, maternal, newborn and child health: key messages from Disease Control Priorities 3rd Edition. Lancet. 2016;pii:S0140-6736(16)00738-8. Epub 9 April 2016.

4 Grove J, Claeson M, Bryce J, Amouzou A, Boerma T, Waiswa P, et al; Kirkland Group. Maternal, newborn, and child health in the era of the Sustainable Development Goals—a call for sustained and improved measurement. Lancet. 2015;386:1511–4. Medline:26530604 doi:10.1016/S0140–6736(15)00517–6

5 Bryce J, Arnold F. Blanc Ann, Hancioglu A, Newby H, Requejo J, Wardlaw T, the CHERG Working Group on Improving Coverage Measurement (2013). Measuring coverage in MNCH: new findings, new strategies, and recommendations for action. PLoS Med. 2013;10:e1001423. Medline:23667340 doi:10.1371/journal.pmed.1001423

6 Eisele TP, Rhoda DA, Cutts FT, Keating J. Ren Ru, Barros AJD, Arnold F. Measuring coverage in MNCH: total survey error and the interpretation of intervention coverage estimates from household surveys. PLoS Med. 2013;10:e1001386. Medline:23667331 doi:10.1371/journal.pmed.1001386

7 Moxon SG, Ruysen H, Kerber KJ, Amouzou A, Fourniers S, Grove J, et al. Count every newborn; a measurement improvement roadmap for coverage data. BMC Pregnancy Childbirth. 2015;15 Suppl 2:S8. Medline:26391444 doi:10.1186/1471–2393–15–S2–S8

8 World Health Organization. Indicator and monitoring framework for the global Strategy for Women's, Children's and Adolescents' Health (2016–2030). Available: http://www.everywomaneverychild.org/images/content/files/EWEC_INDICATOR_MONITORING_FRAMEWORK_2016.pdf. Accessed: 4 June 2016.

9 The WHO Global Reference List of 100 Core Indicators. Available: http://www.who.int/healthinfo/indicators/2015/en/. Accessed: 31 January 2016.

10 World Health Organization and Partnership for Maternal, Newborn and Child Health. Consultation on improving measurement of the quality of maternal, newborn and child care in health facilities. Available: http://apps.who.int/iris/bitstream/10665/128206/1/9789241507417_eng.pdf. Accessed: 1 June 2016.

11 Every Newborn. The Every Newborn Action Plan. Available: http://www.everynewborn.org/every–newborn–event–summary–enap–metrics–meeting/. Accessed: 31 January 2016.

12 World Health Organization. Strategies towards Ending Preventable Maternal Mortality. Available: http://www.who.int/reproductivehealth/topics/maternal_perinatal/epmm/en/. Accessed: 31 January 2016.

13 USAID. The DHS Program, Demographic and Health Surveys. Available: http://dhsprogram.com/. Accessed: 17 September 2015.

14 UNICEF Statistics and Monitoring. Multiple Indicator Cluster Surveys. Available: http://www.unicef.org/statistics/index_24302.html. Accessed: 17 September 2015.

15 Countdown to 2015. Countdown to 2015 for Maternal, Newborn and Child Survival. Available: http://www.countdown2015mnch.org/. Accessed: 17 September 2015.

16 World Health Organization. Implementing the Commission on Information and Accountability Recommendations. Accountability for women's and children's health. Available: http://www.everywomaneverychild.org/resources/archives/coia. Accessed: 17 September 2015.

17 Johns Hopkins Bloomberg School of Public Health. The Lives Saved Tool. Available: http://www.livessavedtool.org/. Accessed: 21 January 2016.

18 Munos MK, Stanton C, Bryce J; Core Group for Improving Coverage Measurement for MNCH. Improving coverage measurement for reproductive, maternal, neonatal and child health: let's get a plan. J Glob Health. Forthcoming 2016.

19 Campbell H, el Arifeen S, Hazir T, O'Kelly J, Bryce J, Rudan I, et al. Measuring coverage in MNCH: challenges in monitoring the proportion of young children with pneumonia who receive antibiotic treatment. PLoS Med. 2013;10:e1001421. Medline:23667338 doi:10.1371/journal.pmed.1001421

20 Hazir T, Begum K, el Arifeen S, Khan AM, Huque MH, Kazmi N, et al. Measuring coverage in MNCH: a prospective validation study in Pakistan and Bangladesh on measuring correct treatment of childhood pneumonia. PLoS Med. 2013;10:e1001422. Medline:23667339 doi:10.1371/journal.pmed.1001422

REFERENCES

21 Countdown to 2015. Countdown to 2015: a decade of tracking progress for maternal, newborn and child survival. The 2015 Report. Available: http://www.countdown2015mnch.org/reports–and–articles/2015–final–report. Accessed: 31 January 2016.

22 Marchant T, Tilley–Gyado R, Tessema T, Singh K, Gautham M, Umar N, et al. Adding content to contacts: measurement of high quality contacts for maternal and newborn health in Ethiopia, north east Nigeria, and Uttar Pradesh, India. PLoS ONE. 2015;10:e0126840. Medline:26000829 doi:10.1371/journal.pone.0126840

23 Scott S, Ronsmans C. The relationship between birth with a health professional and maternal mortality in observational studies: a review of the literature. Trop Med Int Health. 2009;14:1523–33. Medline:19793070 doi:10.1111/j.1365–3156.2009.02402.x

24 Holmer H, Oyerinde K, Meara JG, Gillies R, Liljestrand J, Hagander L. The global met need for emergency obstetric care: a systematic review. BJOG. 2015;122:183–9. Medline:25546039 doi:10.1111/1471–0528.13230

25 Fink G, Ross R, Hill K. Institutional deliveries weakly associated with improved neonatal survival in developing countries: evidence from 192 Demographic and Health Surveys. Int J Epidemiol. 2015;44:1879–88. Medline:26130739 doi:10.1093/ije/dyv115

26 Handley K, Boerma T, Victora C, Evans TG. An inflection point for country health data. Lancet Glob Health. 2015;3:e437–8. Medline:26063456 doi:10.1016/S2214–109X(15)00067–4

journal of global health

Approaches, tools and methods used for setting priorities in health research in the 21st century

Sachiyo Yoshida

Department of Maternal, Newborn, Child and Adolescent Health, World Health Organization, Geneva, Switzerland

Background Health research is difficult to prioritize, because the number of possible competing ideas for research is large, the outcome of research is inherently uncertain, and the impact of research is difficult to predict and measure. A systematic and transparent process to assist policy makers and research funding agencies in making investment decisions is a permanent need.

Methods To obtain a better understanding of the landscape of approaches, tools and methods used to prioritize health research, I conducted a methodical review using the PubMed database for the period 2001–2014.

Results A total of 165 relevant studies were identified, in which health research prioritization was conducted. They most frequently used the CHNRI method (26%), followed by the Delphi method (24%), James Lind Alliance method (8%), the Combined Approach Matrix (CAM) method (2%) and the Essential National Health Research method (<1%). About 3% of studies reported no clear process and provided very little information on how priorities were set. A further 19% used a combination of expert panel interview and focus group discussion ("consultation process") but provided few details, while a further 2% used approaches that were clearly described, but not established as a replicable method. Online surveys that were not accompanied by face–to–face meetings were used in 8% of studies, while 9% used a combination of literature review and questionnaire to scrutinise the research options for prioritization among the participating experts.

Conclusion The number of priority setting exercises in health research published in PubMed–indexed journals is increasing, especially since 2010. These exercises are being conducted at a variety of levels, ranging from the global level to the level of an individual hospital. With the development of new tools and methods which have a well–defined structure – such as the CHNRI method, James Lind Alliance Method and Combined Approach Matrix – it is likely that the Delphi method and non–replicable consultation processes will gradually be replaced by these emerging tools, which offer more transparency and replicability. It is too early to say whether any single method can address the needs of most exercises conducted at different levels, or if better results may perhaps be achieved through combination of components of several methods.

Correspondence to:
Sachiyo Yoshida
Department of Maternal, Newborn, Child and Adolescent Health
World Health Organization, Geneva, Switzerland
Avenue Appia CH–27
Geneva, Switzerland
yoshidas@who.int

Apart from the continuing need to prioritize investments in health systems and health interventions, there is also a need to prioritize health research. Health research is difficult to prioritize, because the number of possible competing ideas for research is large, the outcome of research is inherently uncertain, and the impact of research is difficult to predict and measure [1]. A systematic and transparent process to assist policy makers and research funding agencies in making investment decisions is a permanent need.

At national level several methods have been tried: some of the best examples are the Council on Health Research for Development's approach (COHRED) in Brazil, Cameroon, Peru and Philippines; the Essential National Health Research (ENHR) approach in Cameroon and South Africa; and the Combined Approach Matrix (CAM) in Malaysia, Pakistan and Argentina [2,3]. COHRED, ENHR and CAM were all developed by committees set up by international agencies, such as the World Health Organization (WHO) or the Global Forum for Health Research (GFHR). These methods are useful for organizing the available information so that the research prioritization can take place.

To obtain a better understanding of the landscape of approaches, tools and methods used to prioritize health research I conducted a methodical review of the PubMed database covering the period 2001–2014. My primary aim was not to perform an exhaustive review of the field, which would include searching all available scientific databases and grey literature. Instead, I was interested in identifying the methods and tools that are being commonly used in the papers that are most readily accessible through databases in the public domain such as PubMed, and to assess their relative importance and applicability. The review of PubMed for the period between 2001 and 2014 achieves this aim, because this limits the search of priority–setting tools to health topics only, which is the main interest of this analysis, while drawing on a very large database which is publically available and which should contain the vast majority of relevant studies.

METHODS

My search terms included "research priorit* OR priorit* research". These terms were chosen as the most informative combination of search terms after experimenting with several versions of search terms. The search terms identified 343 publications, 138 of which were excluded from the analysis because their contents were irrelevant to health research priority setting. A further 40 studies were excluded because they were review articles which did not attempt to set priorities. In total, 165 relevant studies were identified and retained for the analysis. **Figure 1** shows a flowchart of the review on all research priority setting exercises conducted between 2001 and 2014.

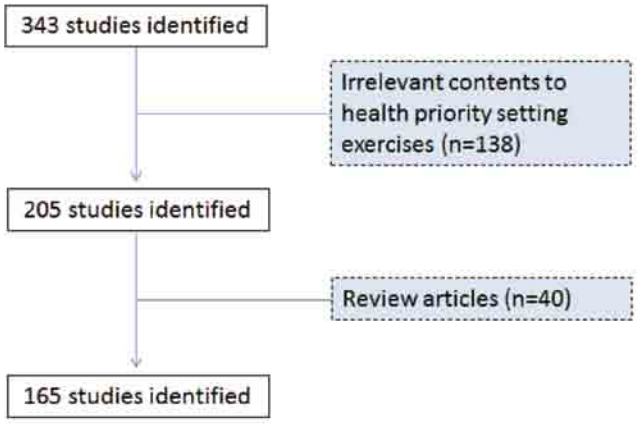


Figure 1. Flowchart of the review on all priority–setting exercises for health research conducted between 2001 to 2014.

RESULTS

Approximately 12 exercises were initiated each year between 2001 and the end of 2014. Since 2012, there has been a steady increase in the number of exercises published with the peak in 2014 with 34 exercises published (**Figure 2**). Of the 165 publications identified, the most frequently used was the CHNRI method (26%), followed by the Delphi method (24%), James Lind Alliance method (8%), the Combined Approach Matrix (CAM) method (2%) and the Essential National Health Research method (<1%). COHRED method, although frequently mentioned and clearly described in the historic context of national–level research priority setting, was not underlying any specific priority–setting process in the time period which I studied. Online surveys that were not accompanied by face–to–face meetings were used in 8% of studies, while 9% used a combination of literature review and questionnaire to scrutinise the research options for prioritization among the participating experts. About 3% of studies reported no clear process and provided very little information on how priorities were set. A further 19% used a combination of expert panel interview and focus group discussion ("consultation process") but provided few details, while a further 2% used approaches that were clearly described, but not established as a replicable method (**Figure 3**). At this point, I would like to clarify that "replicable" refers to the method's description in sufficient detail, so that all other users could apply it in the same way. It does not refer to method's property to yield the same results when repeated, which is a different meaning of the term "replicable" when assigned to a method.

Tables 1 to **6** provide a brief description of the approaches and processes used by the specific methods mentioned in **Figure 3**. The methods range from those that are not described at all, through vaguely described processes of

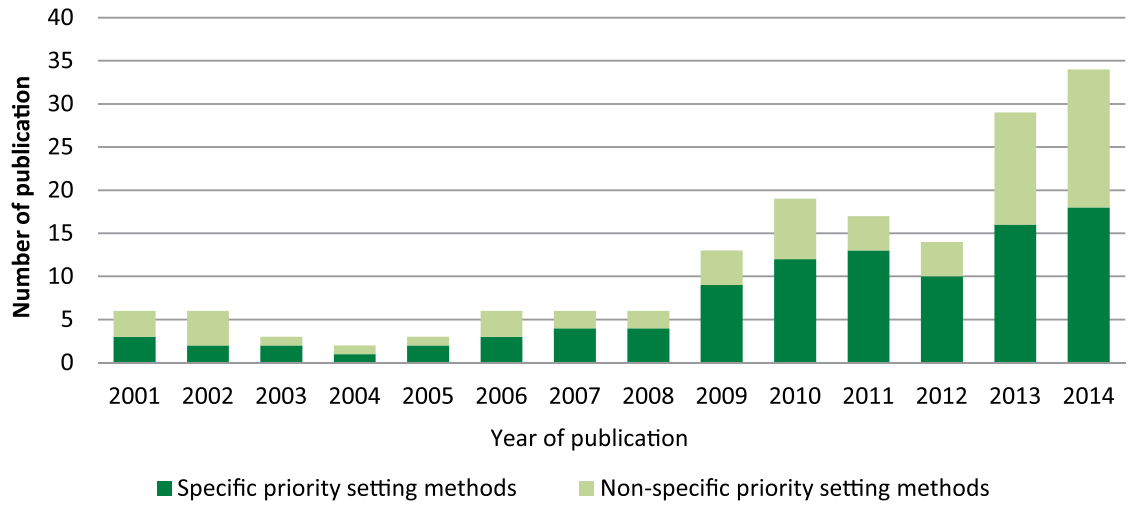


Figure 2. Total number of publication by year (source: PubMed, 2001 to 2014).

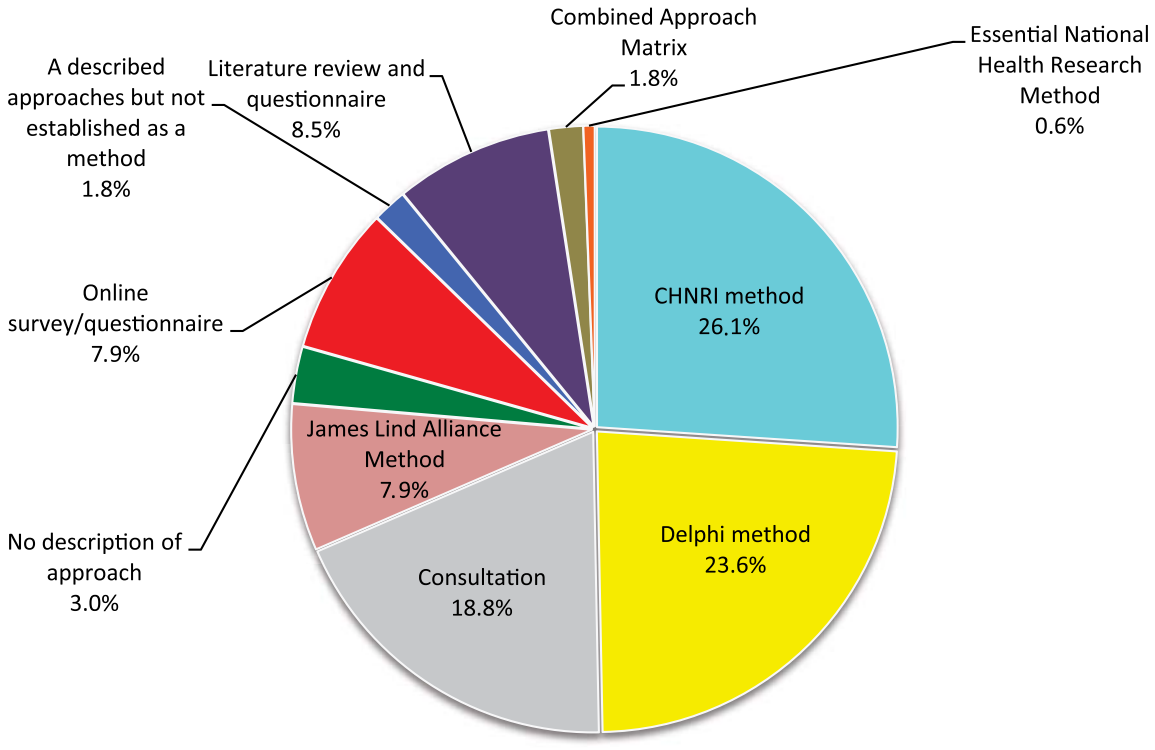


Figure 3. Methods, tools and approaches used for setting health research priorities (source: PubMed, 2001 to 2014).

group decision making, to those that follow a certain structure/process and use transparent criteria. Their output is typically quite general, ie, pointing to broad research areas in which more research activity is needed. As described above, COHRED, ENHR and CAM are used in assembling the evidence that can be used for the consultation but not for the ranking of priorities. Nevertheless, the use of any method, regardless of its limitations, is preferable to the alternative of having no clearly defined approach at all [3].

Among the 165 identified studies that set health research priorities, 21% were conducted at global level, 50% of the exercises were focused on High Income Countries (HICs)

Table 1. Brief explanation of the Essential National Health Research (ENHR) [4–6]

Overall process	ENHR was developed by Commission on Health Research for Development in 1990. It is a step by step guide for national research priority setting, focused on equity in health and development. Strategy focused on inclusiveness in participation, broad–based consultations at different levels, both quantitative and qualitative information used, and stewardship by small working group.
How are participants identified?	Participants are involved through a small representative working group which can facilitate the process, through various consultations. These stakeholders have a major stake in the goal of equity in health and development. The four major categories of participants include: researchers, decision makers, health service providers and communities.
How are research ideas identified	Stakeholders suggest priority areas, via evidence based situation analysis (such as looking at health status, health care system, health research system). Research ideas are gathered from a nomination process from different stakeholders. Consensus building using methods such as brainstorming, multi–voting, nominal group technique, round–table is then used to select research ideas.
Scoring criteria	Criteria is selected as to be: – Appropriate to the level of the action of the priority setting *i.e.* global, national, district; – Detailed in definition; – Independent of each other; – Contain information base; – Reflect equity promotion and development; – Manageable number; – Expressed in a common language. Criteria are agreed on by brainstorming of large collection of possible criteria, eliminating duplicates and clearly defining the meaning of each criterion from stakeholders. Criteria will then be put into representative categories and finally selected depending on purpose and level of action of priority–setting exercise.
Scoring options	Each criteria is scored: Point score to each criteria OR Number of score choices to each criteria
Advantages	– Broad based inclusion and participation of different stakeholders. – Multidisciplinary and cross–sectoral approach – Partnership development – Transparent process – Systematic analyses of health needs
Disadvantages	– Vague criteria and lack of transparency in individual process used by countries – Few countries had guidelines on how to develop nor apply criteria – Needs stronger representation of groups such as private sector, parliamentarians, donors, international agencies– Does not provide methodology for identifying participants

Table 2. Brief explanation of the Combined Approach Matrix (CAM) [7,8]

Overall process	Developed by the Global Forum for Health Research, CAM was to bring together economic and institutional dimensions into an analytical tool with the actors and factors that play a key role in health status of a population. It also aims to organise and present a large body of information that enters the priority setting process. This will help decision makers make rational choices in investment to produce greatest reduction in burden of disease.
How are participants identified?	Institutional approach involving: individual, household and community; health ministry and other health institutions; other sectors apart from health; and macroeconomic level actors.
How are research ideas identified	Five step process including measuring the disease burden, analysing determinants, getting present level of knowledge, evaluating cost and effectiveness, and present resource flows. For each main disease and risk factor, institutions and stakeholders with particular knowledge are brought together to provide information via workshops and brainstorming. Each institution will feed into matrix the information at disposal, regarding a specific disease or factor; the matrix will reveal how little information is available in some areas which can then be candidates for research. Each participant determined the priority research topics based on CAM evidence, then grouping the topics and cutting down to establish the top priorities.
Scoring criteria	Criteria based on questions of what is a research priority in the context, and what is not known but should be.
Scoring options	N/A
Advantages	– Creates framework of information – Identifies gaps in knowledge – Facilitates comparisons between sectors – Broad inclusion of actors – 3D–CAM includes equity
Disadvantages	– Difficult and time–consuming as involves multi–stage discussion – Does not provide algorithm to establish and score research priorities therefore is not repeatable nor systematic – Does not provide methodology for identifying participants

Table 3. Brief explanation of the James Lind Alliance Method [9]

Overall process	Focuses on bringing patients, carers and health professionals in order to identify treatment uncertainties which will become research questions. The method uses a mixture of data gathering, quantitative and qualitative analysis to create research priorities in areas of treatment uncertainty.
How are participants identified?	Participants are identified through Priority Setting Partnerships which brings patients, carers and clinicians equally together and agree through consensus priorities.
How are research ideas identified	Treatment uncertainties are defined as no up to date, reliable systematic reviews addressing treatment uncertainty, or systematic review that shows such uncertainty exists. Step 1: Recommendations by PSPs, or through looking at existing literature, creates a list of uncertainties. Step 2: These are then verified through systematic reviews of databases to verify they are research gaps using Cochrane, DARE, NICE, Sign. An uncertainty is deemed genuine when a reported confidence interval in a systematic review does not cross the line of effect or line of unity. A virtual interim priority ranking, and a final priority setting workshop takes place to agree upon 10 prioritised uncertainties through consensus building.
Scoring criteria	No clear criteria are identified with which to use.
Scoring options	Ranked AND Qualitative consensus
Advantages	– Takes into account underrepresented groups – Applicable to small scale prioritisation (eg, hospital) – Mixture of methods
Disadvantages	– Time consuming to identify and verify treatment uncertainties – Selection of criteria not clear – Not suitable for global level, nor specific disease domains – Very clinically orientated – Disproportionate mix of participants may skew information base

Table 4. Brief explanation of the Council on Health Research for Development (COHRED) [10]

Overall process	COHRED uses a management process for national level exercises to show important steps for priority setting processes.
How are participants identified?	Participants are identified through the chosen methods outlined in the steps of the COHRED guide.
How are research ideas identified	Identification of priority issues much choose method best suited to local context and needs either through compound approaches (ENHR, CAM, Burden of Disease) or foresighting techniques (Visioning, Delphi). Consider using more than one method to optimize usefulness of results.
Scoring criteria	COHRED presents ranking techniques that can be used to rank priority issues including direct and indirect valuation techniques.
Scoring options	Ranked
Advantages	– Overview approach providing steps – Discusses wide range of options – Flexible to contexts and needs
Disadvantages	– Too general and unspecific – Lack of criteria transparency

and 28% were focused on Low and Middle Income Countries (LMIC). At the national level, the countries where research priority exercises were most frequently initiated were the UK (27%), USA (16%), Australia (15%), and Canada (11%) (**Table 7**).

Topic areas for which research priorities were identified included non–communicable diseases (18%), followed by child and adolescent health (17%), mental health (10%), nursing/midwifery (8%) and infectious disease (8%). The remaining exercises (39%) covered a wide variety of topics, including policy and health system, occupational health /therapy, reproductive health/women's health, emergency care, environmental health, occupational health, forensic science and injury prevention (**Table 7**).

DISCUSSION

The number of priority setting exercises in health research published in PubMed–indexed journals is increasing, especially since 2010. These exercises are being conducted at a variety of different levels, ranging from the global level to the level of an individual hospital. With the development of new tools and methods which have a well–defined structure – such as the CHNRI method, James Lind Alliance Method and Combined Approach Matrix – it is likely that the Delphi method and non–replicable consultation processes (see the definition of "replicable" earlier in the text) will gradually be replaced by these emerging tools, which offer more transparency and replicability. This is a process that should be endorsed, as a natural progression of the

Table 5. Brief explanation of the Delphi Process [11]

Overall process	Delphi, mainly developed in the 1950s, is a systematic, interactive forecasting method which relies on a panel of experts and questionnaires.
How are participants identified?	Participants are eligible to be invited if they have related backgrounds and experiences concerning the target issue, are capable of contributing, and are willing to revise their initial judgements in order to reach consensus. Participants are considered and selected through investigators, ideally through a nomination process, or selection from potential leaders or authors through publication. It is suggested that the three groups are used: top management decision makers who will utilise outcomes of Delphi study; professional staff members and their support team; respondents to the Delphi questionnaire. It is recommended to use the minimally sufficient number to generate representative pooling of judgements – however no consensus yet as to optimal number of subjects.
How are research ideas identified	In the first round an open–ended questionnaire is sent to solicit information about a content area from Delphi participants. Investigators will then turn the responses into a well–structured questionnaire to be used as survey for data collection. Through four rounds experts answer questionnaires; the facilitator summarises anonymously the forecast after the first round and the experts are then asked to revise their earlier answer thereby decreasing the range of answers and converging towards the correct answer. Up to four iterations can be used.
Scoring criteria	N/A
Scoring options	Rate or ranking AND Consensus building
Advantages	– Multiple iterations and feedback process – Flexible to change – Anonymity of respondents
Disadvantages	– Does not provide methodology for identifying participants – Lack of criteria transparency – Potential for low response rate due to multiple iterations – Time–consuming – Potential for investigators and facilitators to bias opinions

Table 6. Brief explanation of the CHNRI process [12–15]

CHNRI method Child Health Nutrition Research Initiative	
Overall process	The CHNRI methodology was introduced in 2007 by the Child Health and Nutrition Research Initiative of the Global Forum for Health research. The methodology was developed to address gaps in the existing research priority methods. The CHNRI method is developed to assist decision making and consensus development. The method include soliciting ideas from different carder of participants on the given health topic and use independent ranking system against the pre–defined criteria to prioritise the research ideas.
How are participants identified?	Participants are identified by management team based on their expertise (eg, number of publications, experience in implementation research and programmes etc). Participants includes stakeholders who might not have the technical expertise but have view on the health topic of concern.
How are research ideas identified?	Research ideas are generated by participants or by management team based on the current evidence. If former, usually each participant is asked to provide maximum of three research questions against the predefined domain of health research (eg, descriptive research, development research, discovery research and delivery research). The ideas are usually submitted via online survey and consolidated by the management team.
Scoring criteria	Five standard criteria are usually used: – Answerability – Equity – Impact on burden – Deliverability – Effectiveness. Though the five standard criteria are used in more than 70% of the research priority setting exercises, the method offers optional criteria to be used to replace the standard criteria depending on the needs and context of the exercises. For example, criteria such as low cost, sustainability, acceptability, feasibility, innovation and originality are used to replace or in addition to the standard criteria.
Scoring options	Each criteria is scored: Point score to each criteria in the scale of 0, 0.5 and 1 or in the scale of 0 to 100.
Advantages	– Simple, inclusive and replicable and thus systematic and transparent process. – Independent ranking of experts (avoid having the situation where one strongly minded individual affecting the group decision) – Less costly
	– Potentially represent collective opinion of the limited group of people who were included in the process. – Scoring affected by currently on–going research

Table 7. Distribution of identified studies by geographic context and countries where the research priority setting exercises have been initiated and research priority areas addressed

Geographical area	Number	%	Technical areas	Number	%
Global	35	21	Non–communicable disease	29	18
High income countries	82	50	Child and adolescent health	28	17
Low middle income countries	47	28	Mental health	16	10
Humanitarian settings	1	<1	Infectious disease	14	8
TOTAL	**165**	**100**	Nursing/Midwifery	13	8
National level			Public health in general	10	6
Australia	15	15	Policy and health system	8	5
Brazil	1	1	Occupational health/therapy	6	4
Canada	11	11	Reproductive health/women's health	6	4
Colombia	1	1	Skin disease	5	3
Chile	1	1	Emergency care	3	2
Cuba	1	1	Environmental health	3	2
Hong Kong	2	2	Disability	3	2
India	1	1	Child development potential	2	1
Iran	2	2	Injury prevention	2	1
Ireland	3	3	Maternal and perinatal health	2	1
Italy	1	1	Pharmaceuticals	2	1
Malaysia	1	1	Microbial Forensics	2	1
Nepal	1	1	Behavioural science	1	1
The Netherlands	1	1	Diagnostic accuracy	1	1
Nigeria	1	1	Tuberculosis	1	1
Peru	1	1	Medical science	1	1
Portugal	2	2	Neurological	1	1
South Africa	3	3	Nutrition	1	1
Saudi Arabia	1	1	Surgical	1	1
Spain	3	3	Surveillance system	1	1
United Republic of Tanzania	2	2	Water and sanitation	1	1
United Kingdom	26	27	Primary health care–related disease	1	1
United States of America	16	16	Others	1	1
TOTAL	**97**	**100**	**TOTAL**	**165**	**100**

priority–setting field from the period in which hardly any structured processes existed to fill a need, to the new era which will be increasingly dominated by structured and well–defined tools.

This review is not the first attempt to assess approaches, tools and methods to set health research priorities. Searching the literature, I identified five earlier attempts to review and discuss priority–setting processes. The first review was published by Rudan and colleagues in 2007 in an attempt to develop an evidence base for the development of conceptual framework and guidelines for implementation of the CHNRI methodology [1]. This paper identified ambitious attempts by several large organizations at the international level to define health research priorities for either the whole developing world, large world regions or nationally. These attempts date back to the year 1990, with the "... Commission on Health Research for Development usually being referred to as the first truly significant international initiative aimed toward systematic approach to setting priorities in global health research." Other initiatives that followed were the "Ad Hoc Committee (AHC) on Health Research Relating to Future Intervention Options" (in 1994), the "Global Forum for Health Research" (in 1998), the "Council on Health Research and Development (COHRED)" (in 2000), "The Grand Challenges" proposed at the World Economic Forum in Davos, Switzerland (in 2003) and the "Combined Approach Matrix" as the first specific priority–setting tool for health research (in 2004). The paper concluded that the processes, initiatives and tools fell short of being informative on what the specific research priorities should be and how exactly are they derived [1].

In 2010, Viergever et al. [16] reviewed the articles that set health research priorities and they specifically reviewed exercises coordinated by World Health Organization Headquarters since 2005. This resulted in the total of 230 documents or reports, many of them unpublished (hence, not included in my review). The authors concluded that, at that point in time, there was no "gold standard" approach for health research prioritisation. This was not surprising, given the heterogeneity in the context of research prioritization exercises and different levels at which they were being conducted. Nevertheless, the authors observed several

common themes of "good practice" and proposed a generic framework – in the form of "checklist", like a form of "guidelines" – which also suggested various options for each step of the process. Nine themes were identified through a review of the previously conducted priority–setting processes. They were categorized as the "themes during the preparatory work" (defining context, use of comprehensive approach, ensure inclusiveness of participants, information gathering, planning for implementation), followed by the steps in the process of deciding on the priorities (defining the criteria, methods for deciding on priorities), and two steps in the last phase after the priorities have been set (plan the timing of evaluation in terms of how the research priorities are being used, and write the clear report of the methodology used to ensure the transparency in the process). The authors proposed that the provision of the framework should be of assistance to policy makers and researchers. It could have a dual role: it could not only assist priority–setting process, but also planning the follow up and implementation of the priorities [16].

In the same year, in 2010, the World Health Organization's Department for Research Policy and Cooperation held a consultation between methodology–developing experts to identify optimal characteristics of priority–setting methods that could be applicable at the national level. The aim was to empower low and middle–income countries to take more ownership of their own health research agenda. Tomlinson reviewed the progress made at this meeting and published the main conclusions in 2011 [2]. Three methods emerged as applicable at the national level: the Combined Approach Matrix (CAM), the Council on Health Research and Development (COHRED) and the Child Health and Nutrition Research Initiative (CHNRI). The authors presented and discussed strengths and weaknesses of each method [2]. They also noted that, across the countries surveyed, genuine engagement of stakeholders was difficult to achieve and was typically missing. Countries also varied in the extent to which they would document priority–setting processes, with not a single country having an appeal process for outlined priorities. Another problem was that the identified priorities usually outlined broad disease categories, rather than more specific research questions [2]. The authors concluded that priority–setting processes should aim to include mechanisms for publicizing results, effective procedures to translate and implement decisions and processes to ensure that the revision of priorities eventually does occur.

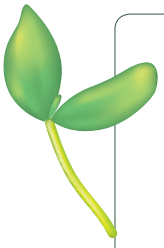
In a more recent report, an independent team from the Kirby Institute in Sydney, Australia, systematically reviewed all studies undertaken in low– and middle–income country (LMIC) settings that attempted to set research priorities over the period from 1966 to 2014. The studies included were not reported but they found 91 studies, including 16 which used the CHNRI method [17]. The authors concluded that almost half of these processes took place at the global level (46%). For regional or national initiatives, a half focused on Sub Saharan Africa (49%), followed by East Asia and the Pacific (20%) and Latin America and the Caribbean (18%). Most commonly, studies were initiated by an international organization or collaboration (46%). Researchers and governments were the most commonly represented stakeholders. The most frequently used process was a conference or workshop to determine priorities (24%), followed by the CHNRI method (18%) [17]. The review revealed inconsistent use of existing methods and approaches in health research prioritization processes. It also showed that while there was strong involvement of government and researchers, participation of other key stakeholders was limited. The authors argued that many processes, regardless of the method used, lacked an implementation strategy to translate the result of the process into implementation of research projects. Finally, the authors concluded that research prioritization exercises would often remain "one–time exercises", given the lack of follow up and implementation strategies involving the funders, researchers and government officials.

Finally, in 2014, as a part of the Lancet series on increasing value and reducing waste in health research generally, one paper of the series (by Chalmers et al. [18]) explored how to increase value and reduce waste when research priorities are set. The group of authors argued that many basic research endeavours do not lead to knowledge that is useful to the end user of the research results. By using various examples, the authors reiterate the same argument: if research does not meet the needs of the users of research, evidence will have little impact on public health and clinical practice. The authors argue that many research studies that fall in the area of basic (fundamental) research were duplicative. Although a replication of positive findings is a welcome process, an excessive repetition of conducting similar research can be prevented by either: (i) conducting systematic reviews and also involving the end user of the research as well as clinicians in the process (where they used the example of hospital based research priority setting exercise using the James Lind Alliance method); and (ii) mapping research portfolios of major agencies, that could help to prevent duplication in the nature of supported research. The main message of the article is, therefore, a need for better co–ordination among the researchers and the funders over the research that is being conducted and increased focus on the translational value of the information that is being generated through research [18].

It is evident from my own methodical review, and from the systematic review undertaken by the researchers from the Kirby Institute, that there is a need for a transparent, replicable, systematic and structured approach to research pri-

ority setting, because the large majority of the previous exercises were not based on processes meeting all of these criteria. The review by McGregor et al. [17] shows how, although a very recent addition to the set of tools, the CHNRI method is set to become the most widely used approach.

The results of my review broadly confirmed the observations of all previous reviews, with an additional insight into time trend – showing an increase in the number of exercises conducted over time, and gradual replacement of poorly defined processes with those that use particular methods and tools, as shown in **Figure 2**. The next step in the field of health research priority setting should therefore involve monitoring whether any single method may address the need for most exercises conducted at different levels, or if better results may perhaps be achieved through combination of strengths of several methods.



Acknowledgements: The author thanks James Gao, an intern at the World Health Organization, for his assistance to the development of Tables 1–6. She also acknowledges helpful input from her mentor, Professor Simon Cousens from the London School of Hygiene and Tropical Medicine.

Disclaimer: SY is an employee of the World Health Organization. The views expressed in this paper are the responsibility of the author and do not necessarily represent the views of the World Health Organization.

Funding: None.

Ethical approval: Not required.

Authorship declaration: SY wrote the paper as a part of her PhD thesis.

Declaration of interest: The author completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). The author reports no conflicting financial or other interest related to the work detailed in this manuscript.

REFERENCES

1 Rudan I, Gibson J, Kapiriri L, Lansang MA, Hyder AA, Lawn J, et al; Child Health and Nutrition Research Initiative (CHNRI). Setting priorities in global child health research investments: Assessment of principles and practice. Croat Med J. 2007;48:595-604. Medline:17948946

2 Tomlinson M, Chopra M, Hoosain N, Rudan I. A review of selected research priority setting processes at national level in low and middle income countries: towards fair and legitimate priority setting. Health Res Policy Syst. 2011;9:19. Medline:21575144 doi:10.1186/1478-4505-9-19

3 Rudan I, Kapiriri L, Tomlinson M, Balliet M, Cohen B, Chopra M. Evidence–based priority setting for health care and research: tools to support policy in maternal, neonatal, and child health in Africa. PLoS Med. 2010;7:e1000308. Medline:20644640 doi:10.1371/journal.pmed.1000308

4 Council on Health Research for Development (COHRED). Essential national health research and priority setting: Lessons learned; 1997. Available: http://www.cohred.org/downloads/586.pdf. Accessed: 21 August 2015.

5 Okello D, Chongtrakul P; The COHRED Working Group on Priority Setting. A manual for research priority setting using the ENHR Strategy; 2000. http://www.cohred.org/downloads/578.pdf. Accessed: 21 August 2015.

6 The Working Group on Priority Setting. Priority setting for health research: lessons from developing countries. Health Policy Plan. 2000;15:130-6. Medline:10837035 doi:10.1093/heapol/15.2.130

7 Ghaffar A, Collins T, Matlin SA, Olifson S. The 3D combined approach matrix: An improved tool for setting priorities in research for health. The Global Forum for Health Research. 2009. Available: http://www.webcitation.org/query.php?url=http://www.globalforumhealth.org/content/download/7860/50203/file/CAM_3D_GB.pdf&refdoi=10.1186/1478–4505–8–36. Accessed: 21 August 2015.

8 Ghaffar A, de Francisco A, Matlin S. The Combined Approach Matrix: a priority–setting tool for health research. Geneva: Global Forum for Health Research; 2004. Available: http://www.webcitation.org/query.php?url=http://www.globalforumhealth.org/Media–Publications/Publications/The–Combined–Approach–Matrix–A–priority–setting–tool–for–health–research&refdoi=10.1186/1478–4505–8–36. Accessed: 21 August 2015.

9 Cowan K, Oliver S. The James Lind Alliance Guidebook (version5). The James Lind Alliance; 2013. Available: http://www.jlaguidebook.org/pdfguidebook/guidebook.pdf. Accessed: 21 August 2015.

10 Montorzi G, Haan S, IJsselmuiden C. Priority setting for research for health: A Management process for countries. Council on Health Research for Development (COHRED); 2010. Available: http://www.cohred.org/downloads/Priority_Setting_COHRED_approach_August_2010.pdf. Accessed: 21 August 2015.

11 Hsu C, Sandford B. The Delphi technique: Making sense of consensus. Pract Assess, Res Eval. 2007;12:1-8.

12 Kapiriri L, Tomlinson M, Gibson J, Chopra M, El Arifeen S, Black RE, et al. Child Health and Nutrition Research Initiative (CHNRI): Setting priorities in global child health research investments: Addressing the values of the stakeholders. Croat Med J. 2007;48:618-27. Medline:17948948

REFERENCES

13 Rudan I, Chopra M, Kapiriri L, Gibson J, Lansang MA, Carneiro I, et al. Setting priorities in global child health research investments: universal challenges and conceptual framework. Croat Med J. 2008;49:307-17. Medline:18581609 doi:10.3325/cmj.2008.3.307

14 Rudan I, Gibson JL, Ameratunga S, El Arifeen S, Bhutta ZA, Black M, et al; Child Health and Nutrition Research Initiative (CHNRI). Setting Priorities in Global Child Health Research Investments: Guidelines for implementation of the CHNRI Method. Croat Med J. 2008;49:720-33. Medline:19090596 doi:10.3325/cmj.2008.49.720

15 Rudan I, El Arifeen S, Black RE, Campbell H. Childhood pneumonia and diarrhoea: Setting our priorities right. Lancet Infect Dis. 2007;7:56-61. Medline:17182344 doi:10.1016/S1473-3099(06)70687-9

16 Viergever RF, Olifson S, Ghaffar A, Terry RF. A checklist for health research priority setting: nine common themes of good practice. Health Res Policy Syst. 2010;8:36. Medline:21159163

17 McGregor S, Henderson KJ, Kaldor JM. How are health research priorities set in low and middle income countries? A systematic review of published reports. PLoS ONE. 2014;9:e108787. Medline:25275315 doi:10.1371/journal.pone.0108787

18 Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gülmezoglu AM, et al. How to increase value and reduce waste when research priorities are set. Lancet. 2014;383:156-65. Medline:24411644 doi:10.1016/S0140-6736(13)62229-1

Electronic supplementary material:
The online version of this article contains supplementary material.

journal of global health

Setting research priorities to improve global newborn health and prevent stillbirths by 2025

Sachiyo Yoshida[1], José Martines[2], Joy E Lawn[3,4], Stephen Wall[4], João Paulo Souza[5], Igor Rudan[6], Simon Cousens[3]; The neonatal health research priority setting group; Peter Aaby[7], Ishag Adam[8], Ramesh Kant Adhikari[9], Namasivayam Ambalavanan[10], Shams El Arifeen[11], Dhana Raj Aryal[12], Sk Asiruddin[13], Abdullah Baqui[14], Aluisio JD Barros[15], Christine S Benn[16], Vineet Bhandari[17], Shinjini Bhatnagar[18], Sohinee Bhattacharya[19], Zulfiqar A Bhutta[20], Robert E Black[21], Hannah Blencowe[22], Carl Bose[23], Justin Brown[24], Christoph Bührer[25], Wally Carlo[26], Jose Guilherme Cecatti[27], Po–Yin Cheung[28], Robert Clark[29], Tim Colbourn[30], Agustin Conde–Agudelo[31], Erica Corbett[32], Andrew E Czeizel[33], Abhik Das[34], Louise Tina Day[35], Carolyn Deal[36], Ashok Deorari[37], Uğur Dilmen[38], Mike English[39], Cyril Engmann[40], Fabian Esamai[41], Caroline Fall[42], Donna M Ferriero[43], Peter Gisore[44], Tabish Hazir[45], Rosemary D Higgins[46], Caroline SE Homer[47], DE Hoque[48], Lorentz Irgens[49], MT Islam[50], Joseph de Graft–Johnson[51], Martias Alice Joshua[52], William Keenan[53], Soofia Khatoon[54], Helle Kieler[55], Michael S Kramer[56], Eve M Lackritz[57], Tina Lavender[58], Laurensia Lawintono[59], Richard Luhanga[60], David Marsh[61], Douglas McMillan[62], Patrick J McNamara[63], Ben Willem J Mol[64], Elizabeth Molyneux[65], G. K Mukasa[66], Miriam Mutabazi[67], Luis Carlos Nacul[68], Margaret Nakakeeto[69], Indira Narayanan[70], Bolajoko Olusanya[71], David Osrin[72], Vinod Paul[73], Christian Poets[74], Uma M Reddy[75], Mathuram Santosham[76], Rubayet Sayed[77], Natalia E Schlabritz–Loutsevitch[78], Nalini Singhal[79], Mary Alice Smith[80], Peter G Smith[81], Sajid Soofi[82], Catherine Y Spong[83], Shahin Sultana[84], Antoinette Tshefu[85], Frank van Bel[86], Lauren Vestewig Gray[87], Peter Waiswa[88], Wei Wang[89], Sarah LA Williams[90], Linda Wright[91], Anita Zaidi[92], Yanfeng Zhang[93], Nanbert Zhong[94], Isabel Zuniga[95], Rajiv Bahl[1]

Affiliations of authors are listed at the end of the manuscript.

Correspondence to:
Dr Rajiv Bahl
Department of Maternal, Newborn,
Child and Adolescent Health
World Health Organization
Geneva, Switzerland
20 Avenue Appia 1211
Geneva–27, Switzerland
bahlr@who.int

Background In 2013, an estimated 2.8 million newborns died and 2.7 million were stillborn. A much greater number suffer from long term impairment associated with preterm birth, intrauterine growth restriction, congenital anomalies, and perinatal or infectious causes. With the approaching deadline for the achievement of the Millennium Development Goals (MDGs) in 2015, there was a need to set the new research priorities on newborns and stillbirth with a focus not only on survival but also on health, growth and development. We therefore carried out a systematic exercise to set newborn health research priorities for 2013–2025.

Methods We used adapted Child Health and Nutrition Research Initiative (CHNRI) methods for this prioritization exercise. We identified and approached the 200 most productive researchers and 400 program experts, and 132 of them submitted research questions online. These were collated into a set of 205 research questions, sent for scoring to the 600 identified experts, and were assessed and scored by 91 experts.

Results Nine out of top ten identified priorities were in the domain of research on improving delivery of known interventions, with simplified neonatal resuscitation program and clinical algorithms and improved skills of community health workers leading the list. The top 10 priorities in the domain of development were led by ideas on improved Kangaroo Mother Care at community level, how to improve the accuracy of diagnosis by community health workers, and perinatal audits. The 10 leading priorities for discovery research focused on stable surfactant with novel modes of administration for preterm babies, ability to diagnose fetal distress and novel tocolytic agents to delay or stop preterm labour.

Conclusion These findings will assist both donors and researchers in supporting and conducting research to close the knowledge gaps for reducing neonatal mortality, morbidity and long term impairment. WHO, SNL and other partners will work to generate interest among key national stakeholders, governments, NGOs, and research institutes in these priorities, while encouraging research funders to support them. We will track research funding, relevant requests for proposals and trial registers to monitor if the priorities identified by this exercise are being addressed

About 2.9 million newborns died in 2011, accounting for 44% of the world's under-5 child deaths [1]. The proportion of neonatal mortality continues to increase because the neonatal mortality rate is declining at a slower rate than the mortality rates for older children [1]. Moreover, 2.7 million stillbirths occur each year, at least 40% of which occur during labour [2]. The leading killers of newborns are preterm birth complications, intrapartum–related events and neonatal infections such as pneumonia, sepsis or meningitis [3]. A high proportion of stillbirths, neonatal and also maternal deaths happen at birth and during the first days after birth – a total of over 3 million deaths [4]. This is also a critical time window to address acute morbidity and long–term impairment associated with preterm birth, intrauterine growth restriction (IUGR), congenital abnormalities, and perinatal or infectious insults [5,6].

With the approaching deadline for the achievement of the Millennium Development Goals (MDGs) in 2015, and the creation of new framework for development goals [7], there is an increasing need to guide the limited research capacity and funding to obtain the maximum impact on maternal and child health. Hence the World Health Organization (WHO) has initiated a set of global research priority–setting exercises in 2007–2008 for improving health of mothers, newborns, children and adolescents [8–12]. The five–year evaluation of that exercise from the perspective of donors, policy–makers and researchers is currently under way and it is showing an increased focus on identified research priorities from all three groups of stakeholders – in terms of investments by the donors [13,14], initiatives launched by policy–makers [15–19] and publication output from researchers [2,20–23], respectively. As part of this initiative, the Department of Maternal, Newborn, Child and Adolescent Health undertook this exercise for setting research priorities in newborn health and stillbirth, in collaboration with Saving Newborn Lives (SNL), a program of Save The Children. The time frame for the expected impact of the research extends to 2025 to allow for medium term and long–term research investments to also be considered. Alongside the persisting urgency of reducing mortality and the findings from previous research priority exercises the group believed that the research should also address morbidity, development, and long–term sequelae of preterm birth, small for gestational age as well as other hypoxic or infectious insults in the neonatal period (**Box 1**). In the exercise, we focused on intrapartum stillbirth as a high proportion of stillbirths occurs during the labour.

Box 1 The purpose and remit of this research priority setting exercise

Population of interest:
Newborns and stillbirths, survival and health, preterm birth, growth and impairment–free development

Time frame:
2013–2025, reaching beyond the timeframe of the Millennium Development Goals

Research domains:
DISCOVERY (new interventions)
DEVELOPMENT (improved interventions)
DELIVERY (implementation of existing interventions)
(note: not including description eg, epidemiology)

Audience (stakeholders):
Governments, researchers in low and middle–income countries, international donors

METHODS

A working group that managed the agenda–setting process consisted of staff responsible for newborn health in WHO and Saving Newborn Lives. The group defined the scope of the priority setting exercise (**Box 1**). Methodology developed by the Child Health and Nutrition Research Initiative (CHNRI) was adapted and used for this priority setting exercise, to enable systematic listing and transparent scoring of many competing research questions [24–26]. This methodology had been used in the previous priority setting exercises by the WHO on five major causes of child deaths: pneumonia, diarrhea, preterm birth and low birth weight, neonatal infections, and birth asphyxia [8–12]. The previous exercise coordinated by the WHO was sharply focused on short–term gains, ie, within the MDG4 target of the year 2015. In addition, the CHNRI methodology has been used by many other subject groups and multiple organizations [27–33]. **Box 2** shows the steps we followed during this priority setting process.

A large group of researchers and program experts were identified and asked to submit three ideas for improving newborn health outcomes by 2025 (**Box 2**). Two hundred of the most productive researchers, representing a broad range of technical expertise and regional diversity, identified through Web of Science® ranking tools, were invited by email to propose research questions on newborn health and birth outcomes. A further 400 program experts in newborn health programmes were also invited to propose research questions.

The proposed research questions and scoring criteria were refined by a small group of 14 experts who were invited by the WHO to participate in a two–day workshop. Each question was assigned to a domain and a technical area. The first of the three domains was "discovery", which included research aimed at finding new solutions such as new medicines, vaccines or other preventive interventions, or new diagnostics. The second domain was "development", which included research questions aimed at improving existing interventions, reducing their costs or mak-

Box 2. Adapted Child Health and Nutrition Research Initiative's (CHNRI) methodology applied to set newborn research priorities

1. Selection of individuals to submit ideas and to score questions:

Individuals representing a wide range of technical expertise in the area of newborn health and birth outcomes were selected by including

- Top 100 most productive researchers in the previous 5 years (2008–2012), according to the Web of Science®, in any research that involved neonates anywhere in the world, including (but not limited to) fundamental research, obstetrics and gynaecology, social science, and other fields;
- Top 50 most productive researchers in the previous 5 years (see above) in research specifically involving neonates in low and middle income countries (LMICs);
- Top 50 most productive researchers in the previous 5 years (see above) in any research involving stillbirths;
- 400 program experts in newborn health, who were contacted through the Healthy Newborn Network Database, representing mainly national–level health programme managers in LMICs.

2. Identification of questions to be scored:

All the identified individuals were approached and asked to submit their three most promising ideas for improving newborn health outcomes by 2025. An expert group meeting was convened to review the 396 questions received from 132 experts. After removing or merging seemingly duplicate ideas, the submissions were consolidated into a set of 205 research questions and clarity of the questions was improved.

3. Scoring of research questions:

A set of 5 criteria to assess the proposed 205 research questions was agreed on.

The scoring criteria were based on CHNRI methodology [8–12]

i. Likelihood of answering the question in an ethical way
ii. Likelihood of efficacy
iii. Likelihood of deliverability and acceptability
iv. Likelihood for an important disease burden reduction
v. Predicted effect on equity

During the preliminary meeting, 14 experts invited from the larger pool of responders completed their scoring to test the methodology. The remaining experts were asked independently to answer a set of questions via an online survey on all the chosen criteria for all listed research options. Scores from a total of 91 experts were received.

4. Computation of scores for competing research options and ranking:

The intermediate scores were computed for each of the five criteria and they could potentially range between 0–100%. Those scores indicate the "collective optimism" of the group of scorers that a given research question would fulfil each given criterion. The overall research priority score for each research question was then computed as the mean of the intermediate scores. The average expert agreement scores were also calculated (**Online Supplementary Document**).

ing them simpler to deliver. The third domain was "delivery", which included research questions that would help deliver existing interventions to more mothers and newborns with high quality. The five separate technical areas included: (i) preterm birth; (ii) intrapartum–related events including intrapartum stillbirths; (iii) newborn infections; (iv) congenital malformations and other specific conditions; and (v) integrated care including the care for mothers and neonates;

The final list of research questions and scoring criteria were sent to the original group of 600 experts with an invitation to score them. Each research question was assessed by the expert and received a score of 1.0, 0.5 or 0 for five preset criteria, with the option of not assigning any score in case the expert did not feel confident to decide on that criterion. Scoring took place over eight weeks and was conducted and returned to the coordinators at the WHO by 91 experts.

Intermediate scores for each research question against the 5 criteria were computed as the sum of the scores for that particular criterion divided by the total number of scorers. This resulted in a number between 0–100% that captured the "collective optimism" of the group of 91 scorers that a given research question would fulfill each given criterion. The overall research priority score (RPS) for each research question was then computed as the mean of the intermediate scores calculated for each of the five criteria: RPS = [(Criterion 1 score %) +(Criterion 2 score %)+(Criterion 3 score %)+(Criterion 4 score %)+(Criterion 5 score %)]/5. The confidence interval was calculated using the bootstrapping methods in STATA version 11.2.

RESULTS

In total, 132 of the 600 invited experts proposed a total of 396 research questions, which were then checked for similarity and consolidated in a final list of 205 questions to be scored. The characteristics of respondents are summarized in **Figure 1**. The 205 research questions were then scored by 91 experts. About 40% of the scorers were based in low and middle income countries (LMICs) in Africa, Asia, and South America. About two–thirds (65%) worked in academic or research institutions and the remainder was divided between program managers (16%), clinicians (7%), donor representatives (7%) and policy makers (5%) (**Figure 1**).

The overall research priority scores given to the 205 proposed questions ranged from 90% (high) to 47% (low; full list of scored questions is presented in the **Online Supplementary Document**). The level of agreement between the 91 experts ranged from 77% (high) to 34% (low), suggesting that on average, for each research question of interest, between three–quarters and one–third of the scorers were in agreement in their responses to each criterion.

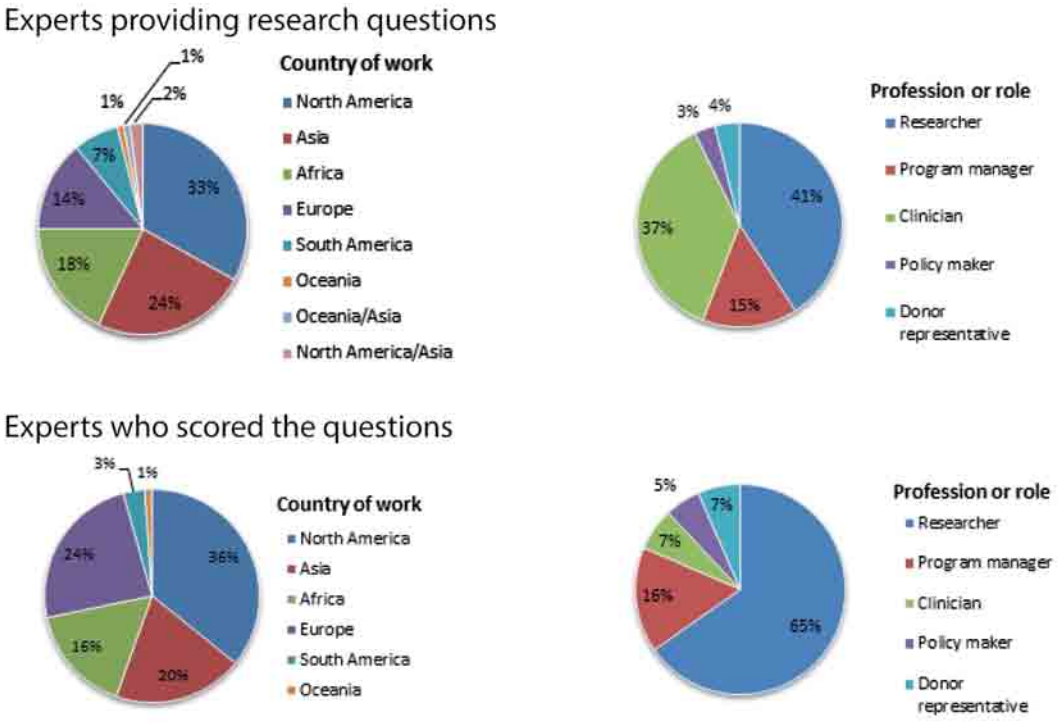


Figure 1. Background characteristics of 132 experts who provided questions and 91 experts who scored the questions.

The overall scores for the highest priority questions ranged from 79% to 90% (**Table 1**). Agreement scores indicated that more than two thirds of the experts had a common view towards the list of research priorities. Nine of the ten top priorities were in the domain of "delivery", with simplified neonatal resuscitation programs and clinical algorithms and improved skills of community health workers leading the list. Among the 11 priorities shown in this table, three addressed preterm birth, four addressed intrapartum–related events and four addressed newborn infections.

In the domain of "development", the top 10 priorities (**Table 2**) were ranked between 8th and 50th on the list of all research questions (displayed in full in **Online Supplementary Document**). They were led by ideas on improved Kangaroo Mother Care, improve accuracy of diagnosis by community health workers, and perinatal audits. Two priorities among the leading ten in this domain were identified in each of the areas of preterm birth, intrapartum related events and newborn infections, while the remaining 4 priorities related to integrated care.

The 10 leading priorities for discovery research (**Table 3**) ranked between 55th and 129th on the list of all research questions (see **Online Supplementary Document**) and they focused on stable surfactant with novel modes of administration, ability to diagnose fetal distress and novel tocolytic agents. Agreement scores for the ten leading questions ranged from 42% to 49%. Three priorities were identified in each of the areas of preterm birth and newborn infections, two on preventing intrauterine growth restriction and one each on intrapartum–related events and antepartum stillbirths.

There was a remarkable similarity in the scoring pattern between experts from a research background and those from a program background for the top 10 ranked priorities (**Table 4**). The programme experts had a tendency to assign somewhat higher overall scores to "delivery" questions, which was mediated through their higher scoring of maximum potential impact and equity criteria. Among "development" questions, the scorers with a background in research gave higher scores for efficacy and deliverability, while programme experts gave higher scores for impact and equity criteria. Surprisingly, the scoring pattern of both groups of experts for "discovery" questions was very similar, both for overall score and for each of the 5 criteria.

DISCUSSION

In this paper, we present global research priorities that have the potential to impact mortality, morbidity, child development, and long–term health outcomes among neonates in the period between 2013–2025. Despite the broad focus on these outcomes and a 12–year timeline, "delivery" questions received highest scores, followed by "development" and "discovery" questions, as was the case in previous exercises with shorter time lines focusing only on reducing mortality [8–12].

Table 1. Top ten research priorities for improving newborn health and birth outcomes by 2025 as ranked by 91 experts

Rank	Research questions	Domain	Total score (confidence interval)	Agreement between scorers	Answerable?	Efficacy?	Deliverability?	Impact?	Equity?
1	Can simplified neonatal resuscitation program delivered by trained health workers reduce neonatal deaths due to perinatal asphyxia?	Delivery	90 (85–91)	77	96	91	94	77	92
2	How can the health worker's skills in preventing and managing asphyxia be scaled up?	Delivery	88 (83–89)	74	96	91	89	75	86
3	Can simple clinical algorithms used by CHW identify and refer neonates with signs of infection and consequently reduce newborn mortality?	Delivery	86 (83–89)	72	92	92	92	66	88
4	How can exclusive breastfeeding in low–resource contexts be promoted to reduce neonatal infections and mortality?	Delivery	85 (79–89)	72	94	89	86	69	86
5	Can the training of CHWs in basic newborn resuscitation reduce morbidity and mortality due to perinatal asphyxia?	Delivery	83 (78–86)	67	94	84	84	64	88
6	How can the administration of injectable antibiotics at home and first level facilities to newborn with signs of sepsis be scaled up to reduce neonatal mortality?	Delivery	82 (78–86)	64	89	88	88	59	84
7	Can community–based initiation of Kangaroo Mother Care reduce neonatal mortality of clinically stable preterm and low birth weight babies?	Development	80 (74–84)	66	86	87	81	69	77
8	How can facility based initiation of Kangaroo Mother Care or continuous skin–to–skin contact be scaled up?	Delivery	80 (71–84)	62	90	82	84	62	81
9	How can chlorhexidine application to the cord be scaled up in facility births and in low NMR setting to reduce neonatal infections and neonatal mortality?	Delivery	80 (70–83)	67	91	85	89	52	81
10	How can quality of care during labour and birth be improved to reduce intrapartum stillbirths, neonatal mortality and disability?	Delivery	79 (71–82)	65	83	84	82	72	75
11*	Can community based "extra care" for preterm/LBW babies delivered by CHWs reduce neonatal morbidity and mortality in settings with poor accessibility to facility care?	Delivery	79 (70–82)	63	87	87	81	62	81

*The overall and criterion specific scores ranged from 0% to 100%.The 11th question added to complete the list of top 10 priorities in the domain of "delivery". The question originally ranked 5th was omitted from this table because it was a variant of question that already received a higher overall score.

Table 2. Top ten development research priorities for improving newborn health and birth outcomes by 2025 as ranked by 91 experts

Rank	Research questions	Total score (confidence interval)	Agreement between scorers
8*	Can community–based initiation of Kangaroo Mother Care reduce neonatal mortality of clinically stable preterm and low birth weight babies?	82 (78–86)	64
26	How can the accuracy of community health workers in detecting key most important high risk conditions or danger signs in pregnant women be improved?	77 (70–80)	61
35	Can perinatal audits improve quality of care in health facilities and improve fetal and neonatal outcomes?	74 (67–79)	58
37	Can intrapartum monitoring to enhance timely referral improve fetal and neonatal outcomes?	74 (67–79)	57
38	Can training community health workers to recognize and treat neonatal sepsis at home with oral antibiotics when referral is not possible reduce neonatal mortality?	74 (62–78)	57
40	Can oral amoxicillin at home for treatment of neonatal pneumonia reduce neonatal mortality?	73 (64–78)	58
43	Can models for strengthening capacity of health Professionals in caring for neonates in peripheral hospitals improve neonatal outcomes?	73 (63–77)	54
44	Can intervention package for CHWs to prevent and manage perinatal asphyxia be delivered by community health workers?	72 (64–77)	55
47	Can low–cost devices for facility care of newborns be developed and tested for the effectiveness at various levels of the health system (eg, CPAP devices, syringe drivers, IV giving sets, phototherapy units, oxygen concentrators, oxygen saturation monitors incubators, ventilators, therapeutic hypothermia technology) ?	72 (65–76)	53
50	Can surfactant reduce preterm morbidity and mortality in low and middle income countries?	72 (65–78)	56

*Also in the overall top 10 priorities.

Table 3. Top ten discovery research priorities in discovery for improving newborn health and birth outcomes by 2025 as ranked by 91 experts

Rank	Research questions	Total score (confidence interval)	Agreement between scorers
55	Can stable surfactant with simpler novel modes of administration increase the use and availability of surfactant for preterm babies at risk of respiratory distress syndrome?	71 (62–73)	49
71	Can the method to diagnose fetal distress in labour be more accurate and affordable?	66 (57–71)	49
97	Can strategies for prevention and treatment of intrauterine growth restriction be developed?	64 (51–68)	46
105	Can novel tocolytic agents to delay or stop preterm labour be developed in order to reduce neonatal mortality and morbidity?	63 (54–68)	42
116	Can major causal pathways and risk factors for antepartum stillbirth be identified?	61 (52–66)	43
118	Can novel point of care diagnostics for congenital syphilis be identified in low resource setting to improve management?	60 (53–64)	49
120	Can novel antibiotic or other biological agents be identified?	60 (51–65)	40
121	Can the new method identify intrauterine growth restriction at the early stage (including biomarkers) and predict abnormal postnatal growth and body composition?	60 (52–63)	43
125	Can novel vaccines for maternal immunization be developed and evaluated to prevent newborn infections (eg, GBS, Klebsiella, *E coli*, Staph)?	60 (51–64)	41
129	Can preterm birth be delayed or averted with antioxidant and/or nutrient supplementation (eg, Vitamin D, omega–3 fatty acids)?	58 (48–63)	42

GBS – group B streptococcus, Staph – staphylococcus

Table 4. Overall scoring pattern by profile of experts

	Median (IQR)		
	All scorers (n = 91)	Researchers (n = 61)	Programme experts (n = 30)
TOTAL SCORE			
Delivery	82 (80–86)	83 (78–86)	86 (81–87)
Development	74 (72–74)	75 (71–76)	75 (68–79)
Discovery	61 (59–64)	62 (60–62)	63 (58–65)
AGREEMENT			
Delivery	67 (65–72)	68 (64–73)	70 (65–75)
Development	57 (55–58)	58 (56–60)	55 (54–62)
Discovery	43 (42–49)	45 (42–47)	44 (39–49)
ANSWERABLE?			
Delivery	92 (87–94)	92 (88–95)	91 (90–94)
Development	84 (82–89)	87 (81–90)	84 (78–89)
Discovery	76 (73–78)	76 (74–79)	76 (70–79)
EFFICACY?			
Delivery	87 (84–91)	87 (83–91)	88 (84–90)
Development	81 (77–83)	84 (79–84)	78 (76–81)
Discovery	68 (64–70)	68 (65–72)	69 (59–72)
DELIVERABILITY?			
Delivery	85 (82–89)	86 (82–91)	87 (82–89)
Development	77 (75–80)	79 (77–81)	74 (70–84)
Discovery	68 (66–72)	69 (64–72)	70 (64–72)
IMPACT?			
Delivery	68 (62–72)	65 (58–70)	73 (69–80)
Development	56 (53–57)	53 (52–58)	62 (52–65)
Discovery	46 (39–50)	46 (38–48)	44 (36–54)
EQUITY?			
Delivery	84 (81–88)	84 (76–89)	87 (79–88)
Development	74 (66–77)	71 (65–76)	76 (75–80)
Discovery	54 (50–59)	52 (50–58)	53 (50–65)

The major emerging themes in the domain of "delivery" included simplifying intervention delivery to implementation at lower levels of the health system, evaluating delivery of interventions by community health workers, developing strategies to improve quality of care during labour and childbirth, and addressing barriers in the scaling up of high impact interventions. It is interesting to note that 5 of the questions were related to neonatal resuscitation. This could be related to neonatal resuscitation being the most dramatic intervention in newborn care. The major themes in the domain of "development" were adapting known interventions to make them deliverable at the community level, adapting effective interventions to increase deliverability in health facilities in low and middle income countries, and approaches such as perinatal audits to improve quality of care to mothers and newborns. The themes in the domain of "discovery" included new, more effective and less expensive medicines for preventing preterm birth and treating sepsis, point of care diagnostics for infections, maternal vaccines to prevent newborn infections, and basic science work on causal pathways for identifying intervention targets and biomarkers for preterm birth, IUGR, and antepartum stillbirths. It is noteworthy that preterm prevention was not ranked highly, even though it may have the largest impact. This appears to be the result of these questions being scored low in answerability.

The relatively lower scores for the "development" and "discovery" groups of research questions may have several possible explanations. First, more than 95% of the neonatal deaths occur in low and middle–income countries (LMICs). Therefore, research addressing neonatal health issues that are relatively more important in wealthy countries may be

perceived to contribute less to global reduction in mortality and morbidity, explaining some of the lower scores received by potentially promising research on novel interventions based on high technologies. Second, "discovery" research often takes longer to be translated into measurable benefits in terms of mortality burden reduction, and by definition the link to reduction in mortality and inequity is less direct. One specific example is research on prevention of preterm birth – while it was likely to have high impact, it was ranked only 129th among the 205 questions. Thereby, respondents sent a message that this research question would likely be difficult to answer given the current stage of knowledge. Third, the process of delivery of novel interventions usually requires specific funding mechanisms, such as PEPFAR or Advance Market Commitment (AMC), which require time for a political agreement [34,35].

The CHNRI process we followed for setting priorities has several strengths. The methodology is transparent, replicable, and feasible to apply via e–mail [8–12, 27–33]. The output is intuitive and easily understood, and it has been refined and improved through many exercises over the past several years [36]. In this particular exercise, further improvements have been introduced to the process. We chose a large number of experts based on their productivity in the previous five years using Web of Science®, thus transparently identifying the group that was most likely to understand the field and its present research challenges and gaps. A very wide global network of programme experts in the Saving Newborn Lives' Network was also invited. Moreover, we used online data collection tools, such as Survey Monkey® and Google Analytics®, which allowed monitoring of the progress of the exercise in real time, ensured adequate representation of experts by their background and region, and increased the efficiency of data management. Finally, 132 experts proposed research questions and 91 scored all the questions in this exercise; this is considerably more than in previous priority setting exercises using CHNRI methodology, where we typically involved fewer scorers, research ideas, and criteria scored by each expert.

There may be concern that the results derived from the CHNRI approach might represent only the collective opinion of the limited group of people who were included in the process. However, we were able to obtain questions and scores from a large number of experts worldwide, who were selected in a transparent and replicable manner, based on their research productivity in the field. The large number of participants and the protection against potential bias provided by the CHNRI approach make our results more credible, although it remains apparent that the highest scored questions may still be biased towards those that researchers are most familiar with and so may bias reflect research already in progress. This issue may be particularly relevant in view that only about a quarter of originally invited researchers, policy makers and programme experts eventually contributed to generating research questions, and only about one in six completed the scoring process, making response bias an important potential concern. Second, even though the list of proposed questions was reviewed and refined before sending for scoring, there were still overlaps in some research questions, possibly creating confusion in scoring such questions. Those and other possible strengths and limitations of CHNRI methodology are described and discussed in greater detail in **Online Supplementary Document**.

A recent analysis of funding committed globally to improving neonatal health and birth outcomes has shown that donor mention of the "newborn" has increased quite sharply since 2005. However, given a total of only 10% of all donor aid to RMNCH mentioning the word "newborn", and only 0.01% referring to interventions expected to reduce newborn deaths, it still seems unlikely that donor aid is commensurate with the large burden of 3.0 million newborn deaths each year, or with the burden of morbidity, developmental and long–term health outcomes [37]. The word "stillbirth" occurred only twice in the OECD database between 2002 and 2010, suggesting even lower attention for the world's 2.7 million stillbirths.

Large inequities in current research funding support exist not only in the amounts invested in newborn health in comparison to other diseases globally, but also between different neonatal conditions themselves. Conditions that affect newborns in high–income countries receive more funding and attention than conditions that largely affect newborns in low–income countries. For instance, the research on care of preterm babies in neonatal intensive care units has received considerably more funding over the past several years in comparison to intrapartum–related birth outcomes or newborn sepsis [38].

The results presented in this paper will assist both the donors and the researchers in setting evidence based priorities to address the key gaps in knowledge, that could make the most difference in saving newborn lives and preventing stillbirth. In addition, attention to many of these questions could also improve maternal and child health outcomes. Likewise, research priorities to address other related areas such as maternal, child and adolescent health and health system issues may have substantial effect on newborn health. Complementary exercises are under way to identify research priorities in these areas. Using the identified research priorities, WHO, SNL and other partners, that are linked to the Every Newborn action plan launched in 2014 [39], will work to generate research interests among key national stakeholders, governments, NGOs, and research institutes, while encouraging research funders to support these priorities. We will track research funding, relevant request for proposals and trial registers to monitor if the priorities identified by this exercise are being addressed, and highlight those that are not being addressed.

AFFILIATIONS OF AUTHORS

[1] Department of Maternal, Newborn, Child and Adolescent Health, World Health Organization, Geneva, Switzerland
[2] Centre for Intervention Science in Maternal and Child Health, Centre for International Health, University of Bergen, Norway
[3] London School of Hygiene and Tropical Medicine, London, UK
[4] Saving Newborn Lives, Save the Children, Washington, USA
[5] Department of Social Medicine, Ribeirão Preto School of Medicine, University of São Paulo, Brazil
[6] Centre for Population Health Sciences and Global Health Academy, The University of Edinburgh Medical School, Scotland, UK
[7] Bandim Health Project, Indepth Network, Guinea–Bissau
[8] Faculty of Medicine, University of Khartoum, Sudan
[9] Kathmandu Medical College, Nepal
[10] Department of Pediatrics, University of Alabama at Birmingham, Birmingham, AL, USA
[11] Centre for Child and Adolescent Health, International Centre for Diarrhoeal Disease Research, Bangladesh
[12] Department of Neonatology Paropakar Maternity and women's Hospital, Nepal
[13] TRAction Bangladesh Project, University Research Co., LLC
[14] Johns Hopkins Bloomberg School of Public Health, USA
[15] Centro de Pesquisas Epidemiológicas, Universidade Federal de Pelotas, Brazil
[16] Research Center for Vitamins and Vaccines, Statens Serum Institut, and University of Southern Denmark/Odense University Hospital, Denmark
[17] Program in Perinatal Research, Yale University School of Medicine, USA
[18] Pediatric Biology Centre, Translational Health Science and Technology Institute, India
[19] University of Aberdeen, UK
[20] Center of Excellence in Women and Child Health, the Aga Khan University, Karachi, Pakistan
[21] Institute of International Programs, Johns Hopkins Bloomberg School of Public Health, USA
[22] The London School of Hygiene & Tropical Medicine, UK
[23] The University of North Carolina at Chapel Hill School of Medicine, USA
[24] Thrasher Research Fund, USA
[25] Department of Neonatology, Charité University Medical Center, Germany
[26] University of Alabama at Birmingham, USA
[27] Department of Obstetrics and Gynaecology, School of Medical Sciences, University of Campinas, Brazil
[28] Departments of Pediatrics, Pharmacology & Surgery, University of Alberta, Canada
[29] Charities, USA
[30] University College London Institute for Global Health, UK
[31] Perinatology Research Branch, Eunice Kennedy Shriver National Institute of Child Health and Human Development/National Institutes of Health/Department of Health and Human Services, Bethesda, Maryland and Detroit, Michigan, USA
[32] Independent consultant maternal health research, Rwanda
[33] Foundation for the Community Control of Hereditary Diseases, Hungary
[34] Biostatistics and Epidemiology, RTI International, USA
[35] LAMB Integrated Rural Health & Development, Bangladesh
[36] Division of Microbiology and Infectious Diseases, National Institute of Allergy and Infectious Diseases National Institute of Health, USA
[37] All India Institute of Medical Sciences, India
[38] Pediatrics and Neonatology, Yıldırım Beyazıt University Medical Faculty, Turkey
[39] Nuffield Department of Medicine & Department of Paediatrics, University of Oxford, UK and KEMRi–Wellcome Trust Research Programme, Nairobi, Kenya
[40] Newborn Health, Family Health Division, The Bill & Melinda Gates Foundation and the University of North Carolina Schools of Medicine and Public Health, USA
[41] Moi University, School of Medicine, Kenya
[42] International Paediatric Epidemiology; Affiliations: Medical Research Council Lifecourse Epidemiology Unit, University of Southampton, UK
[43] Department of Pediatrics, UCSF Benioff Children's Hospital, USA
[44] School of Medicine, Child Health and Pediatrics, Moi University, Kenya
[45] Children's Hospital, Pakistan Institute of Medical Sciences, Pakistan
[46] Eunice Kennedy Shriver NICHD Neonatal Research Network, Pregnancy and Perinatology, Branch, National Institute of Health, USA
[47] Centre for Midwifery, Child and Family Health, University of Technology, Sydney, Australia
[48] Centre for Child and Adolescent Health, International Centre for Diarrhoeal Disease Research, Bangladesh
[49] University of Bergen and Norwegian Institute of Public Health, Norway
[50] Japan International Cooperation Agency (JICA), Bangladesh
[51] Save the Children, USA
[52] Zomba Central Hospital, Ministry of Health, Malawi
[53] Saint Louis University, USA
[54] Paediatrics and Head of Department Shaheed Suhrawardy Medical College, Bangladesh
[55] Centre for Pharmacoepidemiology, Karolinska Institute, Sweden
[56] Departments of Pediatrics and of Epidemiology, Biostatistics and Occupational Health, McGill University Faculty of Medicine, Montreal, Quebec, Canada
[57] Global Alliance to Prevent Prematurity and Stillbirth (GAPPS), USA
[58] University of Manchester School of Nursing Midwifery & Social Work, University of Manchester, UK
[59] Indonesian Midwives Association, Indonesia
[60] Save the Children, Malawi
[61] Save the Children, USA
[62] Department of Paediatrics, Dalhousie University, Canada
[63] Departments of Paediatrics & Physiology, University of Toronto; Physiology & Experimental Medicine program, Hospital for Sick Children, Toronto, Canada
[64] Department of Obstetrics and Gynaecology, Academic Medical Centre Amsterdam, the Netherlands
[65] Paediatrics and Child Health, College of Medicine Malawi
[66] International Baby Food Action Network, Uganda
[67] STRIDES for Family Health, Management Sciences for Health, Uganda
[68] Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, UK
[69] Kampala Children's Hospital Limited and Childhealth Advocacy International, Uganda
[70] United States Agency for International Development /Maternal and Child Health Integrated Program, USA.
[71] Centre for Healthy Start Initiative, Nigeria
[72] Wellcome Trust Senior Research Fellow in Clinical Science, Institute for Global Health, University College London, UK
[73] All India Institute of Medical Sciences, India
[74] University of Tubingen, Germany
[75] Eunice Kennedy Shriver National Institute of Child Health and Human Development National Institutes of Health, USA
[76] Center for American Indian Health, Johns Hopkins Bloomberg School of Public Health, USA
[77] Save the Children, Bangladesh
[78] Department of Obstetrics and Gynecology, University of Tennessee Health Science Center, USA
[79] University of Calgary, Canada
[80] Environmental Health Science Department, University of Georgia, USA
[81] Tropical Epidemiology Group, London School of Hygiene & Tropical Medicine, London, UK
[82] Department of Pediatrics & Child Health, Women & Child Health Division, Aga Khan University, Pakistan
[83] Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, USA
[84] National Institute of Population Research and Training (NIPORT), Ministry of Health and Family Welfare, Bangladesh
[85] Kinshasa School of Public Health, School of Medicine, University of Kinshasa, Democratic Republic of Congo
[86] Department of Neonatology, University of Utrecht, the Netherlands
[87] Institute for Global Health Technologies Rice University, USA
[88] Division of Global Health, Karolinska Institutet, Sweden
[89] School of Medical Sciences, Edith Cowan University, Australia and School of Public Health, Capital Medical University, China
[90] Save the Children UK
[91] Eunice Kennedy Shriver National Institute of Child Health and Human Development National Institutes of Health, USA
[92] Aga Khan University, Pakistan
[93] Department of Integrated Early Childhood Development, Capital Institute of Paediatrics, China
[94] Developmental Genetics Laboratory, New York State Institute for Basic Research in Developmental Disabilities, USA
[95] Médecins sans Frontières, Belgium

Acknowledgements: The authors thank the expert group for their time and effort in this priority–setting exercise, and the contribution of anonymous reviewers to the final version of this report. We acknowledge inputs from Dr Diane Morof from Centers for Disease Control and Prevention, USA.

Disclaimer: SY, RB, JPS, JM are employees of the World Health Organization. The views expressed in this paper are the responsibility of the authors and do not necessarily represent the views of the World Health Organization.

Funding: This work was supported by Save the Children.

Ethical approval: Not required.

Authorship declaration: SY and IR conducted the analysis and wrote the first draft of the paper. RB, JM, JL, SW, JPS and SC revised the draft and finalized the paper. All the named authors had a major role in design, implementation or analysis of the priority setting exercise. The Newborn Health Priority setting group proposed research questions, helped finalize the list of questions and provided scores that served as the dataset for all quantitative analyses

REFERENCES

1 UNICEF, World Health Organization, The World Bank and United Nations. Level & Trends in Child Mortality, Estimates developed by the UN inter–agency group for child mortality estimation. New York: UNICEF, 2014.

2 Cousens S, Blencowe H, Stanton C, Chou D, Ahmed S, Steinhardt L, et al. National, regional, and worldwide estimates of stillbirth rates in 2009 with trends since 1995: a systematic analysis. Lancet. 2011;377:1319-30. Medline:21496917 doi:10.1016/S0140-6736(10)62310-0

3 Liu L, Johnson HL, Cousens S, Perin J, Scott S, Lawn JE, et al. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. Lancet. 2012;379:2151-61. Medline:22579125 doi:10.1016/S0140-6736(12)60560-1

4 Oza S, Cousens SN, Lawn JE. Estimation of daily risk of neonatal death, including the day of birth, in 186 countries in 2013: a vital–registration and modelling–based study. Lancet Glob Health. 2014;2:e635-44. Medline:25442688 doi:10.1016/S2214-109X(14)70309-2

5 Mwaniki MK, Atieno M, Lawn JE, Newton CRJC. Long–term neurodevelopmental outcomes after intrauterine and neonatal insults: a systematic review. Lancet. 2012;379:445-52. Medline:22244654 doi:10.1016/S0140-6736(11)61577-8

6 Horton R. GBD 2010: understanding disease, injury, and risk. Lancet. 2012;380:2053-4. Medline:23245595 doi:10.1016/S0140-6736(12)62133-3

7 United Nations. A new global partnership: Eradicate poverty and transform economies through sustainable development. The report of the high–level Panel of Eminent Persons on the Post–2015 Development Agenda. New York: United Nations, 2013.

8 Fontaine O, Kosek M, Bhatnagar S, Boschi–Pinto C, Chan KY, Duggan C, et al. Setting research priorities to reduce global mortality from childhood diarrhoea by 2015. PLoS Med. 2009;6:e41. Medline:19278292 doi:10.1371/journal.pmed.1000041

9 Rudan I, El Arifeen S, Bhutta ZA, Black RE, Brooks A, Chan KY, et al. Setting research priorities to reduce global mortality from childhood pneumonia by 2015. PLoS Med. 2011;8:e1001099. Medline:21980266 doi:10.1371/journal.pmed.1001099

10 Lawn JE, Bahl R, Bergstrom S, Bhutta ZA, Darmstadt GL, Ellis M, et al. Setting research priorities to reduce almost one million deaths from birth asphyxia by 2015. PLoS Med. 2011;8:e1000389. Medline:21305038 doi:10.1371/journal.pmed.1000389

11 Bahl R, Martines J, Ali N, Bhan MK, Carlo W, Chan KY, et al. Research priorities to reduce global mortality from newborn infections by 2015. Pediatr Infect Dis J. 2009;28:S43-8. Medline:19106763 doi:10.1097/INF.0b013e31819588d7

12 Bahl R, Martines J, Bhandari N, Biloglav Z, Edmond K, Iyengar S, et al. Setting research priorities to reduce global mortality from preterm birth and low birth weight by 2015. J Glob Health. 2012;2:010403. Medline:23198132 doi:10.7189/jogh.01.010403

13 Pitt C, Lawn JE, Ranganathan M, Mills A, Hanson K. Donor funding for newborn survival: an analysis of donor–reported data, 2002–2010. PLoS Med. 2012;9:e1001332. Medline:23118619 doi:10.1371/journal.pmed.1001332

14 Sridhar D. Who sets the global health research agenda? The challenge of multi–bi financing. PLoS Med. 2012;9:e1001312. Medline:23049485 doi:10.1371/journal.pmed.1001312

15 Greenwood B. A global action plan for the prevention and control of pneumonia. Bull World Health Organ. 2008;86:322-2A. Medline:18545727 doi:10.2471/BLT.08.053348

16 Zipursky A, Wazny K, Black R, Keenan W, Duggan C, Olness K, et al. Global action plan for childhood diarrhoea. Developing research priorities: Report from a Workshop of the Programme for Global Paediatric Research. J Glob Health. 2013;3:010406. Medline:23826510 doi:10.7189/jogh.03.010406

REFERENCES

17 Bhutta ZA, Chopra M, Axelson H, Berman P, Boerma T, Bryce J, et al. Countdown to 2015 decade report (2000–10): taking stock of maternal, newborn, and child survival. Lancet. 2010;375:2032-44. Medline:20569843 doi:10.1016/S0140-6736(10)60678-2

18 Lawn JE, Gravett MG, Nunes TM, Rubens CE, Stanton C; Global Action Plan for Prematurity and Stillbirths (GAPPS) Review Group. Global report on preterm birth and stillbirth (1 of 7): definitions, description of the burden and opportunities to improve data. BMC Pregnancy Childbirth. 2010;10 Suppl 1:S1. Medline:20233382

19 World Pneumonia Day. Available: http://worldpneumoniaday.org. Accessed: 1 August 2013.

20 Walker CL, Rudan I, Liu L, Nair H, Theodoratou E, Bhutta ZA, et al. Global burden of childhood pneumonia and diarrhoea. Lancet. 2013;381:1405-16. Medline:23582727 doi:10.1016/S0140-6736(13)60222-6

21 Seale AC, Mwaniki M, Newton CR, Berkley JA. Maternal and early onset neonatal bacterial sepsis: burden and strategies for prevention in sub–Saharan Africa. Lancet Infect Dis. 2009;9:428-38. Medline:19555902 doi:10.1016/S1473-3099(09)70172-0

22 Blencowe H, Cousens S, Oestergaard MZ, Chou D, Moller AB, Narwal R, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. Lancet. 2012;379:2162-72. Medline:22682464 doi:10.1016/S0140-6736(12)60820-4

23 Lawn JE, Cousens S, Zupan J; Lancet Neonatal Survival Steering Team. 4 million neonatal deaths: When? Where? Why? Lancet. 2005;365:891-900. Medline:15752534 doi:10.1016/S0140-6736(05)71048-5

24 Rudan I, El Arifeen S, Black RE, Campbell H. Childhood pneumonia and diarrhoea: setting our priorities right. Lancet Infect Dis. 2007;7:56-61. Medline:17182344 doi:10.1016/S1473-3099(06)70687-9

25 Rudan I, Gibson JL, Ameratunga S, El Arifeen S, Bhutta ZA, Black M, et al. Setting priorities in global child health research investments: Guidelines for implementation of the CHNRI method. Croat Med J. 2008;49:720-33. Medline:19090596 doi:10.3325/cmj.2008.49.720

26 Rudan I, Chopra M, Kapiriri L, Gibson J, Lansang MA, Carneiro I, et al. Setting priorities in global child health research investments: Universal challenges and conceptual framework. Croat Med J. 2008;49:307-17. Medline:18581609 doi:10.3325/cmj.2008.3.307

27 Tomlinson M, Swartz L, Officer A, Chan KY, Rudan I, Saxena S. Research priorities for health of people with disabilities: an expert opinion exercise. Lancet. 2009;374:1857-62. Medline:19944866 doi:10.1016/S0140-6736(09)61910-3

28 Brown KH, Hess SY, Boy E, Gibson RS, Horton S, Osendarp SJ, et al. Setting priorities for zinc–related health research to reduce children's disease burden worldwide: an application of the Child Health and Nutrition Research Initiative's research priority–setting method. Public Health Nutr. 2009;12:389-96. Medline:18426636

29 Wazny K, Zipursky A, Black R, Curtis V, Duggan C, Guerrant R, et al. Setting research priorities to reduce mortality and morbidity of childhood diarrhoeal disease in the next 15 years. PLoS Med. 2013;10:e1001446. Medline:23690756 doi:10.1371/journal.pmed.1001446

30 Tomlinson M, Rudan I, Saxena S, Swartz L, Tsai AC, Patel V. Setting priorities for global mental health research. Bull World Health Organ. 2009;87:438-46. Medline:19565122 doi:10.2471/BLT.08.054353

31 Souza JP, Widmer M, Gülmezoglu AM, Lawrie TA, Adejuyigbe EA, Carroli G. Maternal and perinatal health research priorities beyond 2015: an international survey and prioritization exercise. Reprod Health. 2014;11:61. Medline:25100034 doi:10.1186/1742-4755-11-61

32 Tomlinson M, Swartz L, Officer A, Chan KY, Rudan I, Saxena S. Research priorities for health of people with disabilities: an expert opinion exercise. Lancet. 2009;374:1857-62. Medline:19944866 doi:10.1016/S0140-6736(09)61910-3

33 Hindin MJ, Christiansen CS, Ferguson BJ. Setting research priorities for adolescent sexual and reproductive health in low– and middle–income countries. Bull World Health Organ. 2013;91:10-8. Medline:23397346 doi:10.2471/BLT.12.107565

34 Goosby E, Von Zinkernagel D, Holmes C, Haroz D, Walsh T. Raising the bar: PEPFAR and new paradigms for global health. J Acquir Immune Defic Syndr. 2012;60 Suppl 3:S158-62. Medline:22797738 doi:10.1097/QAI.0b013e31825d057c

35 Zandonella C. If you build it, they will pay. A novel incentive called and Advance Market Commitment could help spur private sector investment in AIDS vaccine research and development. IAVI Rep. 2005;9:13-21. Medline:16276579

36 Rudan I. Global health research priorities: mobilizing the developing world. Public Health. 2012;126:237-40. Medline:22325672 doi:10.1016/j.puhe.2011.12.001

37 Pitt C, Lawn JE, Ranganathan M, Mills A, Hanson K. Donor funding for newborn survival: an analysis of donor–reported data, 2002–2010. PLoS Med. 2012;9:e1001332. Medline:23118619 doi:10.1371/journal.pmed.1001332

38 Howson CP, Kinney MV, Lawn JE, editors. March of Dimes; PMNCH; Save the Children; World Health Organization. Born too soon: The global action report on preterm birth. Geneva: World Health Organization, 2012.

39 Every newborn action plan – an action plan to end preventable deaths. Available: http://www.globalnewbornaction.org/about/. Accessed: 31 July, 2014.

journal of global health

Validating hierarchical verbal autopsy expert algorithms in a large data set with known causes of death

Henry D Kalter[1], Jamie Perin[2,3], Robert E Black[3]

[1] Health Systems Program, Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

[2] Center for Child and Community Health Research, Department of Pediatrics, Johns Hopkins School of Medicine, Baltimore, MD, USA

[3] Institute for International Programs, Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Background Physician assessment historically has been the most common method of analyzing verbal autopsy (VA) data. Recently, the World Health Organization endorsed two automated methods, Tariff 2.0 and InterVA–4, which promise greater objectivity and lower cost. A disadvantage of the Tariff method is that it requires a training data set from a prior validation study, while InterVA relies on clinically specified conditional probabilities. We undertook to validate the hierarchical expert algorithm analysis of VA data, an automated, intuitive, deterministic method that does not require a training data set.

Methods Using Population Health Metrics Research Consortium study hospital source data, we compared the primary causes of 1629 neonatal and 1456 1–59 month–old child deaths from VA expert algorithms arranged in a hierarchy to their reference standard causes. The expert algorithms were held constant, while five prior and one new "compromise" neonatal hierarchy, and three former child hierarchies were tested. For each comparison, the reference standard data were resampled 1000 times within the range of cause–specific mortality fractions (CSMF) for one of three approximated community scenarios in the 2013 WHO global causes of death, plus one random mortality cause proportions scenario. We utilized CSMF accuracy to assess overall population–level validity, and the absolute difference between VA and reference standard CSMFs to examine particular causes. Chance–corrected concordance (CCC) and Cohen's kappa were used to evaluate individual–level cause assignment.

Results Overall CSMF accuracy for the best–performing expert algorithm hierarchy was 0.80 (range 0.57–0.96) for neonatal deaths and 0.76 (0.50–0.97) for child deaths. Performance for particular causes of death varied, with fairly flat estimated CSMF over a range of reference values for several causes. Performance at the individual diagnosis level was also less favorable than that for overall CSMF (neonatal: best CCC = 0.23, range 0.16–0.33; best kappa = 0.29, 0.23–0.35; child: best CCC = 0.40, 0.19–0.45; best kappa = 0.29, 0.07–0.35).

Conclusions Expert algorithms in a hierarchy offer an accessible, automated method for assigning VA causes of death. Overall population–level accuracy is similar to that of more complex machine learning methods, but without need for a training data set from a prior validation study.

Correspondence to:
Henry Kalter, MD, MPH
Health Systems Program
Department of International Health
Johns Hopkins Bloomberg School of Public Health
615 N. Wolfe St.
Suite E8132
Baltimore, MD 21205, USA
hkalter1@jhu.edu

For decades, health officials and program managers in low and middle income countries (LMIC) without well–functioning vital registration systems have used information on causes of death from verbal autopsy (VA) to allocate scarce resources to target the most common causes of child death. Simultaneously, the World Health Organization (WHO) and UNICEF, through their Child Health Epidemiology Reference Group (CHERG), have used VA data from the world's public health literature to model and track the causes of neonatal and child death in LMIC countries [1–4]. However, VA data collection and analysis methods, including those of studies that have contributed input data to the CHERG models, have suffered from a lack of standardization and uncertainty as to the accuracy of their cause of death findings [5].

Until lately most studies have relied on physician analysis of VA findings, which has raised questions regarding the potential introduction of subjectivity and cultural biases into the VA diagnoses, as well as the monetary and health system costs of diverting physicians from patient care to the task of VA analysis [6]. Expert algorithms also have been used for VA analysis, with validation studies demonstrating fair to good accuracy for the diagnosis of several causes of neonatal and child death [7–10]; but this method has more often been used in research settings, with program environments being more comfortable with physician analysis. More recently, several machine learning and probabilistic VA analysis methods have been developed that show promise for providing more accurate diagnoses, as well as the objectivity that comes with automated methods and the efficiency and cost savings of not requiring physicians to conduct the analysis [11]. WHO recently modified its standardized VA questionnaire for use with two of these automated methods, Tariff 2.0 [12] and InterVA–4 [13], and is encouraging the use of these methods instead of the traditional physician review method [14].

However, questions remain as to which method or methods is most accurate, with a recent assessment emphasizing that different methods may work best for different age groups and causes of death [15]. Lastly, none of these studies examined the use of expert algorithms arranged in a hierarchy to select the primary cause of death, which offers the same advantages as other automated methods plus the additional benefit, unlike the Tariff method, of not requiring a training data set from a prior VA validation study, preferably conducted in the same geographic region or disease setting intended for the use of verbal autopsy, and distinct from all other automated methods, is based on clinical algorithms that can be easily explained to non–medical professionals. A later study did examine the performance of hierarchical algorithms, but in a small data set against physician–determined reference standard diagnoses using algorithms refined by physicians at the same sites, and missing some key neonatal causes of death [16]. Therefore, we undertook to validate the hierarchical expert algorithm VA analysis method in a large data set with objective reference standard criteria for a full range of important neonatal and child causes of death, and report the findings of our analyses in this paper.

METHODS

We used source data from the Population Health Metrics Research Consortium (PHMRC) study to validate causes of under–five year–old deaths from verbal autopsy expert algorithms arranged in a hierarchy compared to reference standard causes of death. The design and primary results of the PHMRC study have been described in detail [17]. In brief, the study identified hospital deaths of all ages, including 1629 neonatal deaths and 1456 1–59 months old child deaths, at six study sites in five countries on three continents, determined the main or underlying reference standard cause for each death from available clinical, laboratory and imaging data, and later visited the household of each decedent to conduct a verbal autopsy interview. A large portion of these data are publicly available [18], although some questions about its contents have risen from the verbal autopsy research community [19]. For this reason, we conducted extensive cleaning of the PHMRC data to make it more suitable for our expert algorithm analysis, and have provided the cleaned data, documentation and cleaning information online [20]). We excluded stillbirths and deaths of persons older than five years from our analysis, restricting our interest to deaths of live born children who died before age five, analyzed separately for neonates 0 to 27 days and children 1 to 59 months old.

Verbal autopsy cause of death assignment

Verbal autopsy (VA) expert algorithms are combinations of illness signs and symptoms judged by verbal autopsy researchers to be predictive of particular causes of death. The algorithms validated in the current study were based on those developed by researchers for prior VA validation studies, further consultation with additional verbal autopsy experts, and a literature review to identify illness signs and symptoms commonly associated with particular neonatal and child illnesses. The sources and algorithms themselves are provided in a recent publication [21]. We used the expert algorithms to estimate cause of death given each individual's PHMRC VA questionnaire responses. While the PHMRC questionnaire includes close–ended questions on illness signs and symptoms, an open–ended narrative response and recording of data from medical records and death certificates available in the home, the expert algorithms are based only on the responses to close–ended questions on illness signs and symptoms.

Because the algorithms determine all contributing causes, in the event that more than one cause was identified the primary cause was chosen according to a pre–specified hierarchy. We determined the primary causes of neonatal death utilizing the same algorithms across five hierarchies for neonatal deaths that are currently in use: Arifeen et al. [22], Baqui et al. [23], Kalter et al. [21], Lawn et al. [24], and Liu et al. [25]; and the primary causes of child death (1–59 months of age) utilizing the hierarchies for this age group described by Arifeen et al. [22], Kalter et al. [21], and Liu et al. [25]. Other things being equal, estimating more causes at once will yield lower accuracy than estimating fewer causes [26]. Therefore, for neonatal deaths, we also examined a compromise hierarchy that included four cause categories in common across all five neonatal hierarchies (**Table 1**).

Reference standard cause of death

We used the reference standard causes of death from the PHMRC study to approximate the cause of death distribution in community settings, where verbal autopsy is most relevant. Because the PHMRC study was hospital– as op-

Table 1. Cause assignment hierarchies for determining the main cause of death among co–morbid causes in neonates 0–27 days and 1–59 month–old children

Arifeen et al. 2004 [22]	Baqui et al. 2006 [23]	Kalter et al. 2015 [21]	Lawn et al. 2006 [24]	Liu et al. 2015 [25]	Compromise
Neonates 0–27 days:					
Neonatal tetanus	Neonatal tetanus, Congenital abnormality	Neonatal tetanus	Congenital abnormality	Neonatal tetanus	Congenital abnormality
Congenital abnormality	Preterm delivery	Congenital abnormality	Neonatal tetanus	Congenital abnormality	Birth asphyxia
Birth asphyxia	Birth asphyxia	Birth asphyxia, birth injury	Preterm birth	Birth asphyxia, birth injury	Prematurity
Birth injury	Birth injury	Meningitis	Birth asphyxia	Diarrhea, ARI	Sepsis, pneumonia, meningitis
ARI, diarrhea	Sepsis or pneumonia	Diarrhea	Sepsis, pneumonia, meningitis	Meningitis	
Possible diarrhea, possible ARI, sepsis	Diarrhea	Pneumonia	Diarrhea	Possible pneumonia, possible diarrhea	
Premature birth/LBW	Unspecified	Possible diarrhea	Other	Prematurity/LBW	
Other causes		Possible pneumonia		Sepsis, other possible serious infections	
Unspecified		Sepsis		Unspecified	
		Jaundice			
		Hemorrhagic disease of the newborn			
		Sudden unexplained death			
		Preterm delivery			
		Unspecified			
Children 1–59 months:					
Injury		Injury		Injury	
ARI, diarrhea, measles		AIDS		Measles, diarrhea, ARI	
Possible serious infections		Malnutrition (underlying)		Meningitis	
Malnutrition		Measles		Malaria	
Other causes		Meningitis		AIDS	
Unspecified		Dysentery		Possible diarrhea/ARI	
Undetermined		Diarrhea		Other possible serious infections	
		Pertussis		Unspecified	
		Pneumonia			
		Malaria			
		Possible dysentery			
		Possible diarrhea			
		Possible pneumonia			
		Hemorrhagic fever			
		Other infection			
		Residual infection			
		Malnutrition			
		Unspecified			

LBW – low birth weight, ARI – acute respiratory infection

posed to community based, and the cause distribution in the community and hospital may differ, we resampled from the study deaths to represent a variety of cause distributions.

We approximated three specific mortality settings with the PHMRC data: (1) communities with high under five mortality where malaria is endemic, (2) communities with high under five mortality where malaria is not endemic, and (3) communities with moderate under five mortality. We used the Child Health Epidemiology Reference Group (CHERG) definition of high under–five mortality (more than 35 deaths per 1000 live births) [4], took moderate mortality as 20 to 35 deaths per 1000, and defined malaria endemicity as greater than 5 percent of under–five deaths due to malaria. In addition to these three specific scenarios of interest, we also considered a fourth general scenario, where all cause-specific mortality fractions were randomly varied between 5% and 40%.

Estimated cause proportions of death for all countries in the world, including those where most deaths occur outside the formal health sector, are available from the WHO [4]. We used these estimated causes of neonatal and child mortality as a guide in choosing cause distributions in our scenarios of interest. To generate a possible set of verbal autopsies to represent a given death distribution in a particular mortality scenario, we selected one country at random among all those appropriate, and resampled the PHMRC questionnaire data to correspond approximately to that cause of death distribution. For neonates, we included deaths due to prematurity, birth asphyxia, congenital malformations, meningitis, pneumonia and sepsis; and for children we used deaths from HIV, diarrhea, measles, meningitis/encephalitis, malaria, pneumonia, injuries, other infectious causes, and non–infectious causes. Some causes of interest for Liu et al. [4] do not occur in the PHMRC study data, requiring that we use relative proportions of causes reported by the PHMRC, while unreported causes were not considered. For example, the tetanus mortality fraction for neonatal deaths as reported by WHO is as high as 8%, but there are no neonatal deaths due to tetanus in the PHMRC data.

The PHMRC data include neonatal deaths due to co–morbid preterm delivery, birth asphyxia and/or sepsis; and child deaths due to co–morbid pneumonia and diarrhea. For deaths with co–morbid reference standard causes of death, we used the ICD–10 rules to assign a single underlying cause of death [27]. In accordance with the rule that the mode of perinatal death, including prematurity, should not be classified as the main disease or condition unless it was the only condition known, we assigned deaths due to co–morbid preterm/birth asphyxia to birth asphyxia, preterm/sepsis to sepsis, and preterm/sepsis/birth asphyxia proportionately to sepsis and birth asphyxia. Deaths from conditions directly due to prematurity, such as Respiratory Distress Syndrome, were classified as being due to preterm delivery. For child deaths, we proportionately reallocated co–morbid pneumonia/diarrhea deaths to pneumonia or diarrhea. Using these verbal autopsies for harmonized causes of death, we repeated our selection of cause of death distribution and resampling 1000 times for each of the four scenarios. **Table 2** summarizes our harmonization of the verbal autopsy algorithms and reference standard causes of death.

Accuracy of VA cause of death determination

After resampling the reference standard cause of death data for neonates and children according to the four mortality scenarios as described above, we then, separately for neonatal and child deaths and for each hierarchy in each scenario, used the expert algorithms to estimate cause of death in the resampled reference standard cause of death data given each individual's VA questionnaire responses. We used four metrics to examine the validity of the VA cause of death estimates, two at the population level and two at the level of individual cause assignment. Cause-specific mortality fraction (CSMF) accuracy, as defined by Murray et al. [28], is an overall summary of the estimated and reference standard cause distributions with larger values indicating VA CSMF measurements closer to the reference standard. CSMF accuracy is the sum of absolute errors by cause, scaled by the extent of possible error given the smallest cause fraction, and subtracted from one. It is generally interpretable as percent accuracy. To assess the validity of VA estimates of particular causes of death we examined the absolute difference between VA and reference standard CSMFs for these causes.

The last two metrics estimate the accuracy of VA cause of death assignment at the level of individual deaths. Cohen's kappa is a general measure of agreement between estimated and reference standard causes [29]. Large values of kappa indicate more agreement, where in general values less than zero indicate no agreement, values between 0 and 0.2 are rated as minimal agreement, 0.2 to 0.4 as fair, 0.4 to 0.6 as moderate, 0.6 to 0.8 as substantial, and 0.8 to 1 approach exact agreement [30]. Chance corrected concordance (CCC) is another measure of agreement between VA and reference standard causes at the individual level. This statistic is closely related to Cohen's kappa and average sensitivity across causes or categories [28]. Similar to kappa, large values indicate more agreement. The CCC scale is from 1/(1–N) to 1, for the number of causes N, while the scale for Cohen's kappa is from –1 to 1. We used these two metrics only to generate overall summaries of VA accuracy for all causes together.

Ethics statement

The study data are publically accessible and include no personal identifiers. Therefore, no ethical review of the study protocol or informed consent was necessary.

Table 2. Correspondence of verbal autopsy and reference standard diagnoses in the hierarchies

Verbal autopsy algorithm(s)	PHMRC reference standard group(s)	Placement in hierarchy
Neonates 0–27 days:		
Neonatal tetanus	No PHMRC neonatal tetanus cases	–
Congenital malformation	Congenital malformation	Malformation
Birth injury	No PHMRC birth injury cases	–
Birth asphyxia	Birth asphyxia, preterm delivery (without RDS) and birth asphyxia, preterm delivery (without RDS) and sepsis and birth asphyxia (allocated to birth asphyxia according to the distribution of other deaths due to sepsis and birth asphyxia)	Birth asphyxia
Meningitis	Meningitis (serious infection)	Meningitis
Diarrhea	No PHMRC neonatal diarrhea cases	–
Pneumonia; ARI	Pneumonia (serious infection)	Pneumonia
Possible diarrhea	No PHMRC neonatal diarrhea cases	–
Possible pneumonia, possible ARI	Pneumonia (serious infection)	Possible pneumonia (later to combine with pneumonia)
Sepsis	Sepsis (serious infection), sepsis with local bacterial infection, preterm delivery (with or without RDS) and sepsis, preterm delivery (without RDS) and sepsis and birth asphyxia (allocated to sepsis according to the distribution of other deaths due to sepsis and birth asphyxia)	Sepsis
Jaundice	No PHMRC jaundice cases	–
Hemorrhagic disease of the newborn	No PHMRC hemorrhagic disease of the newborn cases	–
Sudden unexplained death	No PHMRC sudden unexplained death cases	–
Preterm delivery, Preterm delivery with complication specific to prematurity (RDS)	Preterm delivery (<33 weeks gestational age [GA]) with or without RDS, preterm delivery (33–36 weeks GA) with RDS	Preterm delivery
Children 1–59 months:		
Injury	Bite of a venomous animal, burn, drowning, fall, poisoning, road traffic injury, violent death	Injury
AIDS	AIDS	AIDS
Malnutrition (underlying)	No PHMRC malnutrition cases	–
Measles	Measles	Measles
Meningitis	Encephalitis, meningitis	Meningitis
Diarrhea or dysentery	Diarrhea/dysentery	Diarrhea/dysentery
Pneumonia or diarrhea	Pneumonia and diarrhea	Allocated to pneumonia and diarrhea/dysentery according to the distribution of other deaths due to pneumonia and diarrhea/dysentery
Pneumonia	Pneumonia	Pneumonia
Malaria	Malaria	Malaria
Possible diarrhea or dysentery	Diarrhea/dysentery	Possible diarrhea or dysentery (later to combine with diarrhea/dysentery)
Possible pneumonia	Pneumonia	Possible pneumonia (later to combine with pneumonia)
Pertussis, hemorrhagic fever, other infection	Hemorrhagic fever, sepsis, tuberculosis, other infectious diseases	Other infectious causes
Residual infection (possible malaria)	Malaria	Possible malaria (later to combine with malaria)

PHMRC – Population Health Metric Research Consortium, RDS – respiratory distress syndrome, ARI – acute respiratory infection

RESULTS

Neonates

Table 3 shows summary results for the expert algorithm cause of death assignments for all causes together from four mortality scenarios and three measures of accuracy. By the CSMF measure, the Baqui and Lawn hierarchies performed best in the moderate and general mortality scenarios, and the compromise hierarchy did best in both high mortality scenarios. These three hierarchies all did their best in the high mortality scenarios, whereas the Kalter and Liu hierarchies did their best in the general scenario, in which their performance nearly equaled that of the Lawn hierarchy. All the hierarchies did their worst, or nearly so, in the moderate mortality scenario. **Figure 1** also summarizes CSMF accuracy for neonatal deaths in these scenarios.

The Baqui and Lawn hierarchies performed best by the Cohen's kappa measure, followed closely by the compromise

Table 3. Agreement* of reference standard and algorithm cause of death assignment among neonates

Scenario	Arifeen et al. 2004 [22]	Baqui et al. 2006 [23]	Kalter et al. 2015 [21]	Lawn et al. 2006 [24]	Liu et al. 2015 [25]	Compromise
Cause-specific mortality fraction accuracy:						
High U5MR with malaria	0.68 (0.53–0.76)	0.87 (0.77–0.93)	0.68 (0.53–0.76)	0.87 (0.77–0.93)	0.71 (0.56–0.79)	0.89 (0.77–0.93)
High U5MR without malaria	0.65 (0.45–0.74)	0.84 (0.75–0.93)	0.65 (0.45–0.74)	0.84 (0.75–0.93)	0.68 (0.49–0.78)	0.86 (0.71–0.93)
Moderate U5MR	0.49 (0.38–0.63)	0.78 (0.69–0.87)	0.49 (0.38–0.63)	0.78 (0.69–0.87)	0.53 (0.41–0.67)	0.74 (0.61–0.85)
General	0.71 (0.39–0.96)	0.80 (0.57–0.96)	0.74 (0.41–0.94)	0.77 (0.61–0.93)	0.75 (0.44–0.95)	0.76 (0.60–0.93)
Cohen's kappa:						
High U5MR with malaria	0.17 (0.12–0.22)	0.29 (0.23–0.35)	0.17 (0.12–0.22)	0.29 (0.23–0.35)	0.18 (0.13–0.23)	0.26 (0.21–0.31)
High U5MR without malaria	0.17 (0.11–0.22)	0.29 (0.24–0.36)	0.17 (0.11–0.22)	0.29 (0.24–0.36)	0.18 (0.12–0.22)	0.26 (0.21–0.32)
Moderate U5MR	0.15 (0.04–0.21)	0.28 (0.16–0.34)	0.15 (0.04–0.21)	0.28 (0.16–0.34)	0.16 (0.04–0.21)	0.24 (0.11–0.29)
General	0.14 (0.07–0.23)	0.20 (0.08–0.36)	0.15 (0.06–0.24)	0.24 (0.12–0.37)	0.16 (0.07–0.25)	0.22 (0.11–0.33)
Chance corrected concordance:						
High U5MR with malaria	0.13 (0.06–0.19)	0.22 (0.16–0.28)	0.13 (0.06–0.19)	0.22 (0.16–0.28)	0.14 (0.07–0.20)	0.20 (0.14–0.26)
High U5MR without malaria	0.13 (0.09–0.18)	0.22 (0.17–0.28)	0.13 (0.09–0.18)	0.22 (0.17–0.28)	0.14 (0.09–0.19)	0.20 (0.15–0.26)
Moderate U5MR	0.13 (0.09–0.22)	0.22 (0.17–0.44)	0.13 (0.09–0.22)	0.22 (0.17–0.44)	0.14 (0.09–0.23)	0.20 (0.15–0.36)
General	0.14 (0.08–0.22)	0.23 (0.16–0.33)	0.12 (0.05–0.21)	0.23 (0.17–0.32)	0.13 (0.06–0.22)	0.21 (0.14–0.30)

U5MR – under 5 years mortality rate

*Median and range across 1000 simulated instances of the Population Health Metrics Research Consortium study data for cause-specific mortality fraction (CSMF) accuracy, the kappa statistic, and chance corrected concordance (CCC) by mortality scenario and hierarchical method for distributing co–morbid causes of neonatal death, for four causes: birth asphyxia, congenital malformation, prematurity, sepsis/pneumonia or sepsis/pneumonia/meningitis.

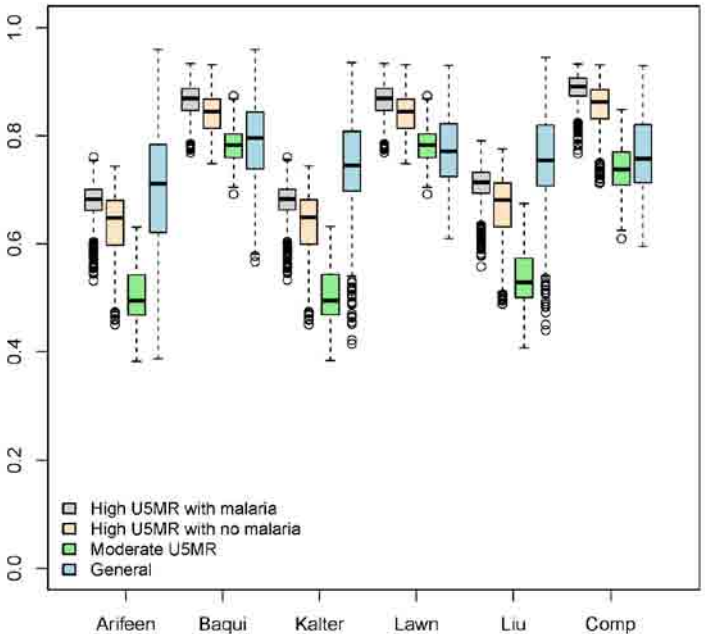


Figure 1. Cause-specific mortality fraction accuracy for six neonatal expert algorithm hierarchies in the resampled Population Health Metrics Research Consortium data, for 1000 simulated cause distributions from four neonatal mortality scenarios, and four neonatal causes (birth asphyxia, congenital malformation, prematurity, sepsis/pneumonia or sepsis/pneumonia/meningitis). Boxes represent interquartile ranges, with a line at the median. Whiskers represent 95% confidence intervals for the median values, and outliers are shown by dots.

hierarchy. Generally, for all algorithms, the Cohen's kappa was between 0.1 and 0.4, indicating minimal to fair agreement between VA estimated and reference standard causes. The CCC statistic also indicates that expert algorithms in the Baqui and Lawn hierarchies provide estimates that are closer to the reference standard causes than either the Kalter or Liu hierarchies, but overall the CCC statistics for all the hierarchies are between 0 and 0.45, indicating small to moderate agreement with the reference standard causes.

The median and range of absolute differences between estimated and reference standard CSMFs are shown in **Table 4** for each neonatal cause of death, along with the proportion of deaths that were not classified by each hierarchy. **Figure 2** shows the simulated reference standard and estimated CSMF in the general mortality scenario. This difference is identical across all hierarchies for the percent of deaths due to congenital malformations, because this cause is the first in each hierarchy. The Baqui and Lawn hierarchies perform best for birth asphyxia, and Baqui is best for sepsis/pneumonia. The compromise hierarchy is best for prematurity, and the Lawn and compromise hierarchies are jointly best for sepsis/meningitis/pneumonia.

Table 4. Absolute difference* between the cause-specific mortality fraction of each estimated and reference standard cause, for the general neonatal mortality scenario

Cause	Arifeen et al. 2004 [22]	Baqui et al. 2006 [23]	Kalter et al. 2015 [21]	Lawn et al. 2006 [24]	Liu et al. 2015 [25]	Compromise
Birth asphyxia	0.11 (0.00–0.22)	0.07 (0.01–0.15)	0.11 (0.00–0.22)	0.07 (0.01–0.15)	0.11 (0.00–0.22)	0.11 (0.00–0.22)
Congenital malformation	0.13 (0.03–0.31)	0.13 (0.03–0.31)	0.13 (0.03–0.31)	0.13 (0.03–0.31)	0.13 (0.03–0.31)	0.13 (0.03–0.31)
Meningitis	–	–	0.09 (0.01–0.24)	–	0.10 (0.01–0.24)	–
Pneumonia	–	–	0.19 (0.01–0.32)		0.20 (0.01–0.32)	–
Prematurity	0.12 (0.02–0.30)	0.08 (0.00–0.17)	0.12 (0.02–0.30)	0.08 (0.00–0.17)	0.11 (0.00–0.28)	0.05 (0.00–0.13)
Sepsis	–	–	0.09 (0.00–0.28)		0.11 (0.00–0.30)	0.11 (0.00–0.30)
Sepsis/pneumonia	0.18 (0.02–0.36)	0.10 (0.00–0.24)	0.12 (0.01–0.29)		0.11 (0.00–0.27)	–
Sepsis/pneumonia/meningitis	–	–	0.22 (0.18–0.30)	0.03 (0.00–0.08)	0.19 (0.15–0.27)	0.03 (0.00–0.08)
Unspecified	0.15 (0.12– 0.19)	0.15 (0.12– 0.19)	0.14 (0.11–0.18)	0.14 (0.11–0.18)	0.14 (0.11–0.18)	0.14 (0.11–0.18)

*Median and range across one thousand simulations. Results are shown for six hierarchies as a proportion of all neonatal deaths

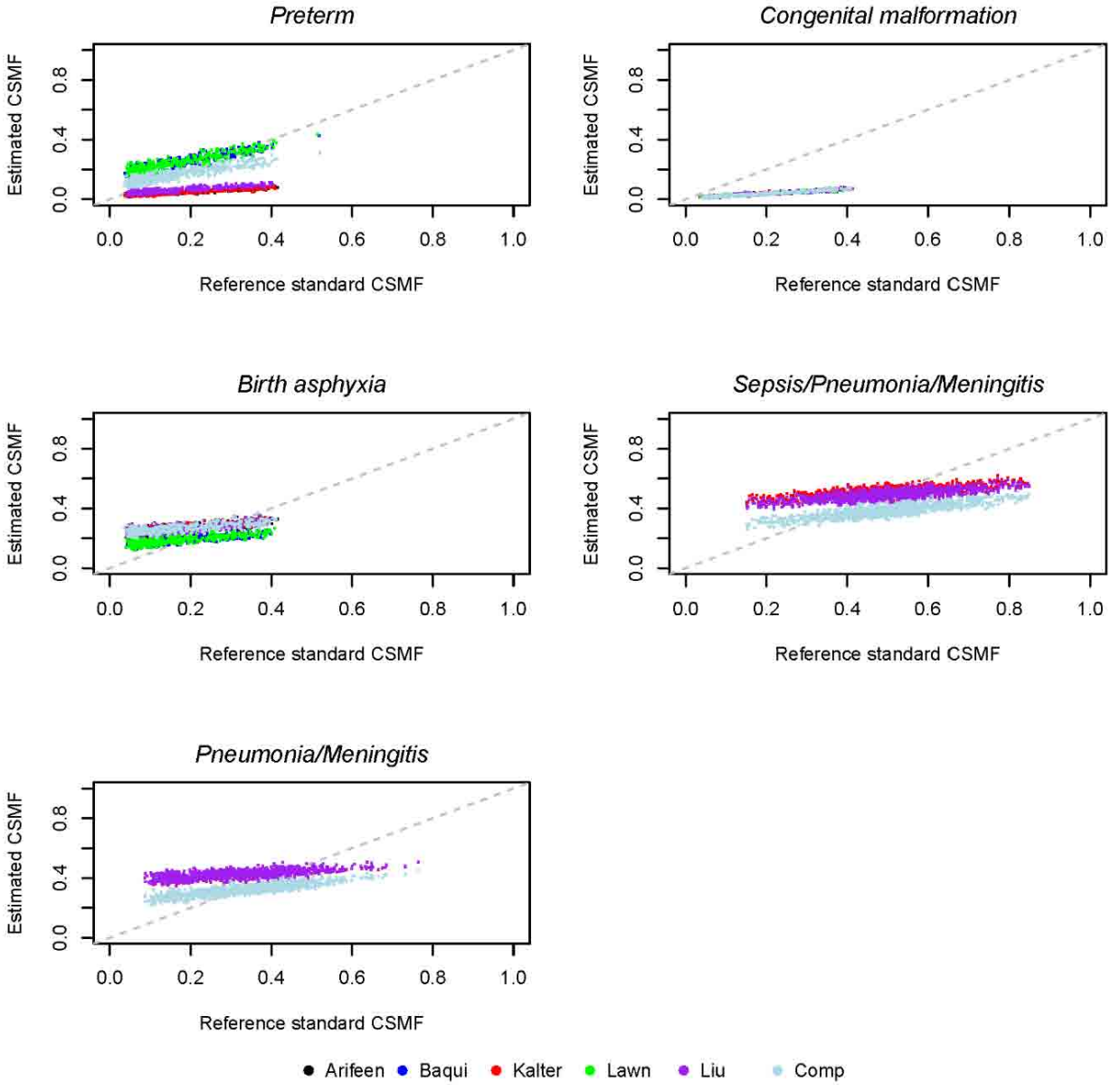


Figure 2. Cause-specific mortality fractions for six neonatal expert algorithm hierarchies in the resampled Population Health Metrics Research Consortium data, for four neonatal causes in the general neonatal mortality scenario, for 1000 simulated cause distributions.

Children

Table 5 shows summary results for the expert algorithm cause assignment of four causes of child deaths in three hierarchies and four mortality scenarios. We used the same three measures of accuracy as for neonatal deaths at the population and individual levels. At the population level, summarized by CSMF accuracy, the Kalter hierarchy performs best in each scenario. This population level comparison is also shown in **Figure 3**.

The hierarchies are not as strongly differentiated at the individual level for child deaths. There is also some counter indication at the individual level between Cohen's kappa and the CCC statistic as to which hierarchy is best in each mortality scenario. By Cohen's kappa, the three hierarchies are very similar in the high mortality with malaria and the general mortality scenarios. Also by Cohen's kappa, the Liu and Arifeen hierarchies are similar in the high mortality without malaria and moderate mortality scenarios, while the Kalter hierarchy has somewhat lower agreement. The Cohen's kappa for these three hierarchies generally range from slight (less than 0.2) to fair agreement (0.2–0.4).

By the CCC statistic, Arifeen's hierarchy has the largest median across the mortality scenarios, although the advantage

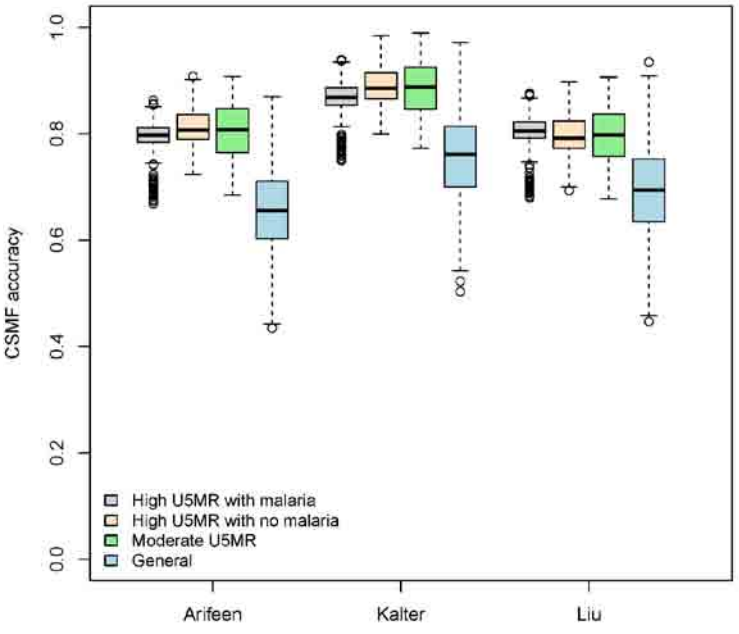


Figure 3. Cause-specific mortality fraction accuracy for three expert algorithm hierarchies in the resampled Population Health Metrics Research Consortium data, for 1000 simulated cause distributions from four mortality scenarios, and four causes of child death (pneumonia/diarrhea, measles, other infectious causes, and injury). Boxes represent interquartile ranges, with a line at the median. Whiskers represent 95% confidence intervals for the median values, and outliers are shown by dots.

Table 5. Agreement* of reference standard and algorithm cause of death assignment among children 1–59 months old

Scenario	Arifeen et al. 2004 [22]	Kalter et al. 2015 [21]	Liu et al. 2015 [25]
Cause-specific mortality fraction accuracy:			
High U5MR with malaria	0.80 (0.67–0.86)	0.87 (0.75–0.94)	0.81 (0.68–0.88)
High U5MR without malaria	0.83 (0.73–0.90)	0.93 (0.79–0.97)	0.80 (0.69–0.90)
Moderate U5MR	0.84 (0.74–0.90)	0.92 (0.79–0.97)	0.80 (0.68–0.91)
General	0.66 (0.43–0.87)	0.76 (0.50–0.97)	0.69 (0.45–0.93)
Cohen's kappa:			
High U5MR with malaria	0.14 (0.06–0.25)	0.13 (0.07–0.22)	0.14 (0.08–0.22)
High U5MR without malaria	0.24 (0.09–0.35)	0.21 (0.08–0.32)	0.23 (0.09–0.34)
Moderate U5MR	0.29 (0.07–0.35)	0.25 (0.08–0.32)	0.28 (0.08–0.35)
General	0.10 (0.02–0.38)	0.10 (0.04–0.33)	0.10 (0.04–0.35)
Chance corrected concordance:			
High U5MR with malaria	0.25 (0.20–0.49)	0.17 (0.12–0.39)	0.20 (0.14–0.42)
High U5MR without malaria	0.23 (0.18–0.45)	0.22 (0.17–0.46)	0.22 (0.18–0.44)
Moderate U5MR	0.40 (0.19–0.45)	0.37 (0.16–0.55)	0.39 (0.17–0.48)
General	0.24 (0.16–0.30)	0.16 (0.10–0.23)	0.19 (0.12–0.25)

U5MR – under 5 years mortality rate

*Median and range across 1000 simulated instances of the Population Health Metrics Research Consortium study data for cause-specific mortality fraction (CSMF) accuracy, the kappa statistic, and chance corrected concordance (CCC) by mortality scenario and hierarchical method for distributing co–morbid causes of child death, for four causes: pneumonia/diarrhea, measles, other infectious causes, and injury.

is small, especially for the high mortality without malaria and moderate mortality scenarios. Overall CCC statistics range from 0.06 to 0.55, indicating small to moderate agreement by the standards for interpreting Cohen's kappa, and somewhat higher agreement than for neonates.

Figure 4 shows the simulated reference standard and estimated CSMF in the general mortality scenario for six causes of child deaths. The median and range of absolute differences between estimated and reference standard CSMFs across these simulated instances of the PHMRC data for each cause of child death are shown in **Table 6**. This difference is identical across all hierarchies for the percent of deaths due to injuries, because this cause occupies the same place in the respective hierarchies. The Kalter hierarchy is best for pneumonia/diarrhea, meningitis/encephalitis, and AIDS. The Arifeen hierarchy is best for other infectious causes, while the Liu hierarchy is generally best for malaria. The Liu hierarchy is especially accurate when malaria is below 0.10 CSMF, while the Kalter hierarchy tends to be more accurate as malaria increases, as shown in **Figure 4**. **Table 6** shows the median absolute difference in CSMF, which may mask differences depending on the reference standard CSMF.

The median absolute differences between estimated CSMF and reference CSMF by cause are also shown for the two high mortality scenarios in **Table 6**, both with and without

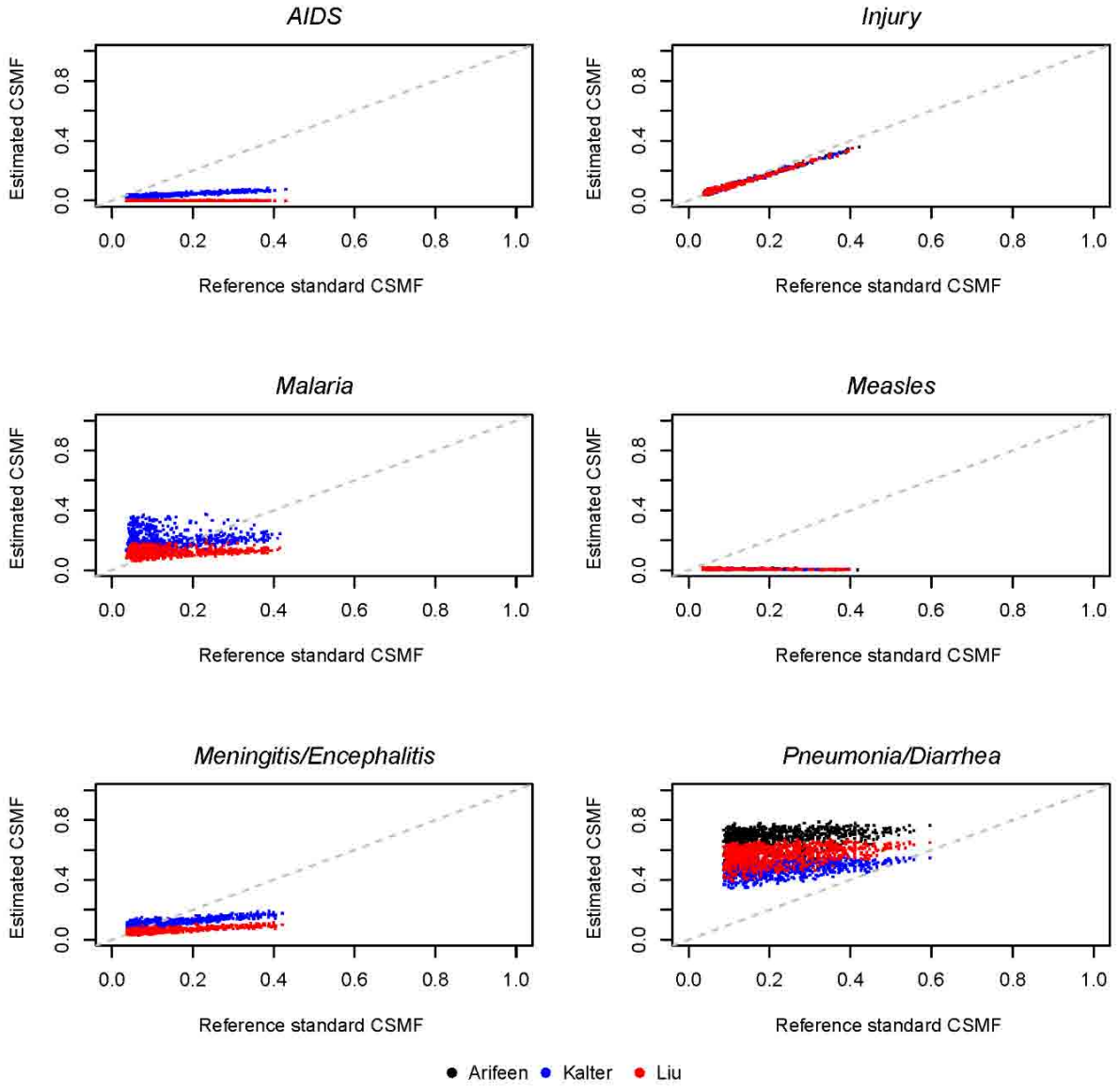


Figure 4. Cause-specific mortality fractions for three expert algorithm hierarchies in the resampled Population Health Metrics Research Consortium data, for six child causes in the general mortality scenario, for 1000 simulated cause distributions.

Table 6. Absolute difference* between the cause-specific mortality fraction of each estimated and reference standard cause, for the general, high mortality with malaria and high mortality without malaria child mortality scenarios

Cause	Arifeen et al. 2004 [22]	Kalter et al. 2015 [21]	Liu et al. 2015 [25]
Child – General mortality scenario:			
AIDS	–	0.08 (0.01–0.28)	0.08 (0.04–0.35)
Diarrhea/dysentery	–	0.08 (0.01–0.15)	–
Injury	0.01 (0.00–0.05)	0.01 (0.00–0.05)	0.01 (0.00–0.05)
Malaria	–	0.10 (0.01–0.26)	0.04 (0.00–0.22)
Measles	0.10 (0.03–0.36)	0.10 (0.03–0.36)	0.10 (0.03–0.36)
Meningitis/encephalitis	–	0.05 (0.00–0.20)	0.06 (0.00–0.27)
Other infectious causes	0.07 (0.00–0.25)	0.08 (0.00–0.32)	0.11 (0.04– 0.35)
Pneumonia	–	0.20 (0.03–0.30)	–
Pneumonia/diarrhea	0.46 (0.24–0.62)	0.26 (0.04–0.40)	0.37 (0.13–0.51)
Unspecified	0.47 (0.50-0.95)	0.11 (0.00-0.51)	0.14 (0.00-0.63)
Child – High mortality with malaria:			
AIDS	–	0.01 (0.00–0.10)	0.02 (0.00–0.13)
Diarrhea/dysentery	–	0.03 (0.00–0.08)	–
Injury	0.01 (0.00–0.01)	0.01 (0.00–0.01)	0.01 (0.00–0.01)
Malaria	–	0.09 (0.00–0.23)	0.15 (0.01–0.30)
Measles	0.01 (0.00–0.08)	0.01 (0.00–0.08)	0.01 (0.00–0.08)
Meningitis/encephalitis	–	0.08 (0.05–0.10)	0.02 (0.00–0.05)
Other infectious causes	0.03 (0.00–0.11)	0.09 (0.05–0.15)	0.12 (0.09–0.17)
Pneumonia	–	0.12 (0.06–0.23)	–
Pneumonia/diarrhea	0.35 (0.25–0.45)	0.14 (0.05–0.26)	0.24 (0.15–0.36)
Unspecified	0.45 (0.26–0.62)	0.06 (0.02–0.29)	0.10 (0.06–0.34)
Child – High mortality without malaria:			
AIDS	–	0.02 (0.00–0.30)	0.01 (0.00–0.37)
Diarrhea/dysentery	–	0.02 (0.00–0.09)	–
Injury	0.01 (0.00–0.02)	0.01 (0.00–0.02)	0.01 (0.00–0.02)
Malaria	–	–	–
Measles	0.02 (0.00–0.08)	0.02 (0.00–0.08)	0.02 (0.00–0.08)
Meningitis/encephalitis	–	0.06 (0.02–0.08)	0.14 (0.07–0.20)
Other infectious causes	0.03 (0.00–0.11)	0.09 (0.05–0.15)	0.12 (0.09–0.17)
Pneumonia	–	0.10 (0.06–0.16)	–
Pneumonia/diarrhea	0.22 (0.17–0.36)	0.09 (0.04–0.22)	0.19 (0.14–0.34)
Unspecified	0.26 (0.17–0.39)	0.10 (0.04–0.23)	0.16 (0.08–0.26)

*Median and range across one thousand simulations. Results are shown for three hierarchies, as a proportion of all child deaths.

malaria. The relative accuracy of the hierarchies by cause was similar to their performance in the general mortality scenario, except that in the high mortality with malaria scenario the Liu hierarchy did best for meningitis/encephalitis and the Kalter hierarchy worked best for malaria. The median absolute difference for pneumonia and diarrhea in the scenario for high mortality with malaria was 0.35, 0.14, and 0.24 for the Arifeen, Kalter, and Liu hierarchies respectively. These same median absolute differences in the high mortality scenario *without* malaria were 0.22, 0.09, and 0.19, indicating an improvement in estimated CSMF for pneumonia and diarrhea when the CSMF for deaths due to malaria was low. In addition, the median difference in the pneumonia CSMFs in the Kalter hierarchy was 0.12 in the high mortality scenario with malaria, and 0.10 in the high mortality scenario without malaria. These results reflect improved estimates for pneumonia, as expected given that high malaria burden may complicate other diagnoses, especially for pneumonia [31].

The software for the best performing neonatal and child algorithms and hierarchies, along with the PHMRC questionnaire needed to collect the input data, are available online [20].

DISCUSSION

We have compared six expert algorithm hierarchies for assigning causes of neonatal death and three for assigning causes of child death, and we compared the resulting cause distributions with reference standard causes. We made these comparisons among the PHMRC study data, resampled to resemble the cause proportions of deaths from a variety of community settings as determined by the Child Health Epidemiology Reference Group on behalf of WHO.

There was minimal to fair agreement between the algorithmic and the reference standard diagnoses at the individual level, both for neonatal and child causes of death, although some hierarchies had slightly higher agreement than others.

Verbal autopsies are generally used to describe populations instead of individuals, and so we have focused on measures of the agreement between algorithm–assigned and reference standard causes at the population level [32]. By this measure the agreement between assigned and reference standard causes was more favorable and the algorithms appear useful. When assessed in this manner, the Baqui, Lawn and compromise hierarchies performed best for neonatal causes, and the Kalter hierarchy performed best for children. The nearly equal performance of several hierarchies for neonatal deaths in the general mortality scenario suggests that several of the VA studies used as input data for the WHO/CHERG modeled estimates, whose cause distributions were the basis for the other mortality scenarios, may have used hierarchies with preterm placed higher up to select among multiple causes, similar to the ordering of diagnoses in the Baqui and Lawn hierarchies.

Hierarchy performance also varied across particular causes of neonatal and child death. For neonatal deaths, the Baqui and Lawn hierarchies performed best for birth asphyxia, and the compromise hierarchy performed best for prematurity. The Baqui hierarchy also performed best for sepsis/pneumonia, while the Lawn and compromise hierarchies performed best for sepsis/pneumonia/meningitis. For deaths in children 1–59 months, there was a striking difference in hierarchy performance for pneumonia, for which the Kalter hierarchy performed best. Clearly some causes are more difficult to classify than others. Hierarchy–estimated CSMF for child deaths due to injury was very close to the reference CSMF across all simulated scenarios. The estimated CSMF for measles, however, was near zero for all simulations, indicating a poor diagnostic ability, contrary to expectations for identifying measles [33]. This was likely due to an aberration in the PHMRC VA interview data, which identified 'rash' in only 3/23 reference standard measles cases [18].

Poor performance for particular causes may be masked by good overall performance as indicated by CSMF accuracy. For example, when an algorithm estimates 52%, 29%, 2% and 4% for neonatal deaths due to sepsis/pneumonia, birth asphyxia, congenital malformation, and prematurity, where the actual CSMFs are 32%, 31%, 11%, and 11%, the CSMF accuracy is 0.79, indicating good overall performance although sepsis/pneumonia is overestimated by 20%. Poor performance was observed for several causes in both neonates and children, where estimated CSMF was relatively flat over a range of reference standard CSMF. The CSMF accuracy as a statistic is limited in its ability to describe these details.

Until very recently the verbal autopsy standard was for questionnaires to be examined individually with cause of death determination by physician review. The new standard is to encourage assignment of cause of death using automated computer programs for the InterVA–4 and Tariff 2.0 methods [14]. The Tariff has been shown to outperform InterVA–4 in population level metrics, although reports vary [11,15]. The Tariff method determines cause based on the relative associations of symptoms and causes of death in a reference standard "training" data set, supplemented with global burden of disease estimates for questionnaires with undetermined cause of death [12].

In a validation study with the PHMRC data, CSMF accuracy of the Tariff 2.0 was reported at 0.81 (uncertainty 0.80, 0.82) for neonatal causes and 0.74 (uncertainty 0.74, 0.75) for child causes [12]. This is within the observed range of the best performing expert algorithm hierarchies (at 0.80 with range 0.57 to 0.96 for neonatal deaths and 0.76 with range 0.50 to 0.97 for child deaths), but with smaller uncertainty. The comparison, however, is not conclusive. Although the CSMF accuracy both of the expert algorithms and the Tariff were determined in the PHMRC data, only the Tariff was built on data from PHMRC study, potentially providing it with an advantage. In addition, the methods for resampling and estimating uncertainty were not the same, and so the reference is not necessarily on the same basis. In addition, the specified causes were not the same. For example, in the assessment of Tariff performance with neonatal causes of death, all deaths with co–morbid prematurity, birth asphyxia and/or sepsis were classified for resampling as being due to prematurity. The Tariff validation included six causes for neonatal deaths, and 21 for children, which is more total causes than in our expert algorithm validation. A definitive comparison of the Tariff and expert algorithm methods is further complicated by computational requirements of the Tariff. A single selection of deaths can be used to validate the expert algorithms, but in addition to these, the Tariff requires a selection of reference deaths for training. This comparison is outside the scope of this paper, but an area for further research.

The expert algorithms are fully deterministic: verbal autopsies with the same responses will be assigned the same cause of death. Algorithms are based on symptom patterns that physicians and medical experts expect to correspond to common causes of death in neonates and children. This determinism is an asset for facilitating use and understanding. While InterVA is also deterministic, it relies on conditional probabilities of the relationships between symptoms and causes of death that operate unseen in the background, rendering it less easily explainable to non–medical profes-

sionals. The Tariff, in contrast to both the VA algorithms and InterVA methods, requires a reference selection, from which symptom patterns are determined. The circumstances that require verbal autopsy are precisely where the cause of death distribution is unknown, precluding the selection of a perfect reference. The Tariff method's sensitivity to this selection is not well understood, as a research friendly version has not been released. There is an unquantified potential for the Tariff to fail in the event that a poor reference is chosen.

This same determinism and predictability in the expert algorithm method that facilitates its use may be a liability in other respects. We observed some outlying cases of poor agreement between the predicted and reference standard cause fractions, although overall there was good agreement at the population level between algorithm and reference standard causes.

CONCLUSION

Verbal autopsy is an invaluable tool in settings where civil registration is unreliable or incomplete. Health policy makers and programmers need verbal autopsy to better understand the causes of neonatal and child deaths and how these deaths might have been prevented. Here we identify the most useful fixed algorithms and hierarchies for assigning cause of death in a deterministic manner. For neonates, these include the Compromise hierarchy to be used in high mortality settings and the Baqui hierarchy otherwise; while for 1–59 month–old children, the Kalter hierarchy performed best overall. These expert algorithms provide an accessible and systematic mechanism for interpreting verbal autopsy, on par with more complex machine learning methods that will soon replace the current standard. Work is ongoing to assess the feasibility of mapping the algorithms to the 2014 WHO VA questionnaire in order to render the method even more accessible

Acknowledgments: We would like to thank Elizabeth Ragan for her assistance in cleaning the PHMRC database.

Funding: Funding was provided by the Bill and Melinda Gates Foundation through a grant to the U.S. Fund for UNICEF for the Child Health Epidemiology Reference Group. The funder had no role in the study design, data collection and analysis, interpretation of data, decision to publish, or preparation of the manuscript.

Authorship declaration: HDK conceived the idea for conducting the hierarchical expert algorithm analysis in the PHMRC data, and developed the analysis plan. JP and HDK worked together to refine the analysis plan, and JP conducted the analysis. JP drafted the manuscript, with substantial contribution by HDK. All authors critically reviewed and commented on the draft paper and read and approved the final manuscript.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). The authors report no conflict of interest.

1 Lawn JE, Wilczynska–Ketende K, Cousens SN. Estimating the causes of 4 million neonatal deaths in the year 2000. Int J Epidemiol. 2006;35:706-18. Medline:16556647 doi:10.1093/ije/dyl043

2 Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis. Lancet. 2010;375:1969-87. Medline:20466419 doi:10.1016/S0140-6736(10)60549-1

3 Liu L, Johnson HL, Cousens S, Perin J, Scott S, Lawn JE, et al. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. Lancet. 2012;379:2151-61. Medline:22579125 doi:10.1016/S0140-6736(12)60560-1

4 Liu L, Oza S, Hogan D, Perin J, Rudan I, Lawn JE, et al. Global, regional, and national causes of child mortality in 2000–13, with projections to inform post–2015 priorities: an updated systematic analysis. Lancet. 2015;385:430-40. Medline:25280870 doi:10.1016/S0140-6736(14)61698-6

5 Soleman N, Chandramohan D, Shibuya K. Verbal autopsy: current practices and challenges. Bull World Health Organ. 2006;84:239-45. Medline:16583084 doi:10.2471/BLT.05.027003

6 Leitao J, Chandramohan D, Byass P, Jakob R, Bundhamcharoen K, Choprapawon C, et al. Revising the WHO verbal autopsy instrument to facilitate routine cause–of–death monitoring. Glob Health Action. 2013;6:21518. Medline:24041439

7 Kalter HD, Gray RH, Gultiano C, Black RE. Validation of postmortem interviews to ascertain selected causes of death in children. Int J Epidemiol. 1990;19:380-6. Medline:2376451 doi:10.1093/ije/19.2.380

REFERENCES

8 Mobley CC, Boerma JT, Titus S, Lohrke B, Shangula K, Black RE. Validation study of a verbal autopsy method for causes of childhood mortality in Namibia. J Trop Pediatr. 1996;42:365-9. Medline:9009566 doi:10.1093/tropej/42.6.365

9 Marsh DR, Sadruddin S, Fikree FF, Mehnaz A, Kisat H, Billoo AG, et al. Validation of diagnostic criteria for cause of child death by delayed interview of bereaved parent: a refined verbal autopsy method. Karachi, Pakistan: Department of Community Health Sciences, The Aga Khan University, 1995.

10 Kalter HD, Hossain M, Burnham G, Khan NZ, Saha S, Ali MdA Ali, et al. Validation of caregiver interviews to diagnose common causes of severe neonatal illness. Paediatr Perinat Epidemiol. 1999;13:99-113. Medline:9987789

11 Murray CJL, Lozano R, Flaxman A, Serina P, Phillips D, Stewart A, et al. Using verbal autopsy to measure causes of death: the comparative performance of existing methods. BMC Med. 2014;12:5. Medline:24405531 doi:10.1186/1741-7015-12-5

12 Serina P, Riley I, Stewart A, James S, Flaxman A, Lozano R, et al. Improving performance of the Tariff Method for assigning causes of death to verbal autopsies. BMC Med. 2015;13:291. Medline:26644140 doi:10.1186/s12916-015-0527-9

13 Byass P, Chandramohan D, Clark SJ, D'Ambruoso L, Fottrell E, Graham WJ, et al. Strengthening standardised interpretation of verbal autopsy data: the new InterVA–4 tool. Glob Health Action. 2012;5:1-8. Medline:22944365

14 World Health Organizationlation. Verbal autopsy standards: the 2014 WHO verbal autopsy instrument. Geneva: WHO, 2015.

15 Leitao J, Desai N, Aleksandrowicz L, Byass P, Miasnikof P, Tollman S, et al. Comparison of physician–certified verbal autopsy with computer coded verbal autopsy for cause of death assignment in hospitalized patients in low– and middle–income countries: systematic review. BMC Med. 2014;12:22. Medline:24495312 doi:10.1186/1741-7015-12-22

16 Mpimbaza A, Filler S, Katureebe A, Quick L, Chandramohan D, Staedke SG. Verbal autopsy: Evaluation of methods to certify causes of death in Uganda. PLoS ONE. 2015;10:e0128801. Medline:26086600 doi:10.1371/journal.pone.0128801

17 Murray CJL, Lopez AD, Black RE, Ahuja R, Ali SM, Baqui A, et al. Population Health Metrics Research Consortium gold standard verbal autopsy validation study: design, implementation, and development of analysis datasets. Popul Health Metr. 2011;9:27. Medline:21816095 doi:10.1186/1478-7954-9-27

18 Population Health Metrics Research Consortium. Population Health Metrics Research Consortium Gold Standard Verbal Autopsy Data 2005-2011. Available: http://ghdx.healthdata.org/record/population-health-metrics-research-consortium-gold-standard-verbal-autopsy-data-2005-2011. Accessed: 14 January 2016.

19 Byass P. Usefulness of the Population Health Metrics Research Consortium gold standard verbal autopsy data for general verbal autopsy methods. BMC Med. 2014;12:23. Medline:24495341 doi:10.1186/1741-7015-12-23

20 John Hopkins Bloomberg School of Public Health, Institute for International Programs, Maternal, Newborn and Child Cause of Death. Publications. *Available:* http://www.jhsph.edu/research/centers-and-institutes/institute-for-international-programs/current-projects/maternal-child-epidemiology-estimation/maternal-newborn-and-child-cause-of-death/index.html. Accessed: 14 January 2016.

21 Kalter HD, Roubanatou AM, Koffi A, Black RE. Direct estimates of national neonatal and child cause–specific mortality proportions in Niger by expert algorithm and physician–coded analysis of verbal autopsy interviews. J Glob Health. 2015;5:010415. Medline:25969734 doi:10.7189/jogh.05.010415

22 National Institute of Population Research and Training (NIPORT), Mitra and Associates, and ORC Macro. Bangladesh Demographic and Health Survey 2004. Dhaka, Bangladesh and Calverton, Maryland, USA: National Institute of Population Research and Training, Mitra and Associates, and ORC Macro, 2005. p.125–35.

23 Baqui AH, Darmstadt GL, Williams EK, Kumar V, Kiran TU, Panwar D, et al. Rates, timing and causes of neonatal deaths in rural India: implications for neonatal health programmes. Bull World Health Organ. 2006;84:706-13. Medline:17128340 doi:10.2471/BLT.05.026443

24 Lawn JE, Wilczynska–Ketende K, Cousens SN. Estimating the causes of 4 million neonatal deaths in the year 2000. Int J Epidemiol. 2006;35:706-18. Medline:16556647 doi:10.1093/ije/dyl043

25 Liu L, Li M, Cummings S, Black RE. Deriving causes of child mortality by re–analyzing national verbal autopsy data applying a standardized computer algorithm in Uganda, Rwanda and Ghana. J Glob Health. 2015;5:010414. Medline:26110053 doi:10.7189/jogh.05.010414

26 Kundel HL, Polansky M. Measurement of observer agreement. Radiology. 2003;228:303-8. Medline:12819342 doi:10.1148/radiol.2282011860

27 World Health Organization. International statistical classification of disease and related health problems, Tenth Revision (ICD–10), Geneva: World Health Organization; 1992.

28 Murray CJL, Lozano R, Flaxman AD, Vahdatpour A, Lopez AD. Robust metrics for assessing the performance of different verbal autopsy cause assignment methods in validation studies. Popul Health Metr. 2011;9:28. Medline:21816106 doi:10.1186/1478-7954-9-28

29 Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20:37-46. doi:10.1177/001316446002000104

30 Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159-74. Medline:843571 doi:10.2307/2529310

REFERENCES

31 Redd SC, Bloland PB, Kazembe PN, Patrick E, Tembenu R, Campbell CC. Usefulness of clinical case–definitions in guiding therapy for African children with malaria or pneumonia. Lancet. 1992;340:1140-3. Medline:1359219 doi:10.1016/0140-6736(92)93160-O

32 Fottrell E, Byass P. Verbal autopsy: methods in transition. Epidemiol Rev. 2010;32:38-55. Medline:20203105 doi:10.1093/epirev/mxq003

33 Anker M. The effect of misclassification error on reported cause–specific mortality fractions from verbal autopsy. Int J Epidemiol. 1997;26:1090-6. Medline:9363532 doi:10.1093/ije/26.5.1090

journal of global health

Verbal/Social Autopsy in Niger 2012–2013: A new tool for a better understanding of the neonatal and child mortality situation

Khaled Bensaïd[1], Asma Gali Yaroh[2], Henry D Kalter[3], Alain K Koffi[3], Agbessi Amouzou[4], Abdou Maina[5], Narjis Kazmi[3]

[1] UNICEF, Niger country office, Niamey, Niger (retired)
[2] Ministry of Health, Niamey, Niger
[3] Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
[4] UNICEF, New York City, NY, USA
[5] Institute National des Statistics, Niamey, Niger

Niger, one of the poorest countries in the world, recently used for the first time the integrated verbal and social autopsy (VASA) tool to assess the biological causes and social and health system determinants of neonatal and child deaths. These notes summarize the Nigerien experience in the use of this new tool, the steps taken for high level engagement of the Niger government and stakeholders for the wide dissemination of the study results and their use to support policy development and maternal, neonatal and child health programming in the country. The experience in Niger reflects lessons learned by other developing countries in strengthening the use of data for evidence–based decision making, and highlights the need for the global health community to provide continued support to country data initiatives, including the collection, analysis, interpretation and utilization of high quality data for the development of targeted, highly effective interventions. In Niger, this is supporting the country's progress toward achieving Millennium Development Goal 4. A follow–up VASA study is being planned and the tool is being integrated into the National Health Management Information System. VASA studies have now been completed or are under way in additional sub–Saharan African countries, in each through the same collaborative process used in Niger to bring together health policy makers, program planners and development partners.

There has been substantial global progress towards achieving the Millennium Development Goal 4 (MDG 4) of reducing child mortality by two–thirds between 1990 and 2015. However, few countries in Sub Saharan Africa will achieve MDG 4, and the regional under 5 mortality rate of 98 deaths per 1000 live births remains the highest in the world. Neonatal mortality has been the most difficult component to overcome, with the level in Sub Saharan Africa remaining at 32 deaths per 1000 live births [1]. Strategies required to further decrease child mortality include expanding health promotive and disease preventive practices in the community and by the health system, improved illness recognition and care-seeking by child caregivers, and increased access to quality health care [2,3]; while for neonates, additional measures include the provision of quality antenatal care, skilled birth attendance and normal newborn care, and increased access to basic and comprehensive emergency obstetric and neonatal care [4-7].

Correspondence to:
Dr. Khaled Bensaïd
10 Route du Paquier
Pringy 1663
Switzerland
kbensaid55@gmail.com

The delivery of health services is guided by health policies and implemented through programs developed at the national or sub–national level. However, the information needed for evidence–based decision making is often lacking or of low quality and not widely accessible in developing countries [8,9]. Obtaining reliable, valid and timely data, especially data related to the causes and determinants of neonatal and child mortality is challenging. National health management information systems (HMIS) collect data on a regular basis and have the potential to support ongoing service improvement and decision–making. Yet, the quality of HMIS data in developing countries is poor; and it also may not be generalizable due to its collection in facilities with less than universal coverage [10,11]. The integrated verbal and social autopsy (VASA) tool and methods were developed to help overcome these limitations by supporting the collection of reliable and high quality population–based data on community– and health system–related maternal morbidity and neonatal and child mortality indicators.

A second critical factor in achieving positive results is the health policy and program planning process whereby study findings are utilized for decision–making. Case studies in Ethiopia highlighted that in addition to the lack of data, other major barriers to using data for decision making were awareness that data exists, accessibility and formatting of data, poor demand and lack of capacity of policy makers to appreciate and use data for informed decision making [12]. A study from Madagascar found that if data users and producers work together from the point of development of tools, the resulting data would be easy to understand and better used both by program implementers and policy makers [13]. As shown by work in Egypt [14] and Mexico [15], interest and commitment on the part of government to use the data to develop needed health programs also are key elements. This was also supported by the fact that the commitment of the government of Uganda to ensure data availability at district level and their willingness to support implementation of strategies improved women's access to the selected health services [16].

Prior evidence supports that the use of verbal autopsy data in health policy development and program planning can lead to significant improvement in health outcomes. A verbal autopsy study in Kenya that focused on data use for decision making found that visual synthesis of data facilitated the use of information in health decision making at the district health system level and promoted program improvement [17]. In Egypt, making use of available verbal autopsy information in designing high mortality impact interventions resulted in a major decline in maternal mortality [14]; a similar pattern was observed in Indonesia [18]. Other countries also have utilized maternal and child mortality data to guide the development of interventions as well as in making informed policy and program decisions [15,19-21].

Niger is one of the few countries in sub–Saharan Africa that have achieved a tremendous reduction in mortality among children under the age of five years, placing the country on track to achieve the MDG 4. In 1990, over three in ten newborns in Niger were dying before reaching age five. The country has taken vigorous measures in the past decade to tackle this high mortality burden and succeeded in reducing child mortality by 65% between 1990 and 2012, with an annual rate of reduction of 4.8% [1]. Encouraged by this success, the Government has decided to strengthen its child survival strategy by addressing additional causes and determinants of mortality with a special emphasis on newborns, whose mortality rate declined insignificantly from 39 deaths per 1000 live births in 1998 to 33 in 2009 [22].

Based on the experiences of other countries in the use of mortality data for guiding program planning, and on the results of a recent study in Niger [22] that demonstrated the power of implementing "high mortality impact" interventions, the Government of Niger decided to conduct a VASA study. The study was implemented in 2012–2013 by the National Statistics Institute (INS) and a VASA working group including the Ministry of Health (MoH), the Niger country office of UNICEF and other partners in collaboration with the Johns Hopkins Bloomberg School of Public Health, which provided technical assistance on behalf of the WHO/UNICEF–supported Child Health Epidemiology Reference Group (CHERG).

The approach was to collect detailed information on the biological causes (verbal autopsy) and household, community and health system determinants (social autopsy) of death. The study was intended to provide critical data needed to revise the child survival strategy by reorganizing and reorienting health services and health interventions for improved care of pregnant and delivering mothers, newborns and children. Also critical to the goal of achieving maximal impact on maternal and child health policy and program development was the plan for the study partners to collaboratively analyze, interpret and disseminate the findings throughout the country.

VASA METHODOLOGY

Mapping deaths identified by the 2010 Niger National Mortality Survey (NNMS) was used to select nationally representative samples of 605 neonatal deaths (0–27 days) and 605 infant and child deaths (1–59 months) that occurred between 2007 and 2010. The VASA study considers the biological causes of death and three levels of determinants, including family (cultural), community (social) and health system factors that affect access to and utilization of health promotive and disease preventive and curative interventions. The data collection tools include the Population

Health Metrics Research Consortium verbal autopsy questionnaire [23] to determine the biological causes of death and the CHERG social autopsy questionnaire [24] to collect data on the determinants. The two questionnaires are chronologically blended together to identify the cause of death and actions that might have been taken before and during the illness to prevent the death.

The analytic methods are fully described elsewhere [25,26]. In brief, the causes of death were determined through the use of expert verbal autopsy algorithms arranged in a hierarchy (EAVA method) to select the primary cause of death, and by one physician certifying the VA underlying cause of death using pre–specified minimal diagnostic criteria together with her clinical judgment (PCVA method), followed by comparison of the EAVA and PCVA diagnoses to help assess their plausibility and reliability. The SA analysis examined the prevalence of preventive factors along the continuum of normal maternal, newborn and child care, and of potential curative factors and constraints to careseeking for the fatal illnesses of the neonates and children along the steps in the Pathway to Survival [24]. In addition, for neonatal deaths we examined maternal pregnancy and delivery complications, as well as careseeking, constraints to careseeking, and the care received for the complications.

It is important to note that conducting the VASA study on the platform of the earlier mortality survey greatly reduced the cost of the VASA compared to a conventional survey through direct interviewing of homes where deaths were reported by the NNMS.

The VASA study was approved by the National Consultative Ethics Committee of the Niger Ministry of Health and the Institutional Review Board of the Johns Hopkins Bloomberg School of Public Health. Informed consent was given by all study participants prior to their being interviewed.

SOME EXAMPLES OF RESULTS

The VASA study provided information previously not available in Niger either at the local or national level. Of particular concern to the country was the need for data on the causes of neonatal and child deaths, maternal complications contributing to neonatal mortality, and some determinants of access to care such as problems faced by mothers to go to a health care facility for their maternal complications late in pregnancy and during labor and delivery and for their sick neonates and children.

The verbal autopsy showed that from 2007–2010 two thirds of neonatal deaths occurred during the first six days of life; more than one in two newborns died from infectious diseases and one in six succumbed to neonatal asphyxia. As for deaths of children aged 1–59 months, just under half occurred before the first birthday, and infectious diseases caused more than 90% of the deaths.

Unexpected results were revealed. For example, in a country where two out of three births take place at home, the pregnant woman herself decided the place of birth in seven of 10 cases, against 18% for her husband and 6% for her mother and mother–in–law.

Paradoxes were identified that require further study to better understand certain behaviors. While five in seven women delivered at home, less than one in three of these women said they had a problem that prevented them from giving birth at a health care provider. Problems of distance and transport were cited most often, by almost 80% of these women, as an obstacle to delivering at formal services, against only 4% of the women who were afraid to be exposed to a male health provider.

In addition, nearly 96% of the mothers and other caregivers of the neonates whose illness began at home reported a possibly severe or severe illness sign or symptom such as fast breathing, fever, or not able to feed, but many either did not recognize or minimized their meaning and importance. As a result, two-thirds of the mothers did not deem it necessary to seek care outside of the home even in the presence of these illness signs. For those who did seek formal health care, the average time before deciding to go was nearly 2 days.

Finally, the study showed very limited access to hospital services, which were received, respectively, by only 3.1% of newborns delivered at home and 13.3% of children aged 1–59 months. Referral was minimal for the vast majority who were seen by a first level provider. Only 8.7% of 69 neonates and 19.0% of 306 children who left the first provider alive were referred.

The complete study results for neonates and children 1–59 months are available in a prior publication [25] and two new papers [27,28] in the current issue of the Journal.

USE OF THE VASA DATA FOR POLICY AND PROGRAM DEVELOPMENT

A favorable environment for effective use of the VASA data existed in Niger from the start due to the government's interest in the study as a means of continuing the momentum toward decreasing child mortality and achieving MDG 4. This was reinforced by the CHERG's emphasis on the early formation of a country working group (CWG) of child survival experts, policy makers and program planners and their active participation in the analysis, interpretation and dissemination of the study findings to a wide circle of stakeholders.

An early meeting of CHERG, the INS and the MoH identified the crucial need for information on the causes and determinants of under–five deaths to help guide the review of the 2011–2015 National Health Development and Child Survival plan and the National Maternal and Neonatal roadmap. The first CWG meeting in 2012 included participants from the INS, MoH, UNICEF and WHO.

Two dissemination meetings were held in 2013. The first brought together regional and central health authorities of Niger and neighboring countries to get acquainted with the VASA data, of a type rarely obtained before and on a national scale, and to broadly assess what the tool itself can produce as strategic information for the planning and evaluation of health interventions. The second involved 42 district health teams, which developed health policy and program recommendations and considered operational aspects of including the VASA in a new pillar of the National HMIS.

RECOMMENDATIONS AND THEIR IMPLEMENTATION

Three sets of recommendations emerged from these meetings. The initial assessment of the implementation of recommendations [29] was quite encouraging, especially as it was conducted only six months after the last dissemination workshop held in November 2013.

1. The first set of recommendations concerned strengthening the existing maternal and child health programs or interventions in light of the VASA study findings. Implementation progressed as follows
   - i. The VASA study found that less than one–third of women with a neonatal death delivered with skilled attendance. As a result, it was recommended that the existing policy of free care for children be extended to deliveries. UNFPA had already funded free deliveries on a pilot basis in four of the country's eight regions in 2010, and following an evaluation extended this from 2011 to 2014. Based on the VASA meeting recommendation, the concept note on the reduction of maternal and neonatal mortality currently being drafted by the MoH is to include national scaling of the program of free deliveries. From 2015, funding is being provided by the Reproductive Maternal and Newborn Child Health (RMNCH) initiative through UNFPA.
   - ii. In response to the VASA finding that nearly 80% of the women with a problem in reaching a health facility for delivery cited lack of transportation and distance as the main barriers, the Government acquired 150 motorcycle ambulances for the transport of pregnant women referred from first level health facilities to hospital and is negotiating with mobile phone companies to improve communications between these facilities for the referral of women in labor. For transportation problems from home to a facility the government is also pilot testing bovine or mule wagons and boats in island areas.
   - iii. The VASA study finding that most neonates who died were not taken outside the home for health care provided an opportunity to discuss the need to accelerate taking community newborn case management to scale. The MoH held a workshop in early November 2014 to discuss the initial experience of scaling up mother and newborn community case management to 45 health posts in three health districts. Extension to 208 health posts in 20 additional health districts is planned, with a training of trainers' workshop having been conducted in May 2015.
   - iv. The VASA finding of limited access to quality health care led to accelerating the recruitment of qualified nurses for health posts and their transformation into health centers from an average of 50 health posts transformed per year before 2014 to nearly 100 units in 2014.
   - v. While introduction of the Pneumococcal vaccine was already planned, this was reinforced by the VASA findings, which showed that pneumonia was one of the main causes of death of under–5 children.
   - vi. The VASA study found that 96% of caregivers were able to report their deceased child's signs of severe illness, yet 69% and 21%, respectively, of newborns' and children's caregivers did not seek health care for the illness. A recommendation was made to develop an integrated communications plan on danger signs in pregnancy and child illness, but as of yet significant progress has not been made. The first step is to articulate a global view of communication and a comprehensive communication plan that can be negotiated with partners for its implementation.
2. The second set of recommendations concerned the use of the VASA findings in the planning and evaluation of district annual action plans and revision of the maternal and neonatal mortality reduction roadmap, the child survival strategy and the national health development plan. Implementation of the recommendations progressed as follows:
   - i. Regarding the maternal and neonatal mortality reduction roadmap:
     - a. A working group was established to write the roadmap concept note, establish consensus and the means of fundraising.
     - b. The VASA findings that antepartum and intrapartum hemorrhage, maternal sepsis and eclampsia were the three main maternal complications contributing to neonatal deaths raised the awareness of

decision makers and led to the implementation of new case management interventions. A census of existing human resources and technical equipment was conducted, followed by the endowment of essential equipment to all 42 districts. A program for capacity building in case management of the three complications was conducted and the interventions have been implemented.

c. Based on the VASA study finding that a very low percentage of the neonates that died were delivered by Caesarean Section, the program for training general practitioners in the "surgery of the district" was accelerated, including abdominal emergency procedures such as Caesarian Section.

ii. Revision of the child survival strategy involved the full utilization of the VASA data:

A ministerial decree on the mission of the national committee for the revision of the strategy document was signed [30], UNICEF mobilized the necessary funding, a National Technical Working Group was officially installed and its terms of reference were finalized. The final revision was entered in the MoH 2014 work plan and the UNICEF annual work plan.

The review coincided with the scaling up of the Reproductive, Maternal, Newborn and Child Health H4+ initiative involving four UN agencies (UNFPA, World Bank, UNICEF, WHO) and other partners and donors, which provided a portion of the necessary funds for the review. Many countries including Niger will benefit from this activity. A concept note was developed based on the revised child survival strategy with the goal of sharing the new strategy for its adoption and funding by all partners and stakeholders to further accelerate the reduction of child mortality.

iii. Regarding the national health development plan, a working group was formed to review the 2010–2015 plan and elaborate the 2016–2020 plan on the basis of available evidence, including the next VASA study.

3. The third set of recommendations concerned strengthening the National HMIS through integration of the VASA tool with the UNICEF Multiple Indicator Cluster Survey (MICS) or the Niger National Mortality Survey. Implementation of the recommendations progressed as follows:

i. The VASA tool was officially adopted by the political leadership, and discussions with the National Institute of Statistics to formalize its integration in the National System for Development of Statistics (NSDS) are under way. This will be accompanied by new regulatory measures to define the roles and responsibilities of the involved parties. Also, the process for adopting and integrating the VASA tool within the National HMIS is being identified.

ii. Possible means of conducting the VASA on a regular basis are being examined, including a combined MICS/VASA survey and integration with all future Niger National Mortality Surveys. Technical assistance may be required for the initial round, both for integration of the VASA with the survey and its articulation with the entire HMIS. A possible obstacle concerns the availability of funding, though some partners, including the World Bank and the Common Fund/Swap partners have already shown interest.

DISCUSSION

Like many countries in the region, the formulation of health policies and strategies in Niger is based on analyses of surveys that are conducted every five or six years, such as the MICS and Demographic and Health Survey. These surveys provide information mostly on system performance in terms of coverage of preventive and curative services, but rarely on aspects related to population demand. The country's routine health information system, which is the mainstay for the daily management of services and programs, deals almost exclusively with the surveillance of notifiable diseases. This pillar also suffers from limited reliability due to uncertainty of the relevant denominators, and so is rarely used for central planning of medium and long term programs. Lastly, some additional studies are occasionally conducted in limited geographic areas and used to analyze the situation of particular programs such as Reproductive Health. Faced with this ongoing gap in required data, as well as the evaluation of the 2010–2015 National Health Development Plan, achieving MDG 4 and the accelerated drive to reduce maternal mortality, reform of the HMIS has been an active topic of recent discussion. This has included a review of all indicators, the frequency of their production, and how they are used for better management of services and resources.

While recent years have seen an improvement in the coordination between departments and programs, difficulties remain in the use of data in the decision–making process. These include, among others, limited dissemination and use of data, often due to the organization of surveys and studies by individual programs leading to weak communication between programs and between programs and sources of support; failure to bring together all relevant departments overseeing programs and resources in the discussion of the data and planning exercises; inadequate financing resulting in weak field monitoring that fails to ensure the necessary feedback to decision makers; and, finally, in the type of data available, usually concerning the

performance of intervention coverage and rarely identifying the demand for services or the barriers to access to care.

Several of the above problems evoke lessons learned in other developing countries regarding how to increase the utilization of data for health policy and program development, from the need for population–based, representative data [10] and the lack of data accessibility and low capacity of policy makers to use data for decision making in Ethiopia [12] to the need for better communication between data producers and users in Madagascar [13]. However, the situation in Niger additionally suggests the need for international donors and partners to not only support large surveys of interest to the global community, but also to more closely examine countries' data needs and contribute to strengthening the processes that will promote the local use of data for improved policy and program development.

In 2007 Niger introduced an effort to use new determinants in the National HMIS, based on the model of Tanahashi [31] and mainly to identify key bottlenecks for preventive services, such as those used to calculate the availability of services and resources, access to services, utilization, adequate coverage and, finally, effective coverage. However, this endeavor was not supported by adequate data, nor by official technical guidelines or training documents for taking the effort to scale.

Thus, the need for population–based data on the determinants of access to care and barriers to the use of services as well as data on causes of death became imperative to accelerate the achievement of MDG 4. The VASA study helped fulfill these data needs, and its implementation model of bringing together a country working group of health policy makers, program planners and international partners to help analyze and interpret the findings, draw conclusions and recommendations and plan next steps, increased the likelihood of the information being used for the improvement of maternal and child health interventions.

The dissemination of the VASA results with the participation of all district health teams further responded to the need for the use of relevant, population–based data in a planning exercise bringing together decision makers, program managers and resource providers, and opened new horizons toward strengthening the health information system in particular and the system of planning, monitoring and evaluation for better governance in general. The VASA study corroborated past evidence of the benefits of utilizing verbal autopsy data for health policy development and program planning [16,17], like some past studies included the collection of extensive data on social determinants in addition to causes of death [19-21], and advanced on this past work through its highly collaborative implementation model.

The assessment exercise undertaken in mid–2014 [29] to examine the follow–up to the recommendations made at the second VASA dissemination meeting in November, 2013 helped revitalize the process. Contacts with the Minister of Health, assembling of the MoH team with INS staff, and providing feedback to the Technical and Financial Partners in the presence of the MoH team facilitated the taking of important decisions. It is apparent that, although the VASA study uniquely met the self–identified data needs of the Government of Niger and was undertaken through a highly collaborative process aimed at promoting the use of the data for evidence–based decision making, continued support from technical and resource partners is still critical to achieve this goal along the road to MDG 4.

CONCLUSIONS

The VASA study allowed the health authorities in Niger to better understand the causes of death and constraints to accessing and utilizing care, as well as weaknesses of the health system in ensuring optimal responses to the health problems of mothers and children. The Niger Government and its health partners will continue their collaboration towards the goal of making the VASA tool a component of the HMIS and a basic reference of choice for developing strategic maternal and child health policies and programs. The experience of the VASA study in Niger helped to refine the tool, which has since been implemented in several additional countries.

Despite the critical mass of recommendations already implemented, continued input by technical and financial partners is still needed to stimulate the full utilization and integration of the VASA tool into the HMIS and the NSDS and monitoring of the development and implementation of the various recommendations. This will accelerate the process including identifying and supporting new funding.

Acknowledgments: Oumarou Habi of the National Statistics Institute of Niger oversaw the collection of the VASA data.

Funding: Funding for the VASA study in Niger was provided by the Bill and Melinda Gates Foundation through a grant to the U.S. Fund for UNICEF for the Child Health Epidemiology Reference Group, and by the UNICEF country office of Niger. The funders had no role in the study design, data collection and analysis, interpretation of data, decision to publish, or preparation of the manuscript.

Authorship declaration: KB conceived the idea to write on the use of the VASA findings in the formulation of health policy and programs in Niger, and its potential for like use in other countries. KB and HDK developed the approach to the paper. HDK, AKK and AM were responsible for the acquisition of the VASA data. AKK and HDK analyzed the data. KB, AGY, HDK, AKK, AA and AM interpreted the findings. KB drafted the paper, with significant input from HDK and NK. AGY ensured of the paper's accuracy in its description of the VASA's contribution to health policy development in Niger. All authors critically reviewed and provided comments on earlier drafts of the paper and read and approved the final manuscript.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). At the time of the VASA study, KB was the Senior Chief of Health and Nutrition of the UNICEF country office of Niger and AGY was the Director General of Reproductive Health of the Niger Ministry of Health. As such, KB influenced the development of health policy in Niger and AGY participated in its formulation and implementation.

REFERENCES

1 UN Inter–agency Group for Child Mortality Estimation (IGME). Levels and trends in child mortality, report 2013.

2 Jones G, Steketee RW, Black RE, Bhutta ZA, Morris SS; The Bellagio Child Survival Study Group. How many child deaths can we prevent this year? Lancet. 2003;362:65-71. Medline:12853204 doi:10.1016/S0140-6736(03)13811-1

3 Bryce J, Victora CG, Black RE. The unfinished agenda in child survival. Lancet. 2013;382:1049-59. Medline:24054535 doi:10.1016/S0140-6736(13)61753-5

4 Darmstadt GL, Bhutta ZA, Cousens S, Adam T, Walker N, de Bernis L. Evidence–based, cost–effective interventions: how many newborn babies can we save? Lancet. 2005;365:977-88. Medline:15767001 doi:10.1016/S0140-6736(05)71088-6

5 Darmstadt GL, Marchant T, Claeson M, Brown W, Morris S, Donnay F, et al. A strategy for reducing maternal and newborn deaths by 2015 and beyond. BMC Pregnancy Childbirth. 2013;13:216. Medline:24261785 doi:10.1186/1471-2393-13-216

6 Bhutta ZA, Das JK, Bahl R, Lawn JE, Salam RA, Paul VK, et al. Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? Lancet. 2014;384:347-70. Medline:24853604 doi:10.1016/S0140-6736(14)60792-3

7 Darmstadt GL, Kinney MV, Chopra M, Cousens S, Kak L, Paul VK. Who has been caring for the baby? Lancet. 2014;384:174-88. Medline:24853603 doi:10.1016/S0140-6736(14)60458-X

8 Qazi MS, Ali M. Health Management Information System utilization in Pakistan: challenges, pitfalls and the way forward. Biosci Trends. 2011;5:245-54. Medline:22281538 doi:10.5582/bst.2011.v5.6.245

9 Zielinski C, Kebede D, Mbondji PE, Sanou I, Kouvividila W, Lusamba–Dikassa PS. Knowledge systems in health in the WHO African Region: results of a questionnaire–based survey. J R Soc Med. 2014;107(1 suppl):22-7. Medline:24643664 doi:10.1177/0141076813518525

10 Abajebel S, Jira C, Beyene W. Utilization of Health Information System at district level in Jimma Zone Oromia Regional State, South West Ethiopia. Ethiop J Health Sci. 2011;21 Suppl 1:65-76. Medline:22435010

11 Nyamtema AS. Bridging the gaps in the Health Management Information System in the context of a changing health sector. BMC Med Inform Decis Mak. 2010;10:36. Medline:20579336 doi:10.1186/1472-6947-10-36

12 Demographic data for development decisionmaking: case studies from Ethiopia and Uganda Available at: http://www.prb.org/pdf09/demographicdatadecisionmaking.pdf. Accessed 1 July 2014.

13 Nutley T, Reynolds H. Improving the use of health data for health system strengthening. Glob Health Action. 2013;6:20001. Medline:23406921

14 Campbell O, Gipson R, Issa AH, Matta N, El Deeb B, El Mohandes A, et al. National maternal mortality ratio in Egypt halved between 1992–93 and 2000. Bull World Health Organ. 2005;83:462-71. Medline:15976898

15 Gutiérrez G, Reyes H, Martinez H, Tomé P, Guiscafré H. Study of the disease–health seeking–death process: another use of the verbal autopsy. Int J Epidemiol. 1994;23:427-8. Medline:8082972 doi:10.1093/ije/23.2.427

16 Krueger K, Akol A, Wamala P, Brunie A. Scaling up community provision of injectables through the public sector in Uganda. Stud Fam Plann. 2011;42:117-24. Medline:21834413 doi:10.1111/j.1728-4465.2011.00271.x

REFERENCES

17 Nutley T, McNabb S, Salentine S. Impact of a decision-support tool on decision making at the district level in Kenya. Health Res Policy Syst. 2013;11:34. Medline:24011028 doi:10.1186/1478-4505-11-34

18 Supratikto G, Wirth ME, Achadi E, Cohen S, Ronsmans C. A district-based audit of the causes and circumstances of maternal deaths in South Kalimantan, Indonesia. Bull World Health Organ. 2002;80:228-34. Medline:11984609

19 Schumacher R, Swedberg E, Diallo MO, Keita DR, Kalter HD, Pasha O. Mortality study in Guinea: Investigating the causes of death for children under 5. The Basic Support for Institutionalizing Child Survival Project (BASICS II). Arlington, VA; 2002. Available: http://www.basics.org/documents/pdf/guinea_mort.pdf. Accessed: 14 May 2014.

20 Aguilar AM, Alvarado R, Cordero D, Kelly P, Zamora A, Salgado R. Mortality survey in Bolivia: the final report. Investigating and identifying the causes of death for children under five. The Basic Support for Institutionalizing Child Survival (BASICS) Project. Arlington, VA; 1998. Available: http://pdf.usaid.gov/pdf_docs/PNACF082.pdf. Accessed: 14 May 2014.

21 Källander K, Kadobera D, Williams TN, Nielsen RT, Yevoo L, Mutebi A, et al. Social autopsy: INDEPTH Network experiences of utility, process, practices, and challenges in investigating causes and contributors to mortality. Popul Health Metr. 2011;9:44. Medline:21819604

22 Amouzou A, Habi O, Bensaïd K. Reduction in child mortality in Niger: a Countdown to 2015 country case study. Lancet. 2012;380:1169-78. Medline:22999428 doi:10.1016/S0140-6736(12)61376-2

23 Murray CJL, Lopez AD, Black RE, Ahuja R, Ali SM, Baqui A, et al. Population Health Metrics Research Consortium gold standard verbal autopsy validation study: design, implementation, and development of analysis datasets. Popul Health Metr. 2011;9:27. Medline:21816095 doi:10.1186/1478-7954-9-27

24 Kalter HD, Salgado R, Babille M, Koffi A, Black RE. Social autopsy for maternal and child deaths: a comprehensive literature review to examine the concept and the development of the method. Popul Health Metr. 2011;9:45. Medline:21819605 doi:10.1186/1478-7954-9-45

25 Kalter HD, Roubanatou AM, Koffi A, Black RE. Direct estimates of national neonatal and child cause–specific mortality proportions in Niger by expert algorithm and physician–coded analysis of verbal autopsy interviews. J Glob Health. 2015;5:010415. Medline:25969734 doi:10.7189/jogh.05.010415

26 Koffi AK, Libite PR, Moluh S, Wounang R, Kalter HD. Social autopsy study identifies determinants of neonatal mortality in Doume, Nguelemendouka and Abong–Mbang health districts, Eastern Region of Cameroon. J Glob Health. 2015;5:010413. Medline:26171142 doi:10.7189/jogh.05.010413

27 Koffi AK, Maina A, Yaroh AG, Habi O, Bensaïd K, Kalter HD. Social determinants of child mortality in Niger: Results from the 2012 National Verbal and Social Autopsy Study. J Glob Health. 2016;6:010603. doi:10.7189/jogh.06.010603

28 Kalter HD, Yaroh AG, Maina A, Koffi AK, Bensaïd K, Amouzou A, Black RE. Verbal/social autopsy study helps explain the lack of decrease in neonatal mortality in Niger, 2007–2010. J Glob Health. 2016;6:010604. doi:10.7189/jogh.06.010604

29 Bensaïd K. Verbal and social autopsy study on the causes and determinants contributing to neonatal and child deaths in Niger. Report to the CHERG of an evaluation trip to Niamey, 12–21 July, 2014.

30 Ministry of Health. Ministerial decree 211 MSP/DGSP/DSME of June 30, 2014, pertaining to the establishment and organization of the working group in charge of the elaboration of the country's proposal on the Global Initiatives to Accelerate Progress in RMNCH. Niamey, Niger: Ministry of Health, 2014.

31 Tanahashi T. Health service coverage and its evaluation. Bull World Health Organ. 1978;56:295-303. Medline:96953

Electronic supplementary material:
The online version of this article contains supplementary material.

journal of global health

Social determinants of child mortality in Niger: Results from the 2012 National Verbal and Social Autopsy Study

Alain K Koffi[1,2], Abdou Maina[3], Asma Gali Yaroh[4], Oumarou Habi[3], Bensaïd[5], Henry D Kalter[1,2]

[1] The Institute for International Programs, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
[2] Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
[3] Institut National de la Statistique, Niamey, Niger
[4] Ministry of Health, Niamey, Niger
[5] UNICEF/Niger country office, Niamey, Niger (retired staff)

Background Understanding the determinants of preventable deaths of children under the age of five is important for accelerated annual declines – even as countries achieve the UN's Millennium Development Goals and the target date of 2015 has been reached. While research has documented the extent and nature of the overall rapid decline in child mortality in Niger, there is less clear evidence to provide insight into the contributors to such deaths. This issue is the central focus of this paper.

Methods We analyzed a nationally representative cross–sectional sample of 620 child deaths from the 2012 Niger Verbal Autopsy/Social Autopsy (VASA) Survey. We conducted a descriptive analysis of the data on preventive and curative care, guided by the coverage of proven indicators along the continuum of well child care and illness recognition and care–seeking for child illnesses encompassed by the BASICS/CDC Pathway to Survival model.

Results Six hundred twenty deaths of children (1–59 months of age) were confirmed from the VASA survey. The majority of these children lived in households with precarious socio–economic conditions. Among the 414 children whose fatal illnesses began at age 0–23 months, just 24.4% were appropriately fed. About 24% of children aged 12–59 months were fully immunized. Of 601 children tracked through the Pathway to Survival, 62.4% could reach the first health care provider after about 67 minutes travel time. Of the 306 children who left the first health care provider alive, 161 (52.6%) were not referred for further care nor received any home care recommendations, and just 19% were referred to a second provider. About 113 of the caregivers reported cost (35%), distance (35%) and lack of transport (30%) as constraints to care–seeking at a health facility.

Conclusion Despite Niger's recent major achievements in reducing child mortality, the following determinants are crucial to continue building on the gains the country has made: improved socio–economic state of the poor in the country, investment in women's education, adoption of the a law to prevent marriage of young girls before 18 years of age, and implementation of health programs that encourage breastfeeding and complementary feeding, immunization, illness recognition, prompt and appropriate care–seeking, and improved referral rates.

Correspondence to:
Dr Alain K. Koffi
Department of International Health
Johns Hopkins Bloomberg School of Public Health
615 N. Wolfe St., Suite E8610
Baltimore, MD, 21205, USA
akoffi1@jhu.edu

With a population over 15 million people in 2011, the Republic of Niger is West Africa's second–largest country [1]. This landlocked country is characterized by chronic food security issues, natural crises, including droughts, floods and locust infestation, and a level of poverty that reflects more than a decade of periodic political instability. Niger's poverty rate of 46.3% makes it one of the world's poorest countries. Per capita income, at $360, puts it at the very bottom of the 187 countries ranked by the United Nations Development Program's Human Development Index [2]. In this fragile nation, women and children suffer the greatest burden of poor health and inequality [3]. Niger's social indicators have improved significantly over the past two decades, as progress toward the Millennium Development Goals (MDGs) is a main priority of the government.

The Government of Niger's policies in support of universal access, provision of free health care for pregnant women and children, and strong nutrition programs have enabled the country to decrease child mortality at a pace that exceeded expectations. These policies are enshrined in general principles and international strategies such as primary health care and the Bamako Initiative. Thus, its health system is organized into three administrative and service levels: local/district, intermediary/regional and central/national. At the local level, public sector services are provided by community health posts (*Case de Santé*), integrated health centers (*Centres de Santé Intégrés*), and district hospitals. About 75% of health posts are staffed by CHWs (the rest by a nurse or midwife), and health centers and hospitals are staffed by at least one nurse, midwife or physician.

Recently, the Niger countdown case study showed far greater reductions in child mortality than in neighboring West African countries. In tandem with its efforts to tackle malnutrition, the government of Niger has put in place several measures to reduce childhood mortality. For the past few years, children under five have received free health care, while significant progress has been made in immunization coverage, recruitment of health staff and in the number of malaria cases treated. Collectively, these factors have contributed to a rapid reduction in the under–five child mortality rate, from 226 deaths per 1000 live births in 1995, to 128 deaths per 1000 in 2009 – a remarkable 43% reduction [4].

In preparing child mortality–reduction strategies in the post–2015 era, progress in reducing child deaths around the globe will require new and different strategies from those used to get the world to the current point. For instance, it is important for each country to know not only the magnitude of under–five mortality, but also the biological causes and social determinants of these deaths in order to assess needs and develop programs that will reduce avoidable child deaths more quickly. Thus, reliable direct estimates of the causes and the determinants of under–five deaths are needed to efficiently tailor evidence–based policies and programs.

A national verbal/social autopsy (VASA) study was conducted in Niger as part of the Child Health Epidemiology Reference Group's (CHERG) recent efforts to directly measure the causes and determinants of neonatal and child mortality in selected high–priority countries.

The current paper aims to complement the recently published verbal autopsy findings [5] and reports on the social autopsy data of post–neonatal deaths. The objective is to provide insights into modifiable family, household, and health system factors that contributed to the deaths of children (1–59 months) from 2007 to 2010 in Niger, information that will be vital to health policymakers in government and non–governmental organizations as they develop new policies and programs for better resource planning in the post–2015 period.

METHODS

Study sample/VASA instruments/Data collection

Details of the statistical sample size calculation, VASA instruments, and data collection are available elsewhere [5]. In summary, the sample of deaths included in the Niger VASA study was identified by the Niger National Mortality Survey (NNMS) conducted in July–August 2010. The VASA study aimed to examine samples of the most recent 605 neonatal (0 to 27 days old) and 605 child (1 to 59 months old) deaths. The final VASA sample consisted of 1166 (96.4%) completed interviews of 1210 attempted, including 93 stillbirths, 453 neonatal deaths and 620 child deaths, with mean interview recall periods of 3.5 years (range 2–5 years) for the neonatal deaths and 2.7 years (range 2–5 years) for the child deaths. The current study focuses solely on the deaths of the 620 children (1 to 59 months).

The VASA questionnaire blends the Population Health Metrics Research Consortium (PHMRC) VA questionnaire with the CHERG SA questionnaire [6]. The interviews were conducted in French and the two main languages of Niger, Haoussa and Zarma, using a CSProX [7] software application developed for the VASA study to assist interviewers to capture responses with minimal data entry errors in the field directly on netbook computers.

There were seven data collection teams, each led by a field supervisor. The interviewers were 12 female and eight male native speakers of Haoussa and/or Zarma. The teams completed the data collection in 55 days.

Statistical analyses

A descriptive analysis was conducted of the data on preventive and curative care, guided by the coverage of key indicators along the continuum of normal well child care and illness recognition and care–seeking for child illnesses encompassed by the BASICS/CDC Pathway to Survival model [8–10]. The 2010 NNMS data require the use of cluster sample weights to obtain nationally representative estimates. Thus all of the results presented here are weighted in order to compensate for threats to external validity inherent to the sample selection approach [11].

The list and definitions of some operational variables used throughout this paper are in the **Online Supplementary Document**.

All the interventions examined by this study have been shown to be efficacious and effective in promoting child survival and thus are included among the interventions examined by the Lives Saved (LiST) tool [10] or recommended by the WHO, and so should be accessible to all children.

The SA data also assessed factors that might help explain why desirable actions were not taken, including socioeconomic and demographic factors, recognition of illness severity, decision makers, and self–identified care–seeking constraining factors.

Ethical approval

The study was approved by the National Consultative Ethics Committee of the Niger Ministry of Health and by the Institutional Review Board of the Johns Hopkins Bloomberg School of Public Health. All the study personnel received training in ethical principles and practices for human subject's research, and informed consent was given by all study participants before the VASA interview was conducted.

RESULTS

A high proportion (91.5%) of the respondents, who were selected to be the child's main caretaker during the fatal illness, were the child's mother.

Socio–demographic characteristics of the deceased children (1–59 months) and their households

The sociodemographic characteristics of the deceased children are presented in **Table 1**. The median age at illness onset was 12 months (SD = 14.0; range 0–day to–48 months) and the median illness duration was 7 days (SD = 40.1; range 0–day to 10 months). The majority of deaths occurred in the post–neonatal (1–11 months of age) and second–year (12–23 months of age) periods, 44.5% and 21.8%, respectively. The data showed slight differentials between deaths of females and males, with a male ratio of 96. Most (71.8%) of the deceased children were born at home; the majority (n = 470 or 75.9%) also died at home. The vast majority (87.2%) of mothers did not have or could not present a vaccination card for the deceased child.

Table 1. General mortality indicators and demographic characteristics of 620 children (1–59 months) deaths, Niger, 2007–2010

Characteristics	Frequency (No.)	Percent
Median age at illness onset (in months)	12 (SD = 13.98)	
Median illness duration (in days)	7 (SD = 40.11)	
Median age at death (in months):	12 (SD = 13.86)	
1–11	276	44.5
12–23	135	21.8
24–59	209	33.7
Sex:		
Male	304	49.1
Female	316	50.9
Masculinity ratio (Boy/Girl×100)	96	
Place of birth:		
Hospital	22	3.5
Other health provider or facility	143	23.0
On route to a health provider or facility	8	1.4
Home	445	71.8
Other	2	0.4
Place of death:		
Hospital	52	8.3
Other health provider or facility	63	10.2
On route to a health provider or facility	24	3.9
Home	470	75.9
Other	10	1.6
Don't know	1	0.1
Child possessed a vaccination card:		
Yes, seen	79	12.8
Yes, not seen	352	56.8
No card	189	30.4

SD – standard deviation

Table 2 shows the characteristics of the mother, her domestic partner, and the household. Approximately 97% of the mothers were married or living with a man at the time of the interview; the vast majority (94%) entered into a union before 20 years of age and had little or no education. Indeed, 87.3% had 0 years of schooling. The occupation most cited for the father as the breadwinner was farmer/agricultural worker (70%). Average household size was 7.7 persons. Only 11% percent of the households had electricity, one in three had access to an improved source of drinking water, 7% used improved sanitation (flush or improved pit toilet) and 98% of the households used firewood for cooking. About 10% of the households had flooring made of cement and 32% had a separate room for cooking. It took on average 80 minutes for the caregiver to reach the usual health care center from her household. The families had been living in the same community for about 18 years

Table 2. Characteristics of the mother and her household, 620 children (1–59 months) deaths, Niger, 2007–2010

Characteristics	Frequency (No.)	Percent
Married or living with a man	599	96.6
Mean age when first married (years):	15.7 (range 12–30)	
12–15	355	59.3
16–19	208	34.7
20–30	36	6.0
Mother's mean age at time of child death (in years):	27.7 (range 11–50)	
11–19	73	11.8
20–24	149	24.0
25–29	152	24.5
30–34	103	16.7
35–50	109	17.6
Don't know	34	5.6
Mean years of maternal schooling:	0.6 (range 0–10)	
0	541	87.3
1–3	15	2.5
4–6	26	4.3
>6	22	3.5
Don't know	15	2.4
Father years of schooling (mean years of schooling):	1.0 (range 0–16)	
0–3	499	80.5
4–6	13	2.1
>6	40	6.5
Don't know	35	5.7
Household characteristics		
Main breadwinner:		
Father	596	96.2
Mother	11	1.8
Other	13	2.1
Main breadwinner is farmer/ agricultural worker	435	70.2
Mean years continuously living in community	17.7 (range 0–69)	
Household size (mean)	7.7 (range 2–25)	
Household has electricity	67	10.8
Use of piped water in–house water supply	208	33.6
Use of improved sanitation (improved pit for toilet)	43	7.0
Separate room for cooking	198	32.0
Household uses firewood for cooking	609	98.2
Floor of the house made of cement	63	10.2
Mean travel time to nearest health facility (min)	80.4 (range 0–1380)	
Social capital:		
In last 3 years, community worked together on at least 1 of the following: schools, health, jobs, credit, roads, public transport, water, sanitation, agriculture, justice, security, mosque/church	540	87.2
Mother was NOT able to turn to any persons or community groups or organizations for help during the pregnancy or child's fatal illness	379	61.2
Mother and her family have never been denied any of the following community	536	86.4

on average, yet 61% of the mothers reported they did not have anyone to help them during their child's illness.

Preventive home care

The exclusive breastfeeding and complementary feeding status among the 414 children whose fatal illnesses started between 0–23 months of age are presented in **Figure 1**. Overall, just 24.4% (n = 101) were appropriately fed. In more detail, about 23.9% (n = 38) of the 159 children whose fatal illnesses began at 0–5 months of age were exclusively breastfed. And only 28.9% (n = 54) of the 187 breastfed children 6–23 months old received the recommended complementary non–liquid feeds each day before the illness began.

Figure 2 shows the preventive home care received by the children along the continuum of care. About one in two (49.1%) of 405 children were likely to be exposed to smoke, ie, he/she was usually near the mother when she cooked inside the house. Sixty–three percent of the children always slept under an insecticide–treated bed net before their fatal illness began.

Preventive health facility care

Figure 2 further shows the percentage of the deceased children (12–59 months of age, n = 344) who received vaccinations against the six major preventable childhood diseases by one year of age. These findings were based on the vaccinations dates documented on the vaccinations cards, seen in about 13% of the cases, parental/respondent recall (87%), or a combination of the two. Overall less than a quarter (23.6%) of the deceased children 12–59 months were fully immunized against these diseases before their fatal illness began. The highest coverage was for BCG, DPT1 or PENTA1, and polio1, ranging from 69.2% to 70.9%. Sixty–two percent of children aged 12–59 months received measles vaccine. The deceased children were least likely to be fully immunized against DPT or PENTA by age one (just 34.8% had had all three doses).

Among the 12–59–month–old children, it took the group of fully immunized children on average 31 minutes less travel time to the nearest health care facility than the not–fully immunized children (63–minute vs 94–minute, $P = 0.050$).

Curative care

Figure 3 presents the breakdowns in the Pathway to Survival that contributed to the deaths of the children. Of the 620 completed interviews, 19 caretakers reported that they took some action at the time the fatal illness was noticed, yet the data on type of action was missing. Thus, the Pathway to Survival analysis included only the 601 children whose caretakers provided information on care–seeking.

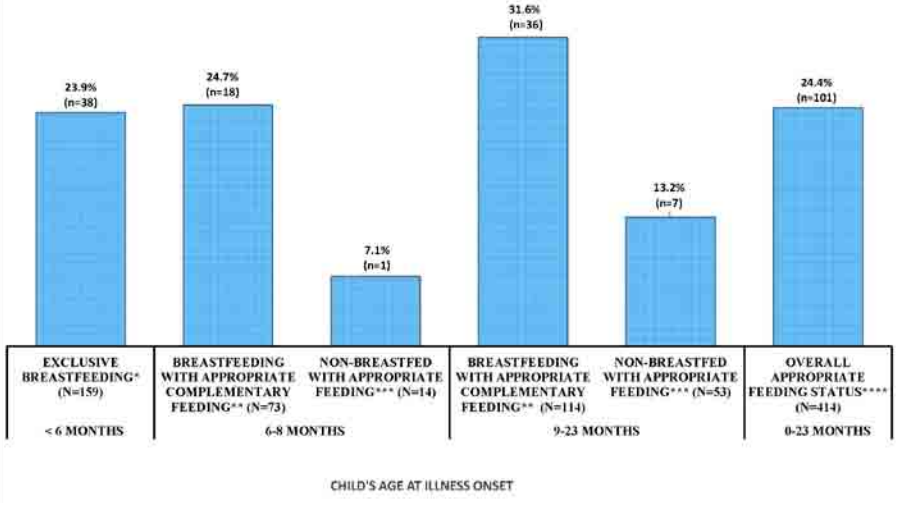


Figure 1. Breastfeeding and complementary or replacement feeding for children whose illness started at age 0-23 months. Legend: *Child's illness began before 6 months of age (0-5 months), he/she was being breastfed at the time of fatal illness and was not given anything but breast milk as food. **Breastfed children whose fatal illness started at 6-8 months old and 9-23 months old who received, respectively, at least two and three complementary non-liquid feedings each day. ***Child's fatal illness started at 6-23 months old and he/she received at least four replacement feeds each day (including milk and solid, semi-solid and soft foods). ****Children whose fatal illness started at 0-23 months and satisfied either of the conditions above.

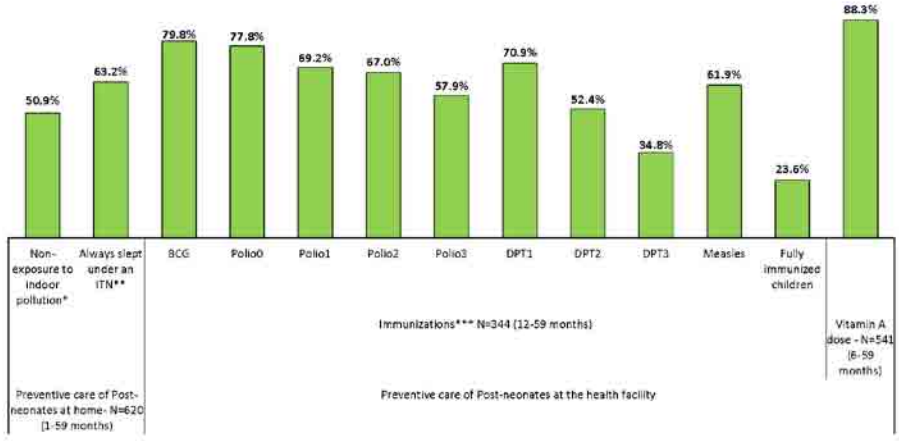


Figure 2. Coverage along the continuum of care for 1-59 month old child deaths in Niger from 2007-2010. Legend: *Proportion of children who were NOT usually nearby their mother when she cooked inside the home. **Insecticide-treated bed net. ***Information on immunizations was obtained either from the vaccination card or when there was no written record, from the respondent (mainly the mother). Polio0 is the Polio vaccination given at birth; fully Immunized children received BCG, measles, and three doses each of DPT and polio vaccine (excluding polio vaccine given at birth).

Nearly all (96.2%) of caretakers of the 601 children recognized that their child had a severe or possibly severe symptom when they first noticed that the child was ill. More than 88% of the children (n = 530) received care, or their caretakers sought or tried to seek care; 17 (2.8%) children "died immediately"; and no care was given or sought for the other 54 (9.0%) children. About 58% of those who "died immediately" and 26% for whom "no care was given or sought" were ranked as being severely ill at the time their caregiver first noticed the illness. For these groups the fatal illness occurred at 7 months of age, and lasted, respectively, 1 and 3 days.

The vast majority (87.2%, n = 462) of the group of 530 children who received or whose caretakers sought or tried to seek care, first sought care outside the home; 68 (12.8%) first received home care, and 36 of these 68 later sought or tried to seek outside care. In total then, 498 received, sought or tried to seek care outside the home, and the median length (or delay) from the illness onset until formal health care–seeking was 1 day.

When care was sought outside of the home, the vast majority (92.6%, n = 461) received only formal care, 16 received both informal and formal care, and 21 received informal care only. The delay in seeking formal care was longer for those who sought informal and formal care than for those who sought only formal care (Median 2 days vs 1 day, $P < 0.045$).

Out of the 477 children for whom formal care was sought, 102 (21.4%) did not reach the health facility because they died before setting out, died en–route or could not reach the health care provider. The remaining 375 (78.6%) children reached the first health care provider after about 67 minutes travel time. About 2 in 3 (n = 243, 64.8%) of these 375 children went to a health center, 5 went to a private clinic, 63 to a health post, 56 (14.7%) went to a non–governmental organization (NGO) or governmental hospital,

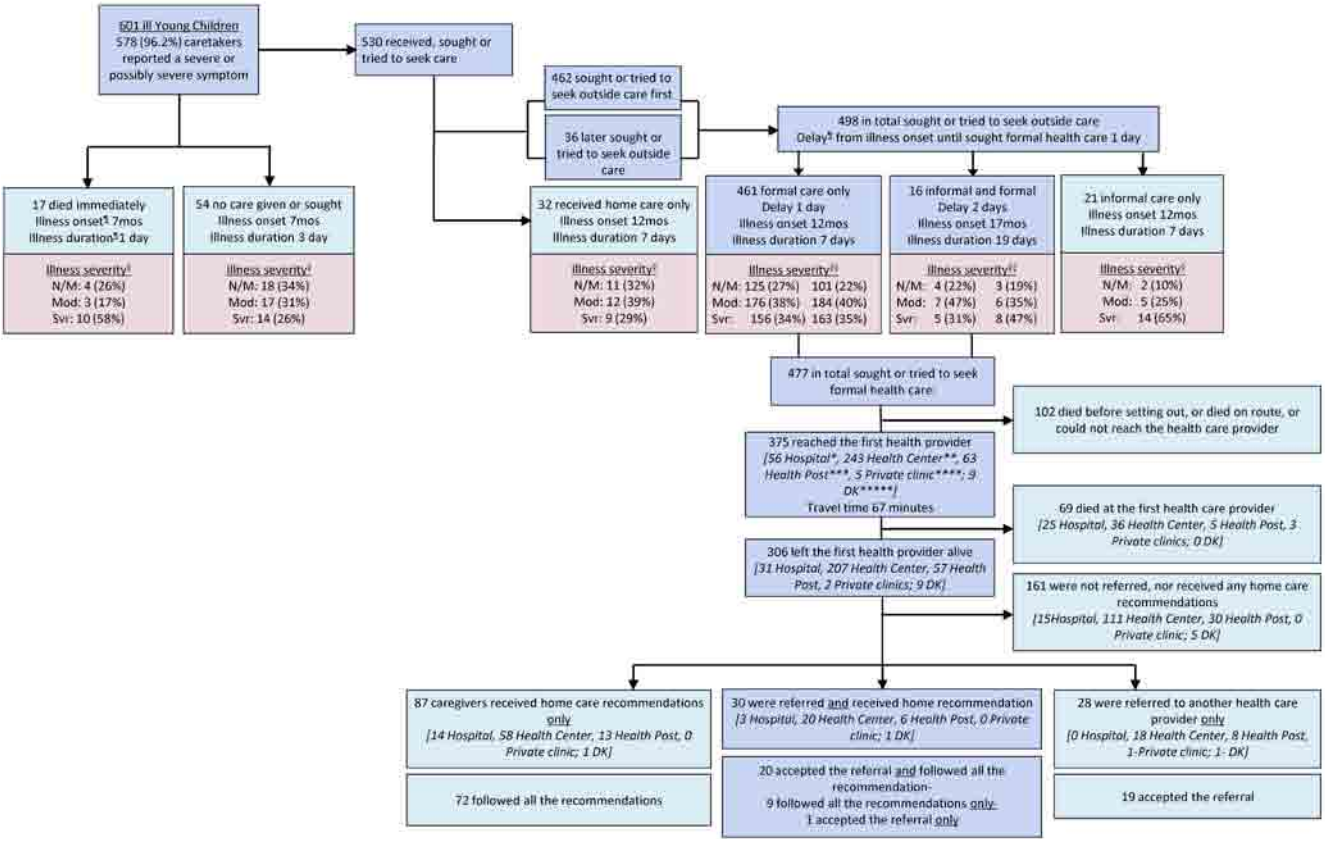


Figure 3. The "Pathway to Survival" for 601 deceased Young Children in Niger, from 2007-2010. Legend: §Illness severity at onset. §§Illness severity at onset and when caregiver decided to seek formal care. N/M=normal/mild, Mod=moderate, Svr=severe. *Hospital includes governmental or non-governmental hospital. **Health center includes governmental or non-governmental health center. ***Health post includes governmental or non-governmental health post. ****Private clinic formal or non-formal. *****DK: Don't know where but with a community health worker, a nurse or a midwife.

and 9 went to a health facility (staffed by a community health worker, a nurse or a midwife) for which the name or type could not be identified with the available data.

Sixty–nine (18.4%) of the 375 children that reached a first provider died at that provider, including the 25 out of 56 (44.6%) that reached a hospital. And 161 children out of the 306 (52.3%) that reached a health provider and left the provider alive were not referred nor given any home care recommendations. The remaining 145 were either only referred (n=28) to a second health care provider, or only received home care recommendations (n=87), or were referred and received home care recommendations (n=30). In summary, just 58 (19.0%) of the 306 that left the first provider alive were referred. However, when recommendations were received, or referrals provided, most of the caregivers (67%–100%) followed all the recommendations or accepted the referral and went to a second health care provider.

Figure 4 explores the care–seeking constraints for fatal child illnesses. In total, 113 caregivers reported that they had some concerns or problems in seeking care from a health care provider, for their child's fatal illness. Cost (35.4%), distance (34.5%) and lack of transport (30.1%) were the primary disincentives for care–seeking at a health provider. Another constraint prevailed among the 36 children who did not seek any formal care: 41.7% of caregivers thought that the child was not sick enough to warrant care.

DISCUSSION

The significance of this study lies in its goal to unveil the modifiable social, behavioral, and health system determinants of post–neonatal and child deaths in Niger between 2007 and 2010.

Findings from this study show that the majority of the deceased young children lived in households with precarious socioeconomic conditions, ie, lacking basic commodities such as electricity, sanitation, and clean water. There is a myriad of evidence to suggest that low standards of living adversely affect child morbidity and survival [12]. Improved safe water supply and community–wide sanitation are crucial interventions.

Exposure to household air pollution in the study setting was considerable: some 90% of the households lacked access to electricity and many relied on kerosene lamps and

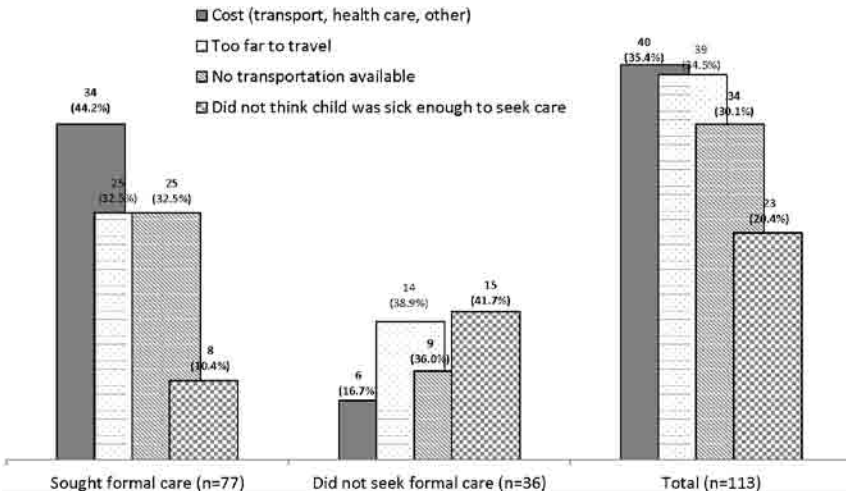


Figure 4. Main care-seeking constraints for child illness (N=113 caregivers).

other polluting alternatives for lighting; about one in three children were usually near their mother while she was cooking indoors. Almost all of the households used biomass fuel for cooking. All of these conditions constitute a large and growing cause of environmental health risks, particularly among children due to the immaturity of their respiratory system. There is consistent evidence that exposure to indoor air pollution can lead to acute lower respiratory infections, and can increase the incidence of pneumonia to twice that of children not exposed [13,14]. Improving access to modern energy that emits less pollution, both in the home and the community, can benefit the health of children in the study setting.

Moreover, this study showed that both parents, especially mothers of deceased children had little or no education. Overwhelming evidence demonstrates the benefits of providing universal education to mothers. The children of mothers with no education are 2.7 times more likely to die than children of mothers who have more than 12 years of education [15].

This study showed that the vast majority of mothers of deceased children entered into marriage before their 20th birthday. The legal age of marriage for girls in Niger is 15 years. A law has been proposed to change the age to 18 years for girls but is yet to be adopted. It is known that early marriage has a direct impact on the health and mortality of children [16]. Understandably, the international community supports measures that aim to prevent early pregnancy and its poor health outcomes by preventing marriage before 18 years of age, by increasing knowledge and understanding of the importance of pregnancy prevention, by increasing the use of contraception and by preventing coerced sex [17].

The low rate of exclusive breastfeeding found in this study is of concern. The benefits of exclusive breastfeeding for 6 months include the reduction of the risk of diarrhea [18] and respiratory illness [19]. For the mother, exclusive breastfeeding is found to delay the return of fertility [20], and accelerate recovery of pre–pregnancy weight [21].

The overall high rate of undernourishment of children whose fatal illness began between 0–23 month–old revealed by this study is problematic. More striking, this study revealed that the vast majority of the non–breastfed children did not receive the daily appropriate recommended feeds. A previous study estimated that a non–breastfed child is 10 times more likely to die from diarrhea in the first six months of life than an exclusively breastfed child [22].

Undernourishment among children leads to malnutrition that in turn affects their immune system. They are more likely to become sick with common illnesses such as malaria, diarrhea, or respiratory infections, and the risk of death is very high [23]. Overall, it is estimated that nutrition–related factors contribute to more than one–third of deaths in children under five years of age [24].

Food taboos and lack of knowledge have been identified as the underlying causes of malnutrition in Niger [25] and are a hindrance when it comes to improving children's health and nutritional status. Awareness–raising among families and communities must be one of the pillars to generate behavioral change in the population in general and among mothers of children under five in particular. More specifically, promotion of exclusive breastfeeding and complementary feeding practices has been proven to improve the nutrition and health of children and mothers [19–21,26].

The current study revealed that just 23.6% of deceased children 12–59 months were fully immunized before the fatal illness began, compared to 37.5% among alive children by 1 year of age as reported by the 2012 Nigerien demographic and health survey [27]. These low proportions

suggest that the need to increase the vaccination coverage within the country cannot be overemphasized.

The study also revealed that the average travel time to the nearest health care facility inversely affected the immunization status of the deceased children. This finding corroborates those of previous studies [28].

Caregivers' ability to recognize the illness and subsequently seek appropriate care for their children is pivotal to control diseases such as malaria, pneumonia, and diarrhea [29]. In the current study, caregivers were questioned only on illness signs and symptoms, but not on their recognition of specific illnesses as seen elsewhere [30]. And while almost all caregivers could report on a severe or possibly severe symptom or sign at the onset of their child's illness, nearly one–fifth did not seek or try to seek care outside of the home. This finding suggests that the mothers failed to recognize the meaning of these signs of common severe childhood illnesses as shown in previous studies [31,32]; therefore many did not seek appropriate care for their children. This echoes the need for the Nigerien government to effectively implement and support the integrated community case management (iCCM) strategy that enables CHWs to provide life–saving interventions closer to home to address common childhood illnesses that were formerly provided only by facility based nurses and doctors [33]. This strategy is expected to result in a greater reduction of childhood deaths, especially if properly delivered at no cost to the country's most disadvantaged communities.

In this study, nearly two–thirds of the children reached a first health care provider, offering the opportunity for effective treatment of the illness or referral if needed. Notably, 63 (16.8%) and 243 (64.8%) of 375 children that reached a first health provider were seen at a health post or health center, respectively. Health posts are able to manage uncomplicated childhood malaria, pneumonia and diarrhea, while providers at health centers treat all types of childhood illnesses and perform some laboratory tests and procedures such as lumbar puncture. Yet, these facilities have at their disposal only a minimal number of materials and drugs [34]. Therefore, providers at those facilities are entitled to refer severe or difficult cases to a higher level of care, usually to a health center, mini–hospital or a nearby district hospital [34].

The management of the children's illnesses at the first provider could have been questionable. Despite that almost all the children were reported by their caregivers to have exhibited signs of severe or possibly severe illness, more than half who reached and left the first health care provider alive were not referred nor received any home care recommendations, suggesting a poor quality of care. The reasons why this significant proportion of sick children was not referred nor received any home care recommendation are unclear and warrant further study. It has been previously reported that Niger has a very low hospitalization rate due to a low referral rate and major accessibility problems [35], and that health care workers often do not refer, and caretakers frequently do not follow referral recommendations [32]. Healthcare workers may also have difficulty in complying with guidelines for referral – especially in rural areas where caretakers may be faced with many communication and transportation barriers [35]. Notably, cost, distance and lack of transport were reported by caregivers as the most important constraints to seeking care from a health provider during the child's fatal illness, followed by not understanding the severity of the child's illness, which was actually the main reason that constrained the caregivers who did not go to a health care provider for their child's fatal illness.

This study has some limitations that were partly discussed in previous papers [5,36]. The long recall period of up to five years was mainly due to the retrospective design adopted for the VASA studies and the importance to include an adequate sample size of deaths. Consequently, this could have compromised the respondents' recall of events, thereby, the validity of the findings.

In addition, while we sought some information from well–child and medical records available in the home, these were rarely available. For example, the majority of immunization records came from parental recall, and infrequently from the vaccination cards or health records. Nevertheless, several problems have been reported for the information provided both by vaccination cards and parental recall [37,38]. Parental recall may be inaccurate if parents forget the type and the numbers of vaccinations received, provide socially desirable responses, are not the person who brought the child to the vaccination session, or received incorrect information on vaccine schedules from providers. Vaccination cards may be incomplete or inaccurate if providers fail to record the doses administered or caregivers forget to bring the card to a vaccination session or, plausibly, unavailable if they were buried along with the deceased child.

Finally, a group of survivors would have allowed the analysis to test whether or not there were significant differences between the coverage of interventions among cases (deceased children) and controls (alive children). However, the lack of a comparison group in SA studies is common and not so necessary since we are studying interventions that should be accessible to all children.

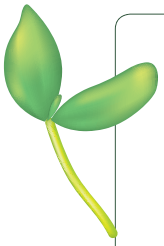
CONCLUSION

As governments and UN institutions work towards agreeing on a post–2015 framework, it is equally imperative that countries such as Niger commit explicitly to completing the job started by the MDGs, by adopting a target date to

end preventable child deaths. The current study is timely in that it provided information previously not available in Niger either at the local or national level.

Of note, the government of Niger and stakeholders are engaged in steps for the use of the overall VASA study results to support the revision of the child survival strategy for the country. For example, in order to improve access to better health services for children, the government approved the transformation of some health posts into health centers with an average of 50 health posts transformed per year by December 2014; and the immunization program is being strengthened with the introduction and widespread provision of the Pneumococcal vaccine in the country [39].

Yet, the country needs to do more by adopting and enforcing the law to prevent marriage of young girls before 18 years of age, and encouraging women's education and empowerment. Implementation of health programs that encourage breastfeeding and complementary feeding, illness recognition, prompt and appropriate care–seeking and improved referral rates for severe or possibly severe child illnesses will also go a long way towards curtailing child mortality in Niger.



Acknowledgments: Funding for the VASA study was provided by the Bill and Melinda Gates Foundation through a grant to the U.S. Fund for UNICEF for the Child Health Epidemiology Reference Group, and by the UNICEF country office of Niger. The funders had no role in the study design, data collection and analysis, interpretation of data, decision to publish, or preparation of the manuscript. We are extremely grateful to Dr Robert E. Black (REB), Director of the Institute for International Programs, Department of International Health, Johns Hopkins University, Bloomberg School of Public Health– for his fruitful comments on an earliest version of this paper.

Authorship declaration: HDK and AKK contributed in designing the study in Niger. OH, AM, KB, AKK, and HDK contributed to acquiring the data. AKK, AM, and AGY analyzed the SA data. AKK drafted the manuscript with significant input from HDK, KB, and AGY. All authors have given final approval of the version to be published.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author), and declare no competing interests.

REFERENCES

1 Institut National de la Statistique. Annuaire Statistique du Niger 2007–2011. Niamey, Niger. Available: http://www.stat-niger.org/statistique/file/Annuaires_Statistiques/AS2007-2011STRUCTUREPOPULATION.pdf. Accessed: 6 June, 2015.

2 UNDP. Human development index 2014. Available: http://hdr.undp.org/en/content/human–development–index–hdi–table. Accessed: 6 June, 2015.

3 Burki TK. Malaria and malnutrition: Niger's twin crises. Lancet. 2013;382:587-8. Medline:23961550 doi:10.1016/S0140-6736(13)61732-8

4 Amouzou A, Habi O, Bensaïd K. Reduction in child mortality in Niger: a Countdown to 2015 country case study. Lancet. 2012;380:1169-78. Medline:22999428 doi:10.1016/S0140-6736(12)61376-2

5 Kalter HD, Roubanatou A, Koffi A, Black RE. Direct estimates of national neonatal and child cause–specific mortality proportions in Niger by expert algorithm and physician–coded analysis of verbal autopsy interviews. J Glob Health. 2015;5:010415. Medline:25969734 doi:10.7189/jogh.05.010415

6 Murray CJ, Lopez AD, Black R, Ahuja R, Ali SM, Baqui A, et al. Population Health Metrics Research Consortium gold standard verbal autopsy validation study: design, implementation, and development of analysis datasets. Popul Health Metr. 2011;9:27. Medline:21816095 doi:10.1186/1478-7954-9-27

7 CSProX. Available: http://serpro.com/csp.html. Accessed: 6 June, 2015.

8 Kalter HD, Salgado R, Babille M, Koffi AK, Black RE. Social autopsy for maternal and child deaths: a comprehensive literature review to examine the concept and the development of the method. Popul Health Metr. 2011;9:45. Medline:21819605 doi:10.1186/1478-7954-9-45

9 Aguilar AM, Alvarado R, Cordero D, Kelly P, Zamora A, Salgado R. Mortality Survey in Bolivia: The Final Report. Investigating and identifying the causes of death for children under five. The Basic Support for Institutionalizing Child Survival (BASICS) Project. Arlington, VA; 1998. Available: http://pdf.usaid.gov/pdf_docs/PNACF082.pdf. Accessed: 14 December 2013.

10 Boschi–Pinto C, Young M, Black RE. The child health epidemiology reference group reviews of the effectiveness of interventions to reduce maternal, neonatal and child mortality. Int J Epidemiol. 2010;39 suppl 1:i3-6. Medline:20348123 doi:10.1093/ije/dyq018

REFERENCES

11 Nevo A. Using weights to adjust for sample selection. When auxiliary information is available. J Bus Econ Stat. 2003;21:43-52. doi:10.1198/073500102288618748

12 WHO. Combating waterborne disease at the household level. Geneva: WHO, 2007. Available: http://www.who.int/household_water/advocacy/combating_disease.pdf. Accessed: 6 June 2016.

13 Dherani M, Pope D, Mascarenhas M, Smith KR, Weber M, Bruce N. Indoor pollution from unprocessed solid fuel use and pneumonia risk in children aged five years: a systematic review and meta–analysis. Bull World Health Organ. 2008;86:390-398C. Medline:18545742 doi:10.2471/BLT.07.044529

14 Smith KR, McCracken JP, Weber MW, Hubbard A, Jenny A, Thompson LM, et al. Effect of reduction in household air pollution on childhood pneumonia in Guatemala (RESPIRE): a randomised controlled trial. Lancet. 2011;378:1717-26. Medline:22078686 doi:10.1016/S0140-6736(11)60921-5

15 Karlsen S, Say L, Souza JP, Hogue CJ, Calles DL, Gülmezoglu AM, et al. The relationship between maternal education and mortality among women giving birth in health care institutions: Analysis of the cross sectional WHO Global Survey on Maternal and Perinatal Health. BMC Public Health. 2011;11:606. Medline:21801399 doi:10.1186/1471-2458-11-606

16 United Nations Population Fund. Marrying too young. New York: UN, 2012. Available: http://www.unfpa.org/sites/default/files/pub–pdf/MarryingTooYoung.pdf. Accessed 3 May 2015.

17 WHO. WHO Guidelines on preventing early pregnancy and poor reproductive outcomes among adolescents in developing countries. Geneva: WHO, 2011. Available: http://www.who.int/maternal_child_adolescent/documents/preventing_early_pregnancy/en/. Accessed: 20 November 2015.

18 Kramer MS, Guo T, Platt RW, Sevkovskaya Z, Dzikovich I, Collet JP, et al. Infant growth and health outcomes associated with 3 compared with 6 months of exclusive breastfeeding. Am J Clin Nutr. 2003;78:291-5. Medline:12885711

19 Chantry CJ, Howard C, Auinger P. Full breastfeeding duration and associated decrease in respiratory tract infection in US children. Pediatrics. 2006;117:425-32. Medline:16452362 doi:10.1542/peds.2004-2283

20 World Health Organization Task Force on Methods for the Natural Regulation of Fertility. The World Health Organization Multinational Study of Breast–feeding and Lactational Amenorrhea. III. Pregnancy during breast–feeding. Fertil Steril. 1999;72:431-40. Medline:10519613 doi:10.1016/S0015-0282(99)00274-5

21 Dewey KG, Cohen RJ, Brown KH, River LL. Effects of exclusive breastfeeding for 4 versus 6 months on maternal nutritional status and infant motor development: results of two randomized trials in Honduras. J Nutr. 2001;131:262-7. Medline:11160544

22 Lamberti LM, Fischer Walker CL, Noiman A, Victora C, Black RE. Breastfeeding and the risk for diarrhea morbidity and mortality. BMC Public Health. 2011;11 Suppl 3:S15. Medline:21501432 doi:10.1186/1471-2458-11-S3-S15

23 Black RE, Allen LH, Bhutta ZA, Caulfield LE, de Oni SM, Ezzati M, et al; Maternal and Child Undernutrition Study Group. Maternal and child undernutrition: global and regional exposures and health consequences. Lancet. 2008;371:243-60. Medline:18207566 doi:10.1016/S0140-6736(07)61690-0

24 Liu L, Johnson HL, Cousens S, Perin J, Scott S, Lawn JE, et al; for the Child Health Epidemiology Reference Group of WHO and UNICEF. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since2000. Lancet. 2012;379:2151-61. Medline:22579125 doi:10.1016/S0140-6736(12)60560-1

25 International Federation of Red Cross and Red Crescent Societies. Case study – Tackling malnutrition in Niger, 2009. Available: https://www.ifrc.org/PageFiles/113731/Tackling%20malnutrition%20in%20Niger–EN.pdf. Accessed June 5, 2015.

26 Bryce J, Coitinho D, Darnton–Hill I, Pelletier D, Pinstrup–Andersen P; Maternal and Child Undernutrition Study Group. Maternal and child undernutrition: effective action at national level. Lancet. 2008;371:510-26. Medline:18206224 doi:10.1016/S0140-6736(07)61694-8

27 Institut National de la Statistique (INS) and ICF International. Enquęte Démographique et de Santé et Indicateurs Multiples du Niger 2012. Calverton, Maryland, USA: INS et ICF International, 2013.

28 Müller I, Smith T, Mellor S, Rare L, Genton B. The effect of distance from home on attendance at a small rural health centre in Papua New Guinea. Int J Epidemiol. 1998;27:878-84. Medline:9839747 doi:10.1093/ije/27.5.878

29 UNICEF. (2003) Communities and families – Improving knowledge of caregivers website. Available: http://www.unicef.org/communities/index_improvingknowledge.html. Accessed 2015 Jan 25.

30 Wilson SE, Ouedraogo CT, Prince L, Ouedraogo A, Hess SY, Rouamba N, et al. Caregiver recognition of childhood diarrhea, care seeking behaviors and home treatment practices in rural burkina faso: A cross–sectional survey. PLoS ONE. 2012;7:e33273. doi:10.1371/journal.pone.0033273. Medline:22428006

31 Geldsetzer P, Williams TC, Kirolos A, Mitchell S, Ratcliffe LA, Kohli-Lynch MK, et al. The recognition of and care seeking behaviour for childhood illness in developing countries: A systematic review. PLoS ONE. 2014;9:e93427. Medline:24718483 doi:10.1371/journal.pone.0093427

32 Cervantes K, Salgado R, Choi M, Kalter HD. Rapid assessment of referral care systems: a guide for program managers. Arlington, Virginia: Basic Support for Institutionalizing Child Survival Project (BASICS II) for the United States Agency for International Development, 2003.

33 UNICEF. Integrated Community Case Management: WHO/ UNICEF Joint Statement. New York: UNICEF, 2012.

REFERENCES

34 Bossyns P, Abache R, Abdoulaye MS, Miyé H, Depoorter A-M, Van Lerberghe W. Monitoring the referral system through benchmarking in rural Niger: an evaluation of the functional relation between health centres and the district hospital. BMC Health Serv Res. 2006;6:51. Medline:16608534 doi:10.1186/1472-6963-6-51

35 Simoes EA, Peterson S, Gamatie Y, Kisanga SF, Mukasa G, Nsungwa–Sabiiti J, et al. Management of severely ill children at first–level health facilities in sub–Saharan Africa when referral is difficult. Bull World Health Organ. 2003;81:522-31. Medline:12973645

36 Koffi AK, Libite PR, Moluh S, Wounang R, Kalter HD. Social autopsy study identifies determinants of neonatal mortality in Doume, Nguelemendouka and Abong–Mbang health districts, Eastern Region of Cameroon. J Glob Health. 2015;5:010413. Medline:26171142 doi:10.7189/jogh.05.010413

37 Suarez L, Simpson DM, Smith DR. Errors and correlates in parental recall of child immunizations: effects on vaccination coverage estimates. Pediatrics. 1997;99:E3-5. Medline:9113960 doi:10.1542/peds.99.5.e3

38 Valadez JJ, Weld LH. Maternal recall error of child vaccination status in a developing nation. Am J Public Health. 1992;82:120-2. Medline:1536315 doi:10.2105/AJPH.82.1.120

39 Bensaïd K, Yaroh AG, Kalter HD, Koffi AK, Amouzou A, Maina A, et al. Verbal/Social Autopsy – Niger 2012–13: A new tool for a better understanding of the neonatal and child mortality situation. J Glob Health. 2016;6:010601.

journal of global health

PAPERS

Verbal/social autopsy study helps explain the lack of decrease in neonatal mortality in Niger, 2007–2010

Henry D Kalter[1], Asma Gali Yaroh[2], Abdou Maina[3], Alain K Koffi[1], Khaled Bensaïd[4], Agbessi Amouzou[5], Robert E Black[6]

[1] Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

[2] Ministry of Health, Niamey, Niger

[3] Institute National des Statistics, Niamey, Niger

[4] UNICEF, Niger country office, Niamey, Niger (retired staff)

[5] UNICEF, New York City, NY, USA

[6] The Institute for International Programs, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Background This study was one of a set of verbal/social autopsy (VASA) investigations undertaken by the WHO/UNICEF–supported Child Health Epidemiology Reference Group to estimate the causes and determinants of neonatal and child deaths in high priority countries. The study objective was to help explain the lack of decrease in neonatal mortality in Niger from 2007 to 2010, a period during which child mortality was decreasing.

Methods VASA interviews were conducted of a random sample of 453 neonatal deaths identified by the 2010 Niger National Mortality Survey (NNMS). Causes of death were determined by expert algorithm analysis, and the prevalence of household, community and health system determinants were examined along the continuum of maternal and newborn care, the Pathway to Survival for newborn illnesses, and an extended pathway for maternal complications. The social autopsy findings were compared to available data for survivors from the same cohort collected by the NNMS and the 2012 Niger Demographic and Health Survey.

Findings Severe neonatal infection and birth asphyxia were the leading causes of early neonatal death in the community and facilities. Death in the community after delayed careseeking for severe infection predominated during the late neonatal period. The levels of nearly all demographic, antenatal and delivery care factors were in the direction of risk for the VASA study decedents. They more often resided rurally ($P<0.001$) and their mothers were less educated ($P=0.03$) and gave birth when younger ($P=0.03$) than survivors' mothers. Their mothers also were less likely to receive quality antenatal care ($P<0.001$), skilled attendance at birth ($P=0.03$) or to deliver in an institution ($P<0.001$). Nearly half suffered an obstetric complication, with more maternal infection (17.9% vs 0.2%), antepartum hemorrhage (12.5% vs 0.5%) and eclampsia/preeclampsia (9.5% vs 1.6%) than for all births in Niger. Their mothers also were unlikely to seek health care for their own complications (37% to 42%) as well as for the newborn's illness (30.6%).

Conclusions Niger should scale up its recently implemented package of high–impact interventions to additional integrated health facilities and expand the package to provide antenatal care and management of labor and delivery, with support to reach a higher level facility when required. Community interventions are needed to improve illness recognition and careseeking for severe neonatal infection.

Correspondence to:
Dr Henry Kalter
Department of International Health
Johns Hopkins Bloomberg School of Public Health
615 N. Wolfe St., Suite E8132
Baltimore, MD 21205
USA
hkalter1@jhu.edu

The 2010 Niger National Mortality Survey (NNMS) found that from 1998 to 2009 the mortality rate of children less than 5 years old decreased significantly by 43.4%, from 226 (95% confidence interval CI 207–246) to 128 (95% CI 117–140) deaths per 1000 live births, but mortality of neonates less than 28 days old declined insignificantly from 39 (95% CI 32–46) to 33 (95% CI 28–39) deaths per 1000 live births [1,2]. The reduction in child deaths was attributed to improvements in the nutritional status of children less than 2 years old and increased coverage of key child survival interventions, including insecticide–treated bed nets, vitamin A supplementation, treatment of diarrhea with oral rehydration salts and zinc, careseeking for childhood pneumonia and fever or cough, and vaccinations. The rapid uptake of interventions was achieved through government policy decisions to implement the Integrated Management of Childhood Illness (IMCI) approach, integrated community case management for children with fever or malaria, suspected pneumonia and diarrhea, and to provide free health care for all pregnant women and children including scaling up access to a minimum package of high–impact interventions at integrated health centers and health posts.

Interventions effective against neonatal mortality that were examined, including antenatal care, maternal tetanus toxoid, skilled birth attendance, early initiation of and exclusive breastfeeding, showed smaller increases in coverage to endpoint levels well below 50%, likely inadequate to decrease neonatal mortality [1]. In addition, an earlier study on the quality of maternal and newborn care found that few health workers present at birth had the knowledge, skills and access to basic equipment needed to effectively manage obstetric and newborn problems. Only 2.5% of Centres de Santé Intégrés (CSI), which are meant to have at least two nurses or midwives on duty at all times and which are the main health centers throughout the country intended to provide Basic Emergency Obstetric and Neonatal Care (BEmONC), had the full capacity for this service; and the national met need for EmONC stood at 2.3%, varying by region from 1.4% to 6.5% [3]. Health posts (Case de Santés), only about one–fourth of which have a nurse or midwife on staff and are not intended to provide EmONC, were not examined.

Neither of these studies, however, examined several other interventions critical to neonatal survival nor did they assess the causes of and events leading up to the deaths of the newborns along the continuum of antenatal and delivery care of the mother and immediate postnatal care of the newborn, maternal complications and the severe newborn illnesses these can lead to, mothers' perceptions and knowledge of how to respond to such critical events, their careseeking attempts for themselves and their newborns, and factors affecting these behaviors.

The fact that maternal complications occur at a fairly constant level, severe enough to kill the mother in about 1.0% to 1.4% of pregnancies and to kill the baby at a much higher rate, and that it cannot be reliably predicted which women will experience these complications, is the basis for the maternal mortality reduction strategy of universal access to skilled birth attendance and emergency obstetric care when needed [4–7]. This strategy is no less important to the survival and health of the neonate, as it has been shown that pregnancy and delivery complications are the most important risk factors for neonatal mortality [8–12], with care directed at the intrapartum period providing the greatest mortality reduction [13]. Integrated maternal–neonatal care packages and linkages of community with facility maternal and newborn care provide further reductions in stillbirths and neonatal deaths [14,15]. The addition of newborn–specific strategies, including fetal monitoring, access to Caesarean section for fetal distress, clean delivery and cord care, neonatal resuscitation, early initiation of and exclusive breastfeeding, timely and appropriate thermal care of the baby, kangaroo mother care for stabilized preterm infants, recognition of and early careseeking for newborn illness, access to quality health care, and urgent referral to neonatal intensive care when needed, are required to maximize newborn survival [13,16–20].

Examining such vital information on maternal and newborn care provided for babies that died is needed to help explain why the deaths occurred and how they might have been prevented. Collecting comparison data for newborns that suffered a severe but non–fatal illness during the same time period as the deaths would require the difficult task of identifying households where such an illness occurred; and the inability to appropriately match deaths with other cases on the basis of illness severity and the timing of clinical signs has led to a misleading situation where one could falsely conclude that treatment increased mortality risk [21]. Moreover, promoting neonatal health and preventing the death of sick newborns requires well–proven interventions for which the population levels established by already–completed surveys can provide reasonable comparisons for the surveyed factors.

Social autopsy (SA) is a method of inquiring about deaths that adds questions on household, community and health system determinants of mortality to complement a verbal autopsy (VA) interview on the illness signs and symptoms used to establish the biological cause of death [22]. We undertook to assess the biological causes and social determinants of recent neonatal deaths in Niger by conducting a verbal/social autopsy (VASA) study of neonatal deaths that occurred in 2007–2010 and were identified by the 2010 NNMS. Where possible, we compared the VASA findings for the deaths to the same factors for surviving children from the same cohort determined by recent population sur-

veys. In this way, we sought to further explain the reasons for the limited decrease in neonatal mortality in Niger from 2007 to 2010.

METHODS

The VASA was a descriptive study of the causes of death and the prevalence of key determinants of a national random sample of neonatal deaths derived from the 2010 NNMS's full birth history interview of women aged 15–49. Where possible, the levels of key determinants for the decedents were compared to the same factors for surviving children from the same cohort determined by the 2010 NNMS and the 2012 Niger Demographic and Health survey (NDHS); and maternal complications for surviving children were ascertained by the 2010–2011 WHO Multicountry Survey on Maternal and Newborn Health [12].

Data

The study sample has been fully described elsewhere [23]. In brief, the deaths included in the VASA study were identified by the lifetime birth history interview conducted of all women 15 to 49 years old who participated in the 2010 NNMS [24]. The VASA study considered only the 2380 under 5 years old (734 neonatal, 0 to 27 days old and 1646 child, 1 to 59 months old) deaths as far back from the survey period as four years. From these, in order to minimize the interview recall period, we started with the most recent death and moved backwards, taking the one most recent death in each household with at least one under 5 years old death until the desired sample sizes of 605 neonatal and 605 child deaths had been achieved.

The final VASA sample consisted of 1166 (96.9%) completed interviews of 1203 attempted, including 453 neonatal deaths, 620 child deaths and 93 stillbirths. Although the NNMS was designed to identify only live births and child deaths, some survey–classified (mainly) neonatal deaths were determined by the more detailed VASA interview to have been stillbirths, as defined by the caregiver's report that the child was born dead and never cried, breathed or moved. These discrepancies, as well as some movement between the neonatal and child categories, were checked during revisits to the households in question. The final VASA–determined birth status and age at death were taken as the correct data for this study. This paper examines the 453 neonatal deaths.

VASA interview

The VASA questionnaire, its translation, and the study's interview methods also have been fully described [23]. To sum up, the questionnaire blends the Population Health Metrics Research Consortium (PHMRC) VA questionnaire [25] with the Child Health Epidemiology Reference Group (CHERG) SA questionnaire [22]. The interviews were conducted in French and the two main languages of Niger, Haoussa and Zarma, using a CSProX [26] software application developed for the VASA study to assist interviewers to capture responses with minimal data entry errors in the field directly on netbook computers. Most of the fieldwork was conducted from March–April 2012. Revisits to some households extended the data collection until September 2012.

The interviewers were 12 female and eight male native speakers of Haoussa and/or Zarma, all secondary school graduates and 86% with some post–secondary education. They received 10 days of classroom training in all aspects of the VASA study and three days of field practice in conducting the interview. The seven teams, each with its interviewers and one supervisor, completed the data collection in 55 days. The respondent was the person most closely involved in caring for the child during the fatal illness, which typically is the mother. Secondary respondents were allowed, if necessary, to capture information on all phases of the illness, including the mother's pregnancy and delivery, during which she may herself have been ill and so less aware of the child's condition. In case of any disagreement the main respondent's answer was always taken as final.

Neonatal cause of death assignment

Verbal autopsy algorithms arranged in a hierarchy were used to assign the main cause of death for each neonate. The development of the algorithms and hierarchy and analysis to determine the causes of death also have been fully described [23]. Briefly, the algorithms were based on prior validation studies, additional verbal autopsy expert consultation, a literature review to identify illness signs and symptoms associated with particular neonatal illnesses, and the development of new algorithms for previously non–validated conditions. The hierarchy was developed to select the main, usually underlying, cause of death among all co–morbid conditions identified by the algorithms.

Social determinants of neonatal death

The Pathway to Survival [27] conceptual model was used to organize the collection and analysis of the social autopsy data on the health promotive, disease preventive and curative actions taken for children inside– and outside–the–home. An extended pathway for neonatal survival was developed to examine the mothers' antenatal and delivery care, pregnancy and delivery complications and careseeking for these. All the indicators examined along the pathways are of proven interventions against neonatal mortality contained in the Lives Saved (LiST) tool [28], judged by an evidence review or recommended by the WHO. The VASA study also assessed factors that might help explain why desirable actions were not taken, including socio–economic and demographic factors, recognition of illness se-

verity, who were the decision makers, and self–identified careseeking constraining factors.

Illness severity

Caregivers' reports of their child's symptoms indicating severe or possibly severe illness at the time of illness onset were used to evaluate the appropriateness of the first action taken in response to the illness. Symptoms' severity was rated according to their use in the Integrated Management of Childhood Illness (IMCI) approach [29], with any symptom signifying the need for urgent referral being ranked as "severe" and other symptoms yielding an IMCI disease classification requiring treatment ranked as "possibly severe." The severity of symptoms included in the verbal autopsy but not in the IMCI was rated by two of the study authors (HDK and AKK, both physicians). In addition to the rating of individual symptoms' severity, illness severity syndromes of symptoms recognized by mothers as indicating the need for health care [30-33] were formed by combining caregivers' reports of their child's feeding behavior, alertness and activity level at key points during the illness and used to evaluate the appropriateness of actions taken at those times. The method used to rank the severity of the syndromes has been previously described [34].

Maternal complications

As part of the extended pathway for neonatal survival, the presence of and careseeking for seven pregnancy and seven delivery complications were assessed. Because there is much overlap of the symptoms between major obstetrical conditions that can lead to over counting of complications, we examined the complications as defined by the symptom syndromes displayed in **Box 1**.

Box 1. Definitions of maternal complications

Pregnancy complications

Antepartum hemorrhage: Any vaginal bleeding before the onset of labor

Preeclampsia/eclampsia: (Puffy face and (blurred vision or severe headache or high blood pressure)) and/or (Convulsions and no fever and no history of convulsions)

Maternal infection: Fever and (severe abdominal pain or smelly vaginal discharge)

Maternal anemia: (Severe anemia or (pallor and shortness of breath)) and (too weak to get out of bed or fast or difficult breathing)

Gestational diabetes: Diabetes that started during pregnancy and before labor began

Premature rupture of the membranes: Water broke 6 hrs or more before labor began

Malaria: Convulsions and fever

Labor/delivery complications (start after labor onset)

Intrapartum hemorrhage: Excessive bleeding during labor or delivery

Preeclampsia/eclampsia: same as for pregnancy

Maternal infection: Fever and (severe abdominal pain or smelly vaginal discharge or foul smelling amniotic fluid)

Maternal anemia: same as for pregnancy

Preterm delivery: Less than 9 months

Prolonged labor: Labor for 12 or more hours

Malaria: same as for pregnancy

Comparison data

The 2010 NNMS was examined for data on factors comparable to those included in the VASA study for surviving children who were their mother's most recent birth in 2007–2010 (**Table 1** and **Table 2**). The NNMS collected data on maternal care variables only for births in the 12 months prior to the survey. Some variables not assessed by the 2010 NNMS were available from the 2012 NDHS [35]. These data also were examined for surviving children who were their mother's last birth in 2007–2010 (**Table 2**). **Table 3** displays baseline levels of maternal complications prior to surviving births established by the 2010–2011 WHO Multicountry Survey on Maternal and Newborn Health [12], including 10871 births in Niger.

Statistical analysis

This study was mainly descriptive. Percentages, means and medians are reported for demographic factors, causes and social determinants of neonatal deaths, maternal complications associated with the deaths, and the available comparison data. The χ^2–test was used to assess differences between proportions for the VASA findings and comparison data. The VASA analysis was adjusted for sampling weights, taking into account the cluster design nature of the 2010 NMMS that identified the deaths. The 2010 NMMS and 2012 NDHS analyses that provided the comparison data also were adjusted for cluster sampling using the sampling weights for those studies.

Ethical approval

The study was approved by the National Consultative Ethics Committee of the Niger Ministry of Health and the Institutional Review Board of the Johns Hopkins Bloomberg School of Public Health. Informed consent was given by all study participants prior to their being interviewed.

RESULTS

Ninety–seven percent of the respondents for the 453 neonatal deaths were the mother of the deceased child. Nearly two–thirds (65.8%) of the deaths occurred during the first

Table 1. Comparison of demographic characteristics of 453 neonatal deaths with those of the general population, Niger, 2007–2010

	Neonatal deaths			Comparison data*				
Characteristic	N	%, mean or median	(Q1, Q3)	N	%, mean or median	(Q1,Q3)	χ^2	*P*–value
Sex:								
Male	263	58.1		8360	50.8		9.3	0.002
Female	190	41.9		8082	49.2			
Birth order:								
1	114	25.3		2317	14.1		44.3	<0.001
2–3	92	20.4		4878	29.7			
4+	245	54.3		9212	56.2			
Mother's age at first marriage:†								
Median age (years)	425	15.0	(15, 17)	8938	15.0	(14, 17)		
Mother's age at birth of index child:								
<20	101	22.8		2559	15.7		16.3	<0.001
20–24	121	27.2		4112	25.2			
25–29	97	22.0		4302	26.4			
30+	124	28.0		5338	32.7			
Mean age (years)	443	25.6	(20, 31)	16312	26.9	(21, 31)		
Mother's age at first birth:								
<15	41	9.2		1190	7.4			
15–19	273	61.8		9400	58.8			
20+	128	29.0		5409	33.8		4.5	0.034
Mean age (years)	442	18.4	(16, 20)	15999	18.7	(16, 20)		
Mother's education:								
None	386	86.6		13521	82.7		4.7	0.030
Primary	47	10.6		1890	11.6			
Secondary+	13	2.9		948	5.8			
Median years	446	0.0	(0, 0)	16359	1.0	(1, 2)		
Father's education:								
None	369	84.7		–	–			
Primary	46	10.6		–	–			
Secondary+	20	4.7		–	–			
Median years	434	0.0	(0, 0)	–	–			
Residence:								
Urban	44	9.7		2994	18.2			
Rural	409	90.3		13450	81.8		21.5	<0.001
Travel time (min) to usual health facility:								
<30	167	39.2		–	–			
30–59	54	12.6		–	–			
60+	206	48.3		–	–			
Median minutes	427	40.0	(10, 120)	–	–			

*Unless otherwise noted, comparison data are for surviving children who were their mother's most recent birth in 2007–2010 and for their families from the 2010 Niger National Mortality Survey.

†VASA and comparison data are for ever–married women 20–49 years old at the time of the survey. The comparison data are from the 2012 Niger Demographic and Health Survey.

six days of life. More than half (62.9%) the newborns died from an infectious disease, nearly 20% succumbed to intrapartum-related complications (including 17.9% with birth asphyxia alone, 1.1% with signs of birth injury alone, and 0.9% with both birth asphyxia and injury), and 2.6% each to congenital malformations and complications of preterm delivery (**Figure 1**).

Demographic factors

As shown in **Table 1**, the deceased neonates were significantly different from surviving children born in the same population during the same years for several demographic characteristics. There was a greater male predominance among the deaths than among surviving children (58.1% vs 50.8%, $\chi^2 = 9.3$, $P = 0.002$). More of the deceased neonates than the survivors were their mother's first born child (25.3% vs 14.1%, $\chi^2 = 44.3$, $P < 0.001$). Their mothers also gave birth at a younger age—more of the decedents' mothers were less than 20 years old, both at the time of the index child's birth (22.8% vs 15.7%, $\chi^2 = 16.3$, $P < 0.001$) and when they had their first birth (71.0% vs 66.2%, $\chi^2 = 4.5$, $P = 0.034$), than mothers of surviving children. In addition,

Table 2. Comparison of antenatal and delivery care indicators for 453 neonatal deaths with those of the general population, Niger, 2007–2010

	Neonatal deaths		Comparison data*		χ^2	*P*–value
Characteristics	N	%	N	%		
Antenatal care:†						
At least 1 visit	336	74.1	3269	83.7	25.7	<0.001
4+ visits	140	31.6	1275	32.6	0.2	0.671
Antenatal care content:†						
Blood pressure	249	75.1	2432	74.4	0.1	0.81
Urine test	86	25.7	1252	38.3	20.8	<0.001
Blood test	129	38.7	1560	47.7	9.8	0.002
Danger sign counseling	117	35.1	1882	57.6	61.8	<0.001
Quality antenatal care (blood pressure, urine& blood test, counseling)	40	12.0	801	24.6	26.7	<0.001
Tetanus vaccination	283	62.4	4209	65.7	2.0	0.162
Antimalarial†	233	51.5	2760	70.6	69.3	<0.001
Delivery place:						
Hospital	26	5.9	113	1.8	36.1	<0.001
Other formal provider	101	22.3	2294	35.8	33.9	<0.001
Institutional delivery	127	28.1	2407	37.6	16.6	<0.001
En route to provider	12	2.6	–	–		
Home	313	69.1	3940	61.5	10.4	0.001
Other	1	0.2	57	0.9		
Birth attendant:						
Skilled	129	28.5	2148	33.5	4.8	0.028
Traditional birth attendant	148	32.7	3096	48.3		
Mother herself	92	20.4	–	–		
Other	84	18.5	1160	18.1		
Delivery mode:						
C–Section†	11	2.4	73	1.9	0.6	0.443

*Unless otherwise noted, comparison data are for mothers of surviving children who were their mother's most recent birth in the 12 months prior to the 2010 Niger National Mortality Survey.

†Comparison data are for mothers of surviving children who were their mother's most recent birth in 2007–2010 from the 2012 Niger Demographic and Health Survey.

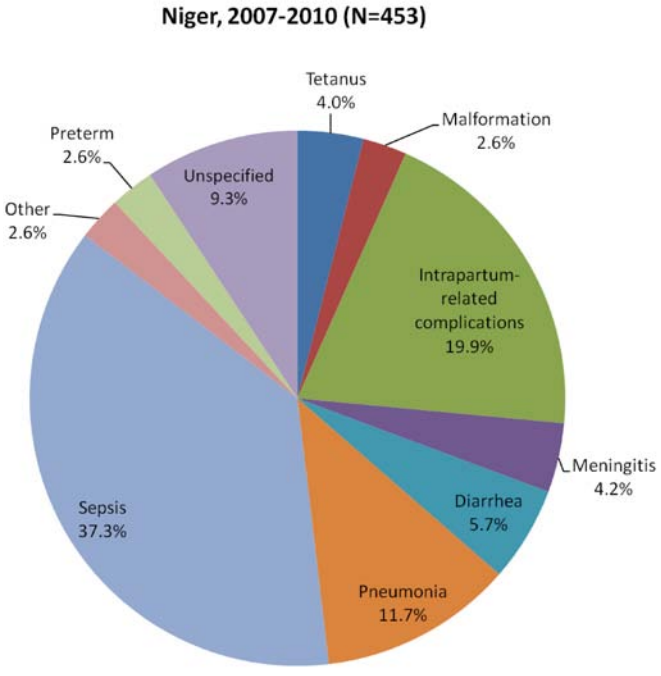


Figure 1. Expert algorithm, hierarchical verbal autopsy causes of death for 453 neonatal deaths, Niger, 2007–2010.

more of the deceased children's mothers had no formal education (86.6% vs 82.7%, $\chi^2=4.7$, P=0.030). The households of the children also differed in that more families of the deceased than the survivors resided in a rural area (90.3% vs 81.8%, $\chi^2=21.5$, $P<0.001$). Comparison data are lacking for travel time to the usual health facility used. Nevertheless, the data for the decedents reveals a long travel time (40 minutes). The mothers both of decedents and survivors first married when they were very young.

Antenatal and delivery care

Fewer decedents' than survivors' mothers accessed any antenatal care (ANC) (74.1% vs 83.7%, $\chi^2=25.7$, $P<0.001$), although only about 32% of both groups made at least the recommended four visits (**Table 2**). However, twice as many mothers of the survivors (24.6% vs 12.0%, $\chi^2=26.7$, $P<0.001$) who made at least one ANC visit received quality care consisting of all of four key ANC interventions. The biggest gap in the individual ANC components was in counseling on the danger signs of pregnancy requiring urgent careseeking. Survivors' mothers also were more likely to take a prophylactic anti–malarial during their pregnancy

(70.6% vs 51.5%, $\chi^2=69.3$, $P<0.001$), to have an institutional delivery (37.6% vs 28.1%, $\chi^2=16.6$, $P<0.001$) and to be cared for by a skilled birth attendant (33.5% vs 28.5%, $\chi^2=4.8$, $P=0.028$). Only about 2% of both groups were delivered by Caesarean section.

Maternal complications and careseeking

Table 3 shows that more than half (52.6%) of the 298 mothers with an early neonatal death had a serious pregnancy or labor and delivery complication, with the level for individual pregnancy complications ranging up to 17.9% for maternal infection and for individual labor and delivery complications up to 16.1% for intrapartum hemorrhage. While there were no comparison data for these findings in the NNMS or NDHS, a district hospital study in Kenya [8], community–based studies in Bangladesh [9] and Palestine [10], and a multi–country hospital–based survey [12] found comparable, some higher and some lower, levels of these same complications in women with a perinatal or early neonatal death and, by country, uniformly lower levels in women with a surviving neonate. The multi–country survey included Niger, which had much lower levels for three maternal complications among all births than found for the deaths in our study (**Table 3**).

Because early onset neonatal infection is common in newborns whose mothers have maternal infection or colonization [36], we also explored the relationship of maternal infection to early onset (at less than 2 complete days of life) neonatal infection (sepsis, meningitis or pneumonia) as the primary cause of neonatal death. We demonstrated a significant positive association between these maternal and neonatal conditions, both when comparing early onset to late onset neonatal infection (28/95 [30.0%] vs 17/146 [11.3%], $\chi^2=13.2$, $P<0.001$) as well as to all other neonatal deaths from any cause (28/95 [30.0%] vs 65/358 [18.1%], $\chi^2=6.5$, $P=0.011$) [23].

Figure 2 shows the maternal complications and careseeking for these for all 453 women with a neonatal death. Fewer than half the women with a pregnancy complication (65/155, 42.0%) or labor and delivery complication that began at home (45/122, 37%) sought any formal health care for their complications, and they were no more likely to deliver at a health facility than women without a complication (any pregnancy complication: 32.7% vs 25.8% without,

Table 3. Comparison of maternal complications among early neonatal deaths in Niger, 2007–2010, with those of all births in Niger and perinatal and early neonatal deaths and neonatal survivors in other countries

			Multi–country survey* [12]							
	Niger VASA	Niger	All countries		Kenya [8]		Bangladesh [9]		Palestine [10]	
	ENM N = 298 %	All births N = 10 871 %	ENM N = 2528 %	Surv N = 298 912 %	PNM N = 108 %	Surv N = 802 %	PNM N = 86 %	Surv N = 1498 %	PNM N = 80 %	Surv N = 808 %
Pregnancy complications:										
Maternal infection†	17.9	0.2	2.6	0.5	–	–	–	–	–	–
Antepartum hemorrhage	12.5	0.5	5.7	0.6	8.3	0.4	12.8	2.7	12.5	0.7
Eclampsia/preeclampsia	9.5	1.6	9.7	2.2	–	–	19.8	8.6	26.3	18.4
Premature rupture of membranes	6.8	–	–	–	12.0	1.6	–	–	–	–
Malaria	6.7	–	0.5	0.1	–	–	–	–	–	–
Anemia	3.1	–	4.8	1.2	–	–	–	–	–	–
Diabetes	0.0	–	–	–	–	–	–	–	–	–
Any pregnancy complication	36.4	–	–	–	–	–	–	–	37.5	7.4
Labor and delivery complications:										
Intrapartum hemorrhage	16.1	–	1.4	0.2	–	–	–	–	–	–
Prolonged labor	5.2	–	–	–	17.6	7.2	–	–	7.5	4.3
Preterm labor	8.2	–	52.2	6.0	12.0	1.6	45.3	21.7	57.5	9.3
Maternal infection	4.5	–	–	–	–	–	–	–	–	–
Eclampsia/preeclampsia	0.8	–	–	–	–	–	–	–	–	–
Malaria	0.7	–	–	–	–	–	–	–	–	–
Anemia	0.3	–	–	–	–	–	–	–	–	–
Any labor/delivery complication	28.0	–	–	–	60.2	14.0	–	–	45.0	24.3
Any maternal complication	52.6	–	–	–	–	–	–	–	–	–

ENM – early neonatal mortality, PNM – perinatal mortality, Surv– neonatal survivors

*The referenced study did not distinguish between antepartum and intrapartum complications. The distinctions made here are based on information provided in the paper; for example, placenta previa was categorized as antepartum hemorrhage, and ruptured uterus as intrapartum hemorrhage.

†For Niger VASA ENM, maternal infection = sepsis; for all multi–country survey results, maternal infection = any one or more of pyelonephritis, influenza–like illness, other systemic infection/sepsis.

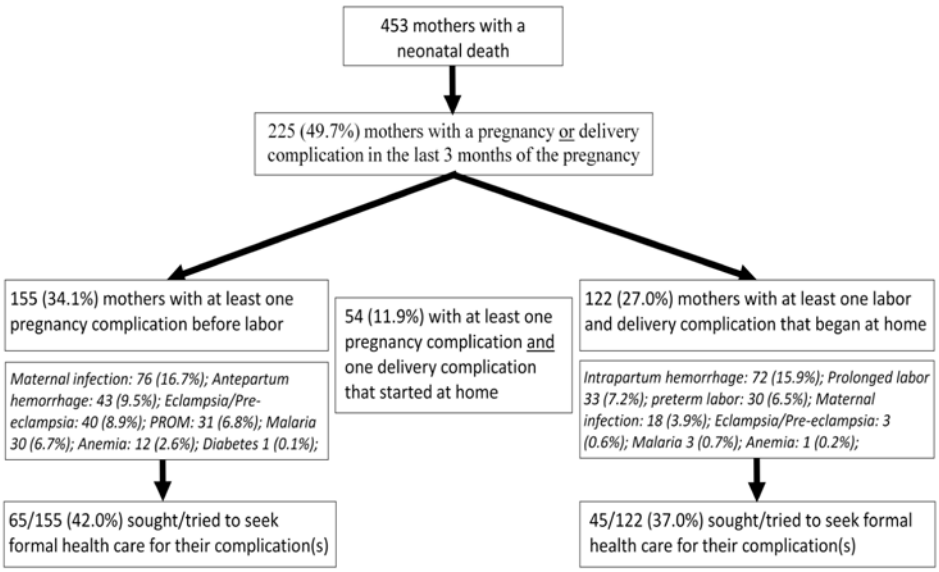


Figure 2. Maternal complications and careseeking during the pregnancy and delivery for 453 neonatal deaths, Niger, 2007–2010.

$\chi^2 = 2.45$, $P = 0.118$; any delivery complication: 25.1% vs 29.3% without, $\chi^2 = 0.79$, $P = 0.373$). Almost all this care was sought at the primary care level. Of the 45 women who sought or tried to seek formal care for a labor and delivery complication, 21 (46%) went to a CSI, 5 (12%) to a Case de Santé, 11 (24%) to a primary care facility of undetermined type, and 3 (7%) said that they saw a nurse or midwife in the community. Only four (10%) of the women first went and four (10%) more later went to a hospital. Eight (17%) delivered in a hospital, 20 (44%) in a CSI, Case de Santé or undetermined primary care facility, 3 (7%) on route to a facility and 14 (32%) at home. Of the labor and delivery complications that occurred in more than 3% of the women, formal health care was sought most often for maternal sepsis (9/18 women, 51.7%), followed by preterm labor (14/30, 48.6%), prolonged labor (14/33, 42.9%) and intrapartum hemorrhage (26/72, 35.8%). However, as for all delivery complications, women with these complications were no more likely to deliver at a health facility than women without a delivery complication.

Normal newborn care

Almost a third (144, 31.7%) of the babies that died were bathed within one hour after birth, and 351 (79.1%) were bathed before 24 hours after birth, which is the recommended lower time limit for first bathing [37]. An appropriate measure was taken to keep 406 (90.0%) of 451 newborns warm after birth, but only 42 (9.9%) of 423 were breastfed in the first hour after birth, compared to 2350 (42.9%) of 5478 surviving children born in the prior 24 months identified by the 2010 NNMS ($\chi^2 = 177.1$, $P < 0.001$). In all, only 1 of 408 deceased neonates received quality postnatal care in the first day of life (ie, sterile blade used to cut the cord, baby not bathed in the first 24 hours after birth, baby dried and wiped or wrapped in a blanket or given skin to skin contact or placed in an incubator after birth, and baby breastfed within 1 hour after birth).

Failures in the pathway to survival

Figure 3 illustrates the careseeking process from home for the 385 deceased neonates who either were born at home or delivered at a health facility and left alive. Although nearly all (95.8%) caregivers reported that the first symptom of their child's illness was either a severe or possibly severe symptom, 232 (60.3%) neonates, 64 of whom were said to have died "immediately," received no care for their fatal illness.

The mean age at illness onset for the 232 newborns who received no care was 3.5 days and their mean illness duration was 1.7 days, compared to illness onset at age 6.8 days and duration of 3.7 days for the 153 (39.7%) neonates whose caregiver's first action was to provide home care or seek care outside the home. Caregivers' reports of their child's illness severity as rated according to their feeding behavior, alertness and activity level confirmed that more of the 232 newborns who received no care were severely ill at the start of their illness than of the 153 newborns who received some care (62.4%, vs 44.0%; $\chi^2 = 11.6$, $P < 0.001$), corresponding with the earlier onset and more rapid progression of their illnesses.

The causes of death of the two groups (**Figure 4**) also reflected their age and illness course, with more deaths due

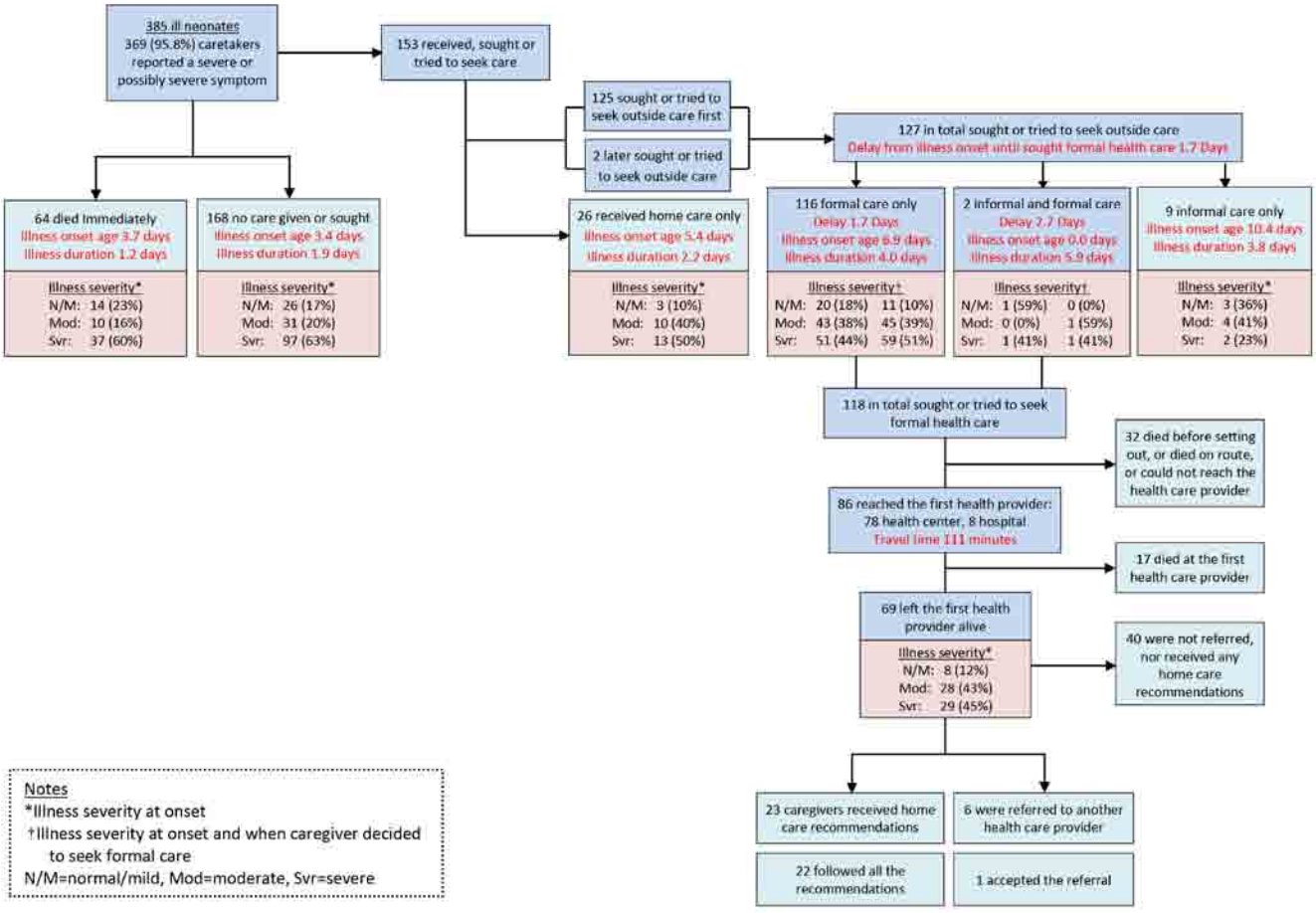


Figure 3. The "Pathway to Survival" for 385 neonatal deaths born at home or left the delivery facility alive, Niger 2007–2010.

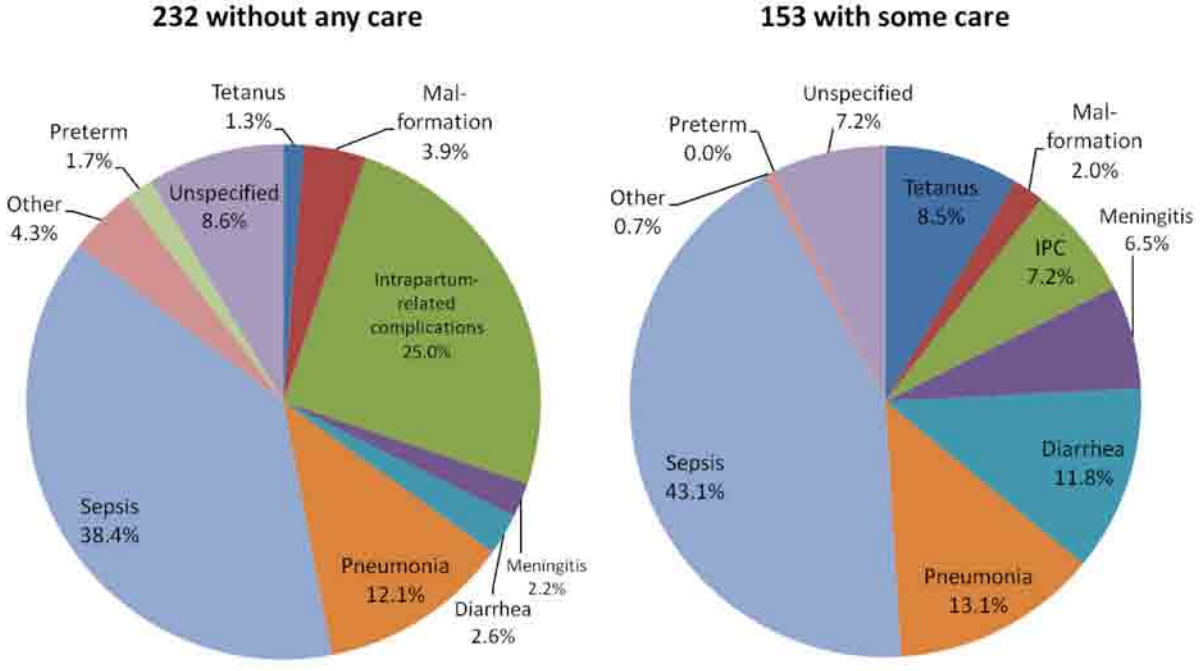


Figure 4. Expert algorithm, hierarchical verbal autopsy causes of death for 385 neonates born at home or left the delivery facility alive, with and without any care given or sought from home, Niger, 2007–2010. IPC – intrapartum-related complications.

to birth asphyxia and preterm delivery in the younger group with faster illness progression and no care provided or sought than among the older neonates who received care (26.7% vs 7.2%), and more infectious deaths in the older group with longer illnesses and some care than among the younger neonates without care (83.0% vs 56.6%). Of note is that 8.5% of the 153 late neonatal deaths that received some care were due to tetanus.

For comparison, **Figure 5** shows the causes of death of 59 neonates not included in **Figure 3** because they died in their delivery facility without ever leaving. The cause distribution was more similar to that of the 232 neonates who died at home without any care given or sought, though even more skewed toward birth complications and preterm delivery over infectious causes (45.8% vs 39.0%), while their mean age at illness onset (1.0 days) was younger and their illness duration (2.3 days) was somewhat longer. Most remarkable was that eight of the 12 deaths due to preterm delivery in all 453 neonates occurred among the 59 babies that were born and died in a health facility without leaving. Of the other four preterm deaths (all among the 232 newborns that died at home without seeking care), three were delivered in a health facility and one at home.

Figure 3 also shows that formal health careseeking was attempted for only 118 of the 153 neonates who received any care. Among these 153 newborns, there was no difference in formal careseeking for males and females (78.2% vs 76.6%, $\chi^2=0.5$, $P=0.816$). Formal careseeking was delayed for 1.7 to 2.7 days, nearly half way into the illness course, by which time half the children were severely ill. Though the children for whom formal care was sought were less severely ill at the onset of their illness and their illnesses progressed more slowly than those of the children who received no care, the delay in careseeking was associated with 32 (27.1%) of the 118 neonates dying before reaching the formal provider. However, there was no significant difference in the proportions of newborns whose illness severity increased between those who reached and did not reach the provider (20.2% vs 9.6%, $\chi^2=1.76$, $P=0.416$), and equal proportions were severely ill at the time the decision was taken to seek formal care (48.4% vs 55.8%, $\chi^2=0.51$, $P=0.47$).

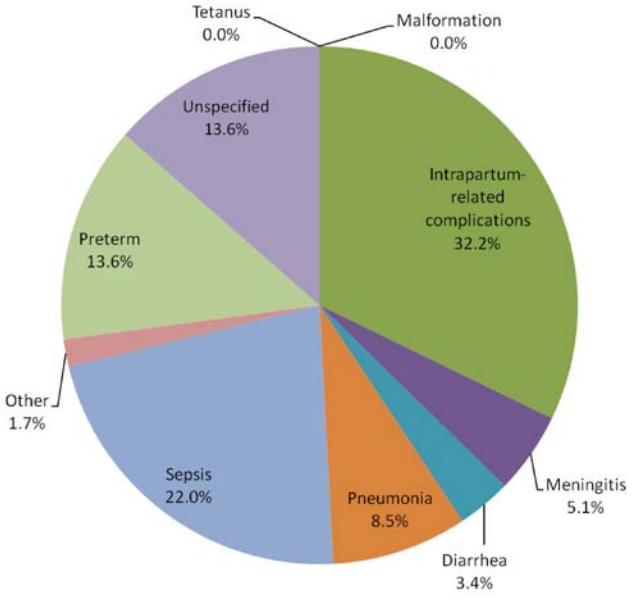


Figure 5. Expert algorithm, hierarchical verbal autopsy causes of death for 59 neonates born and died in the delivery facility, Niger, 2007-2010.

Almost all formal care was sought at lower level health facilities; only eight (6.8%) of the 118 newborns initially went to a hospital, only six (8.7%) of 69 who left a first level facility alive were referred, and only one of the six accepted the referral. The most common actions taken by the first health facilities reached by the 86 neonates included: 34 (39.9%) were given an intramuscular medication, 20 (23.7%) received an oral anti–malarial, 18 (20.7%) received another oral medicine, and 13 (15.6%) were given an oral antibiotic. On average, each neonate received 1.7 treatments; the health provider did "nothing" for only 6 (7.0%) neonates. One–third of the 69 caregivers of neonates discharged alive from a first level facility received recommendations for home care, and all but one of those were able to follow all the recommendations. **Figure 3** also shows that 29 (45%) of the 69 children that left the first provider alive were still rated by their caregiver as being severely ill at discharge.

Few comparison data for surviving children in Niger are available for these careseeking findings. The 2010 NNMS asked about formal careseeking for children with a fever or cough in the two weeks prior to the survey. For fever, 71.6% of 6 neonates, 56.5% of 20 1–month olds, and 62.8% of 71 2–month olds sought formal care, for a total of 62.8% of 97 children under 3 months of age; while for cough, all 3 neonates, 8.5% of 8 1–month olds, and 52.7% of 49 2–month olds sought formal care, for a total of 48.5% of 60 children under 3 months of age. These figures compare with the 118 (30.6%) and 86 (22.3%) of 385 neonates with a fatal illness who, respectively, sought formal health care and reached the first formal provider. Limiting the comparison to the 151 deceased neonates who survived at least one week, which is closer to the age distribution of the surviving children, 73 (48.3%) and 54 (35.7%), respectively, sought formal health care and reached the first provider.

Constraining factors for maternal and newborn careseeking

Concerns of caregivers that contributed to delays in careseeking were similar for pregnancy complications, deliver-

ing at a health facility and newborn illnesses, with some notable exceptions (**Table 4**). Most women, even those who did not seek formal health care, said they had no concerns that kept them from seeking care, although more of those who did not seek care reported one or more concerns. Of 90 women who did not seek formal care for their pregnancy complication(s), 26 (28.7%) reported, on average, 2.0 constraints each. The most common concerns were the cost for transportation or health care (13), distance to a facility (12) and the lack of transportation (9). Cost and distance also were the main constraints for those who sought formal care, but only 7 (10.8%) of the 65 women who sought care reported that they had any careseeking constraints. Similarly, distance and transportation were the main constraints women had for delivering at a health facility and seeking care for their newborn's illness, with cost being a lesser issue for both. Unlike for the other two situations, underestimating illness severity was an often stated constraint to careseeking for newborn illnesses. However, an incongruity to be considered is that 7 of the 17 caregivers who reported that their baby was not sick enough to need health care also ranked the child as being severely ill.

DISCUSSION

We undertook a verbal/social autopsy study of recent neonatal deaths in Niger to examine critical factors that might help explain the non–decrease in neonatal mortality in light of the significant decline in child mortality from 1998 to 2009. The study deaths were identified by the 2010 NNMS, the same survey used to establish the mortality trends [2], and that provided most of the data on surviving children and their families that we compared to the deaths. We determined the cause distribution of the deaths and related maternal complications, as these can highlight needed interventions, and focused on social, behavioral and health system determinants that influence the strategies required to effectively deliver maternal and child survival interventions [13].

Demographic factors, normal maternal and newborn care

The examination of demographic factors revealed several significant differences between deaths and survivors of the same cohort, with all factors in the direction of risk for the deaths. Some of these are potentially modifiable in the long term, while knowledge of some others might help in targeting interventions.

The excess in rural residence of deceased neonates, combined with their long travel time to the usual health facility, suggests that some remote communities remain at risk due to limited access to primary care despite recent efforts in Niger that have brought a package of high–impact interventions at integrated health centers and posts to within 5 km of 80% of the population [38].

The predominance of deaths of male neonates agrees with a well–established pattern of excess male mortality that suggests this is due to an unmodifiable, biological effect [39,40]. On the other hand, the excess in neonatal deaths of firstborns, among women under age 20 and women with

Table 4. Constraints for formal health careseeking for three situations contributing to neonatal deaths, 2007–2010, Niger

	Pregnancy complications				Health facility delivery				Newborn illness			
	65 sought formal care		90 did not seek formal care		127 delivered at a facility		326 did not deliver at a facility		118 sought formal care		202 did not seek formal care	
	N	%	N	%	N	%	N	%	N	%	N	%
Constraints:												
Did not think she/the baby was sick enough to need health care	0	0	2	1.9	0	0.2	4	1.1	0	0	17	8.3
No one available to go with her	0	0	1	1.3	0	0.2	8	2.4	0	0	1	0.3
Too much time from her regular duties	1	1.7	1	1.2	0	0.2	2	0.7	0	0	0	0
Someone else had to decide	0	0	6	6.6	0	0	3	0.9	0	0	4	2.1
Too far to travel	3	4.7	12	12.9	2	1.3	49	15.0	10	8.5	18	8.8
No transportation available	1	1.7	9	10.4	3	2.4	43	13.1	12	10.2	17	8.4
Cost (transport, health care, other)	3	5.1	13	14.4	5	3.9	17	5.3	4	3.0	6	3.0
Not satisfied with available health care	2	2.4	1	1.6	0	0	7	2.1	2	1.4	2	1.2
Symptom(s) required traditional care	0	0	3	2.8	0	0	0	0	1	0.7	2	0.8
Thought she/baby was too sick to travel	0	0	0	0	0	0	2	0.7	0	0	3	1.5
Thought she/baby will die despite care	0	0	0	0	0	0	0	0	0	0	6	2.9
Was late at night (transportation or provider not available)	–	–	–	–	1	0.6	5	1.5	0	0	1	0.4
Fears exposure to male health provider	0	0	1	0.8	0	0	2	0.7	–	–	–	–
Other	0	0	4	4.0	1	1.1	21	6.3	1	1.0	3	1.3
Total careseekers:	6	8.8	26	28.7	7	5.7	89	27.3	16	13.6	44	21.8
Total constraints:	10		52		12		163		30		80	

no formal education, suggests the need to provide young women and new mothers with information and support to better care for themselves and their newborn children. A community approach is necessary to accomplish this goal [41]. Ensuring universal access to four high quality antenatal care visits, including counseling on maternal health, normal newborn care, and pregnancy, delivery and newborn danger signs, also could help achieve this aim. The low median age at first marriage and low levels of formal education and antenatal care for mothers of surviving as well as deceased neonates shows that much work remains to be done to improve these indicators, and that encouraging girls and young women to delay marriage and instead go to school might help achieve these objectives.

Particular essential antenatal and delivery interventions that remain at low levels in the general population of Niger were found to be at very low levels among the mothers of deceased neonates, including taking a prophylactic anti–malarial medication during pregnancy, institutional delivery and attendance at birth by a skilled person. These finding suggest the need to expand the scope of the essential care package provided at integrated health centers and posts to better cover the needs of pregnant women as well as the accessibility of facilities.

The very low level of breastfeeding within one hour of birth found for the newborns who died, one–fourth of the general population's 43%, might have been partly due to sick newborns being unable to feed. However, the mean age at illness onset of 1.3 days for the two–thirds of neonates that died in the first week of life and the odds of not being breastfed within one hour of birth for neonates with illness onset at age 0 days vs 1 to 6 days (2.03, 95% CI=0.87, 4.73), argues that this was not a major factor. There was no comparison data for other normal newborn care measures, but the low levels for deceased neonates, summarized as the 1 of 408 newborns that received quality postnatal care in the first day of life, suggest that these factors contributed to the deaths. As with many of the maternal care indicators for which there was comparison data, the very low levels for the newborn care indicators among the deaths appear to be the extreme of overall low population levels that have helped maintain the high neonatal mortality rate in Niger. This is supported by the comparison of coverage data for newborn care indicators that are available both for Niger and other sub–Saharan Africa (SSA) countries, for postnatal care of all neonates within 2 days of birth (Niger, 13% vs 35% for 15 other SSA countries) and for registration of live births in 2012 by age 1 year (Niger, 30% vs 50% for 44 other SSA countries) [42].

Maternal complications and careseeking

Multiple studies conducted in low and middle income countries have found that maternal complications constitute the greatest risks for perinatal and early neonatal mortality [8–12]. Two–thirds of the neonatal deaths in our study occurred in the first week of life, and half the women with an early neonatal death had a pregnancy or labor and delivery complication. The much lower levels of maternal infection, antepartum hemorrhage and preeclampsia/eclampsia in mothers with a surviving neonate in Niger [12] and the generally far lower levels for these and other complications in women with a surviving neonate found by other studies [8–10] strongly suggest that the high rate of maternal complications identified by the VASA study significantly contributed to the neonatal deaths. This conclusion is supported by the positive relationship we identified between maternal infection and early onset severe neonatal infection as the primary cause of neonatal death. We are unaware of any prior study that has demonstrated this association at community level using verbal autopsy methodology.

Fewer than half the women with a pregnancy or labor and delivery complication sought formal health care. The most common reasons stated for not seeking care for a pregnancy complication, as well as for not delivering at a health facility, were distance to the facility, the lack of transportation and cost. However, the most striking finding was that nearly three–fourths of the women who did not seek care or deliver at a facility could not state a constraint to their use of the facility. This points to the need for further research to better understand this phenomenon, as well as the need to bring the required care to, or closer to, the community and to strengthen the links between communities and community–based providers with the health system. Recent studies have examined and shown promise for several potential strategies, including increased referrals to health facility for pregnancy related complications [43], community mobilization to increase institutional births, financial incentive plans and community referral/transport systems to increase rates of skilled birth attendance and the use of emergency obstetric care [14]; but it is not yet clear whether skilled birth attendance can be successfully provided in the community [15]. To help monitor such programs, periodic national surveys such as the DHS should consider adding questions on maternal complications, careseeking for these, and women's success in receiving care for these important risk factors for maternal and perinatal morbidity and mortality.

Failures in the pathway to survival

While almost all caregivers reported one or more signs of severe illness at the onset of their newborn's sickness, less than half sought or provided care. This decision appears to have been influenced by caregivers' perception of their child's illness severity, rather than their recognition of severe illness signs. We included feeding behavior, alertness

and activity level in ranking perceived illness severity, as these are signs that caregivers both recognize and perceive to indicate the need to seek formal health care for sick neonates [29–32] and, indeed, we found that careseeking was associated with these illness signs. The lower level of careseeking for the most severely ill newborns with early onset illness supports the hypothesis that caregivers believed they could not take any effective action for these children. Alternatively, their lack of action might have been due to there being less time to provide care before the children died.

In a multivariable analysis of careseeking for fatal newborn illnesses in Bangladesh that accounted for the competing risk of dying before careseeking, we also found that caregivers' reporting that their child was "not moving" at illness onset did not positively affect careseeking, while a child that was "less active than normal" was more likely to be taken for formal care [44], lending support to the severity perception theory. While several newborn illness danger signs are included in the WHO/UNICEF IMCI algorithm [29], it would appear to be more efficient and effective for careseeking messages at household and community level to focus on the signs that are most intrinsically perceived as indicating the need to seek care. The incongruity that we found between the illness severity ranks and some caregivers' apparent underestimate of their child's illness severity suggests that there is a need to provide caregivers with information even on the more well–perceived signs.

For the children whose illnesses began after the early neonatal period, in the aggregate careseeking was delayed until their illnesses had progressed from mild or moderate to severe. This may have contributed to the deaths, although this conclusion is tempered by the finding of equal severity among those who reached and did not reach the first health provider alive. Other studies both of fatal and nonfatal child illnesses have found that first careseeking to an informal provider contributed to delays in formal careseeking, which was undertaken only after the child's illness progressed to a severe state [31,33,45]. This was not an important factor in the current study, in which both informal and formal careseeking were sought for only two newborns and nine others received only informal care.

Illness severity was related to younger age of the child at illness onset and to faster illness progression, which in turn were related to the cause of death. While infectious diseases caused 60% of the neonatal deaths overall, and birth asphyxia and preterm delivery together accounted for another 27%, the distributions of these causes varied by age, place of death and care provided for the fatal illness. Birth asphyxia and preterm delivery were more common causes of early neonatal illnesses and deaths of children who died in the facility of their birth without leaving or at home without receiving or seeking care, compared to the predominance of infectious causes of the late neonatal illnesses and deaths of children who received home care or sought care from home. Yet, infections also were common causes of early neonatal deaths in facilities and caused twice as many early neonatal deaths in the community as did birth asphyxia.

The more similar cause distributions of neonates who died in their birth facility without leaving and at home without care suggests that the causes of death were as or more related to the children's ages than to any care they received or did not receive, though the still greater proportion of early neonatal infectious deaths in the community suggests that facilities may have had somewhat more success in treating infectious causes than intrapartum–related events and preterm birth complications. A facility–based study of presenting illnesses, treatments provided, quality of care, case fatality and cause–specific mortality rates would be needed to accurately determine the impact of facility birth on mortality and its cause distribution. The excess deaths from birth asphyxia in both settings and from preterm delivery in facilities, combined with the finding that women with delivery complications were equally likely to deliver at home as in a facility, suggests the need to improve the quality of intrapartum care to decrease deaths from these causes in the community and in health facilities. This strategy, along with postnatal care such as exclusive breastfeeding and providing antibiotics to reduce deaths from infectious causes, has been shown to provide the greatest reduction in neonatal mortality [13].

Fully 83% of the late neonatal deaths, that is, the deaths for which most affected children received some formal health care, were due to infectious causes. The limited comparison data available for these children suggest that young infants in Niger with a non–fatal fever or cough are taken for health care as or more often than severely ill late neonates. This corresponds with our finding that neonates who were moderately ill at illness onset were more likely to be taken for care than those who were severely ill. However, the decedents' younger age and the inability to match deaths with survivors on illness severity and the timing of careseeking underscores the difficulty in identifying an appropriate comparison group for examining careseeking in fatal illnesses. The finding that 8.5% of the late neonatal deaths were due to tetanus deserves further investigation that is beyond the scope of the current study.

As discussed for the early neonatal deaths, accurately assessing the quality of care provided in health facilities for the late neonatal deaths would require conducting a facility–based study. However, the VASA study offered some indication of problems that point to the need for just that. Caregivers reported that nearly one–fourth of the 86 neonates treated in a health facility were given an anti–malarial medication. If

correct, that would most likely be a treatment error, as such young children would be much more likely to have had bacterial sepsis than malaria [46]. The fact that caregivers rated 45% of the 69 children that left the first health provider alive as severely ill, while only 9% of these children were referred, is another indicator of a potentially serious deficit in the quality of care provided at these facilities.

Study limitations

The VASA study was not designed to examine changes in any factors over time and so cannot strictly conclude that a lack of change in the examined factors led to the non–decrease in mortality. However, the study provides plausible evidence by identifying low levels among the decedents of factors known to be critical to neonatal well–being and survival, and for many of these factors by showing that their levels in the decedents were significantly below those of surviving newborns from the same cohort. Also, the VASA study was retrospective in design, with an average recall period of 3.5 years. This could lead to inaccuracies in respondents' recall of events. Yet, the recall period for the survivors data examined for this analysis was similar to that of the decedents, which should minimize any recall bias in comparisons of findings for the decedents and survivors.

In conclusion, the VASA study revealed multiple factors contributing to the non–decrease in neonatal mortality that can be most effectively tackled through an integrated maternal–neonatal care package in the community and at health facilities. The predominance of rural residence and the role of distance and transport constraints to reaching a facility point to the need for Niger to scale up its recently implemented package of high–impact interventions to additional integrated health centers and posts. The low level of quality antenatal care and skilled birth attendance, high level of maternal complications, and many deaths from birth asphyxia and early onset severe neonatal infection in the community and health facilities call for expanding the package to provide antenatal and intrapartum care, with support for reaching a higher level facility when needed; while community education, mobilization and support are needed to improve illness recognition and careseeking for early and late onset severe neonatal infection. The quality of intrapartum and neonatal infectious disease care in first–level facilities and hospitals should be assessed and, if found to be required, improved.

Acknowledgments: We thank Narjis Kazmi for assisting with referencing of the paper. Oumarou Habi of the National Statistics Institute of Niger oversaw the collection of the VASA data as well as of the 2010 NNMS and 2012 NDHS data.

Funding: Funding was provided by the Bill and Melinda Gates Foundation through a grant to the U.S. Fund for UNICEF for the Child Health Epidemiology Reference Group, and by the UNICEF country office of Niger. The funders had no role in the study design, data collection and analysis, interpretation of data, decision to publish, or preparation of the manuscript.

Authorship declaration: REB conceived the overall VASA study project. HDK conceived the study, and HDK, AGY, AM, AKK, KB and AA contributed to designing the study and interpreting the data. HDK, AM and AKK contributed to acquiring the data. AM and AKK determined the levels of the comparison indicators from the 2010 NNMS and 2012 NDHS data. HDK and AKK analyzed the VASA data and HDK interpreted the comparisons with the 2010 NNMS, 2012 NDHS and other survivors' data and drafted the manuscript. All authors critically reviewed and provided comments on the draft paper and read and approved the final manuscript.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). The authors declare no competing interests.

REFERENCES

1 Amouzou A, Habi O, Bensaïd K. Reduction in child mortality in Niger: a Countdown to 2015 country case study. Lancet. 2012;380:1169-78. Medline:22999428 doi:10.1016/S0140-6736(12)61376-2
2 Institut National de la Statistique (INS). Enquete nationale sur la survie des enfants de 0–59 mois et la mortalité au Niger 2010. Niamey: INS, 2011.
3 Institut National de la Statistique. Enquete nationale sur les besoins en soins obstétricaux et néonatals d'urgence au Niger. Niamey, République du Niger, 2011.
4 The World health report: 2005: Make every mother and child count. World Health Organization, Geneva, 2005.
5 L'approche des besoins obstétricaux non couverts pour les interventions obstétricales majeures. Etude comparative Bénin, Burkina Faso, Haiti, Mali, Maroc, Niger, Pakistan et Tanzanie. [Tackling unmet needs for major obstetric interventions. Case studies in Benin, Burkina Faso, Haiti, Mali, Morocco, Niger, Pakistan and Tanzania]. Antwerp, Unmet Obstetric Needs Network, 2002:1–47 (www.uonn.org).

REFERENCES

6 Say L, Pattinson RC, Gülmezoglu AM. WHO systematic review of maternal morbidity and mortality: the prevalence of severe acute maternal morbidity (near miss). Reprod Health. 2004;1:3.. Medline:15357863 doi:10.1186/1742-4755-1-3

7 AbouZahr C. Global burden of maternal death and disability. In: Rodeck C, ed. Reducing maternal death and disability in pregnancy. Oxford, Oxford University Press, 2003:1–11.

8 Weiner R, Ronsmans C, Dorman E, Jilo H, Muhoro A, Shulman C. Labour complications remain the most important risk factors for perinatal mortality in rural Kenya. Bull World Health Organ. 2003;81:561-6. Medline:14576887

9 Mamun AA, Padmadas SS, Khatun M. Maternal health during pregnancy and perinatal mortality in Bangladesh: evidence from a large–scale community–based clinical trial. Paediatr Perinat Epidemiol. 2006;20:482-90. Medline:17052283 doi:10.1111/j.1365-3016.2006.00752.x

10 Kalter HD, Khazen RR, Barghouthi M, Odeh M. Prospective community–based cluster census and case–control study of stillbirths and neonatal deaths in the West Bank and Gaza Strip. Paediatr Perinat Epidemiol. 2008;22:321-33. Medline:18578745 doi:10.1111/j.1365-3016.2008.00943.x

11 Debelew GT, Afework MF, Yalew AW. Determinants and Causes of Neonatal Mortality in Jimma Zone, Southwest Ethiopia: A Multilevel Analysis of Prospective Follow Up Study. PLoS ONE. 2014;9:e107184. Medline:25232842 doi:10.1371/journal.pone.0107184

12 Vogel JP, Souza JP, Mori R, Morisaki N, Lumbiganon P, Laopaiboon M, et al; WHO Multicountry Survey on Maternal and Newborn Health Research Network. Maternal complications and perinatal mortality: findings of the World Health Organization Multicountry Survey on Maternal and Newborn Health. BJOG. 2014;121 Suppl. 1:76-88. Medline:24641538 doi:10.1111/1471-0528.12633

13 Darmstadt GL, Bhutta ZA, Cousens S, Adam T, Walker N, de Bernis L; Lancet Neonatal Survival Steering Team. Evidence–based, cost–effective interventions: how many newborn babies can we save? Lancet. 2005;365:977-88. Medline:15767001 doi:10.1016/S0140-6736(05)71088-6

14 Lee ACC, Lawn JE, Cousens S, Kumar V, Osrin D, Bhutta ZA, et al. Linking families and facilities for care at birth: What works to avert intrapartum–related deaths? Int J Gynaecol Obstet. 2009;107 Suppl 1:S65-S85. Medline:19815201 doi:10.1016/j.ijgo.2009.07.012

15 Darmstadt GL, Lee ACC, Cousens S, Sibley L, Bhutta ZA, Donnay F, et al. 60 million non–facility births: Who can deliver in community settings to reduce intrapartum–related deaths? Int J Gynaecol Obstet. 2009;107 Suppl 1:S89-112. Medline:19815200 doi:10.1016/j.ijgo.2009.07.010

16 Lawn JE, Kerber K, Enweronu–Laryea C, Bateman MO. Newborn survival in low resource settings—are we delivering? BJOG. 2009;116 Suppl 1:49-59. Medline:19740173 doi:10.1111/j.1471-0528.2009.02328.x

17 Fauveau V, de Bernis L. 'Good obstetrics' revisited: too many evidence–based practices and devices are not used. Int J Gynaecol Obstet. 2006;94:179-84. Medline:16844125 doi:10.1016/j.ijgo.2006.05.020

18 Engmann C, Wall S, Darmstadt GL, Valsangkar B, Claeson M. on behalf of the participants of the Istanbul KMC Acceleration Meeting. Consensus on kangaroo mother care acceleration. Lancet. 2013;382:e26-7. Medline:24246562 doi:10.1016/S0140-6736(13)62293-X

19 Blencowe H, Cousens S, Mullany LC, Lee ACC, Kerber K, Wall S, et al. Clean birth and postnatal care practices to reduce neonatal deaths from sepsis and tetanus: a systematic review and Delphi estimation of mortality effect. BMC Public Health. 2011;11 Suppl 3:S11. Medline:21501428 doi:10.1186/1471-2458-11-S3-S11

20 Mosha F, Winani S, Wood S, Changalucha J, Ngasalla B. Evaluation of the effectiveness of a clean delivery kit intervention in preventing cord infection and puerperal sepsis among neonates and their mothers in rural Mwanza Region, Tanzania. Tanzan Health Res Bull. 2005;7:185-8. Medline:16941946

21 Peters DH, Gray RH, Becker S, Gultiano SA, Black RE. A nested case-control study of case-fatality from diarrhea and pneumonia in Filipino children. Unpublished manuscript, available upon request from David H Peters at *dpeters@jhu.edu*.

22 Kalter HD, Salgado R, Babille M, Koffi AK, Black RE. Social autopsy for maternal and child deaths: a comprehensive literature review to examine the concept and the development of the method. Popul Health Metr. 2011;9:45. Medline:21819605 doi:10.1186/1478-7954-9-45

23 Kalter HD, Roubanatou A-M, Koffi A, Black RE. Direct estimates of national neonatal and young child cause–specific mortality proportions in Niger by expert algorithm and physician–coded analysis of verbal autopsy interviews. J Glob Health. 2015;5:010415. Medline:25969734 doi:10.7189/jogh.05.010415

24 Institut national de la statistique du Niger. Évaluation de la qualité des données de l'Enquęte nationale sur la survie des enfants de 0 ŕ 59 mois et la mortalité au Niger 2010. Available : http://www.stat-niger.org/statistique/file/ESM/2012/Evaluation_ESM2010.pdf. Accessed: 22 January 2016.

25 Murray CJL, lopez AD, Black R, Ahuja R, Ali SM, Baqui A, et al. Population Health Metrics Research Consortium gold standard verbal autopsy validation study: design, implementation, and development of analysis datasets. Popul Health Metr. 2011;9:27. Medline:21816095 doi:10.1186/1478-7954-9-27

26 SERPRO S.A. Santiago, Chile. CSProX software description, at http://serpro.com/csp.html

27 Claeson M, Waldman RJ. The evolution of child health programmes in developing countries: from targeting diseases to targeting people. Bull World Health Organ. 2000;78:1234-45. Medline:11100618

28 Friberg IK, Bhutta ZA, Darmstadt GL, Bang A, Cousens S, Baqui AH, et al. Comparing modelled predictions of neonatal mortality impacts using LiST with observed results of community–based intervention trials in South Asia. Int J Epidemiol. 2010;39:i11-20.. Medline:20348113 doi:10.1093/ije/dyq017

REFERENCES

29 Integrated Management of Childhood Illness. Chart Booklet. World Health Organization, Geneva, Switzerland, March 2014.

30 Sreeramareddy CT, Shankar RP, Sreekumaran BV, Subba SH, Joshi HS, Ramachandran U. Care seeking behaviour for childhood illness–a questionnaire survey in western Nepal. BMC Int Health Hum Rights. 2006;6:7. Medline:16719911 doi:10.1186/1472-698X-6-7

31 Mohan P, Iyengar SK, Agarwal K, Martines JC, Sen K. Care–seeking practices in rural Rajasthan: barriers and facilitating factors. J Perinatol. 2008;28 Suppl 2:S31-7. Medline:19057566 doi:10.1038/jp.2008.167

32 Hill Z, Kendall C, Arthur P, Kirkwood B, Adjei E. Recognizing childhood illnesses and their traditional explanations: exploring options for care–seeking interventions in the context of the IMCI strategy in rural Ghana. Trop Med Int Health. 2003;8:668-76. Medline:12828551 doi:10.1046/j.1365-3156.2003.01058.x

33 Awasthi S, Srivastava NM, Pant S. Symptom–specific care–seeking behavior for sick neonates among urban poor in Lucknow, Northern India. J Perinatol. 2008;28 Suppl 2:S69-75. Medline:19057571 doi:10.1038/jp.2008.169

34 Koffi AK, Libite P-R, Moluh S, Wounang R, Kalter HD. Social autopsy study identifies determinants of neonatal mortality in Doume, Nguelemendouka and Abong–Mbang health districts, Eastern Region of Cameroon. J Glob Health. 2015;5:010413. Medline:26171142 doi:10.7189/jogh.05.010413

35 Institut National de la Statistique (INS) et ICF International. 2013. Enquęte Démographique et de Santé et ŕ Indicateurs Multiples du Niger 2012. Calverton, Maryland, USA: INS et ICF International.

36 Chan GJ. Lee ACC, Baqui AH, Tan J, Black RE. Prevalence of early–onset neonatal infection among newborns of mothers with bacterial infection or colonization: a systematic review and meta–analysis. BMC Infect Dis. 2015;15:118. Medline:25886298

37 Guidelines on Maternal, Newborn, Child and Adolescent Health approved by the WHO Guidelines Review Committee; Recommendations on newborn health. 2013, WHO, Geneva.

38 Ministčre de la Santé. Accélération de l'atteinte des Objectifs du Millénaire pour le Développement (OMD 4, 5 et 6) Santé. Développement d'une approche communautaire intégrée autour de la case de santé. Promotion du paquet minimum d'activités de la case de santé. Direction Générale de la Santé; Direction de l'Organisation des Soins. Niamey, République du Niger, 2012.

39 Sullivan JM, Rutstein SO, Bicego GT. Infant and Child Mortality. DHS Comparative Studies No. 15. Calverton, Maryland: Macro International Inc., June 1994. Available at: https://dhsprogram.com/pubs/pdf/CS15/CS15.pdf. Accessed: 22 January 2016.

40 The World Health Report 2003: Shaping the Future. Chapter 1. WHO, Geneva 2003.

41 Bhutta ZA, Lassi ZS, Blanc A, Donnay F. Linkages among reproductive health, maternal health, and perinatal outcomes. Semin Perinatol. 2010;34:434-45. Medline:21094418 doi:10.1053/j.semperi.2010.09.002

42 Healthy Newborn Network. Global and National Newborn Health Data and Indicators; last updated Jan. 22, 2015. Available: http://www.healthynewbornnetwork.org/resource/database–global–and–national–newborn–health–data–and–indicators. Accessed: 22 January 2016.

43 Lassi ZS, Haider BA, Bhutta ZA. Community–based intervention packages for reducing maternal and neonatal morbidity and mortality and improving neonatal outcomes. Cochrane Database Syst Rev. 2010;(11):CD007754. Medline:21069697

44 Nonyane BAS, Kazmi N, Koffi A, Begum N, Ahmed S, Baqui AH, Kalter HD. Factors associated with delay in care–seeking for fatal neonatal illness in the Sylhet district of Bangladesh: results from a verbal and social autopsy study. J Glob Health. 2016;6:010605.

45 Schumacher R, Swedberg E, Diallo MO, Keita DR, Kalter HD, Pasha O. Mortality study in Guinea: Investigating the causes of death for children under 5. Save the Children Federation, Inc. and the Basic Support for Institutionalizing Child Survival (BASICS II) Project, 2002.

46 Tann CJ, Nkurunziza P, Nakakeeto M, Oweka J, Kurinczuk JJ, Were J, et al. Prevalence of bloodstream pathogens is higher in neonatal encephalopathy cases vs. controls using a novel panel of real–time PCR assays. PLoS ONE. 2014;9:e97259. Medline:24836781 doi:10.1371/journal.pone.0097259

journal of global health

Electronic supplementary material:
The online version of this article contains supplementary material.

PAPERS

Factors associated with delay in care–seeking for fatal neonatal illness in the Sylhet district of Bangladesh: results from a verbal and social autopsy study

Bareng AS Nonyane[1,2], Narjis Kazmi[1], Alain K Koffi[1], Nazma Begum[2], Salahuddin Ahmed[2], Abdullah H Baqui[1,2], Henry D Kalter[1]

[1] Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

[2] International Center for Maternal and Newborn Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Background We conducted a social and verbal autopsy study to determine cultural–, social– and health system–related factors that were associated with the delay in formal care seeking in Sylhet district, Bangladesh.

Methods Verbal and social autopsy interviews were conducted with mothers who experienced a neonatal death between October 2007 and May 2011. We fitted a semi–parametric regression model of the cumulative incidence of seeking formal care first, accounting for competing events of death or seeking informal care first.

Results Three hundred and thirty–one neonatal deaths were included in the analysis and of these, 91(27.5%) sought formal care first; 26 (7.9%) sought informal care first; 59 (17.8%) sought informal care only, and 155 (46.8%) did not seek any type of care. There was lower cumulative incidence of seeking formal care first for preterm neonates (sub–hazard ratio SHR 0.61, $P=0.025$), and those who delivered at home (SHR 0.52, $P=0.010$); and higher cumulative incidence for those who reported less than normal activity (SHR 1.95, $P=0.048$). The main barriers to seeking formal care reported by 165 mothers included cost (n=98, 59.4%), believing the neonate was going to die anyway (n=29, 17.7%), and believing traditional care was more appropriate (n=26, 15.8%).

Conclusions The majority of neonates died before formal care could be sought, but formal care was more likely to be sought than informal care. There were economic and social belief barriers to care–seeking. There is a need for programs that educate caregivers about well–recognized danger signs requiring timely care–seeking, particularly for preterm neonates and those who deliver at home.

Correspondence to:
Bareng Nonyane
Department of International Health
Johns Hopkins Bloomberg School of Public Health
615 N Wolfe Street
21205 Baltimore ML
USA
bnonyan1@jhu.edu

Recent estimates show that 6.3 million children died in 2013, which is a decline from 12.7 million in 1990 [1]. Of these 6.3 million, 2.8 million babies died in the neonatal period (within 28 days of birth) making up 44% of all under 5 deaths, a trend that has been observed over the last two to three decades [2–4]. Most of the under–5 deaths occur in developing countries where the estimated rate in 2012 was 53 per 1000 live births (90% uncertainty bound 51,56) compared to 6 (6,7) per 1000 live births in developed countries [5]. Most of the deaths in developing coun-

tries occur due to preventable causes for which care could be made available by targeted interventions [6]. Liu et al [7] estimated that in 2013, 2.761 million (44%) of the under–5 deaths were in the neonatal period, and among those 0.42 million neonates died of sepsis. Appropriate and timely care–seeking for infections could substantially improve neonatal survival. A systematic review of care–seeking behaviors for neonatal and childhood illnesses in low and middle–income countries showed that levels of care–seeking vary considerably by geographic region, with levels in South Asia being particularly low [8].

Delays in care–seeking and constraining factors

Care–seeking by itself is important, but even more important is the ability to reach out to available qualified health care providers as soon as the illness signs are recognized. The 3–delays model for care–seeking for maternal illnesses was developed by Thadeus and Maine (1994) [9] and has been applied to characterize the delay in seeking care for childhood illnesses [10–14].This model identifies the delay in 1) recognizing danger signs and deciding to seek care, 2) time to get to the health facility and 3) receiving adequate and appropriate care after reaching the facility.

Traditional beliefs and cultural practices have been shown to influence decision–making and the time to seeking care [15–19]. A specific example is that in Rajasthan, India, care–seeking for sick neonates was found to be less than that for older infants and children, mainly due to caretakers believing that even qualified providers lack the expertise to treat newborns [16–18]. Qualitative research in three South Asian countries (Bangladesh, Nepal and Pakistan) found that local traditions, lack of knowledge about the importance of care–seeking and recognition of danger signs and perceived poor quality of health services were important factors [19]. A focused ethnographic study in India found that despite caretakers recognizing danger signs indicating that their child needed health care, inability to discriminate among the available health care sources and perceived poor quality of health services led them to delay seeking care or to seek care from unqualified providers [20].

Similar results have been found in Sub–Saharan Africa and Latin America. For example, in some cases, caregivers assume that they know what the illness is and treat it at home instead of seeking formal care [12]. Other influences include lack of knowledge about dangerous illness symptoms [21], and various factors that ranged from social and traditional beliefs as well as poor health systems [10,11,22–25].

Care–seeking in Bangladesh

Bangladesh has a total population of 160 million and 9.6% of the population is under 5 years of age. The Bangladesh Demographic and Health Survey shows that the country had an under–5 mortality rate of 53 per 1000 live births and neonatal mortality rate of 32 per 1000 live births in 2011 [26]. Even though the country has met the Millennium Development Goal (MDG) 4 of reducing child mortality by two–thirds between 1990 and 2015, the neonatal mortality rate is still high and it requires targeted interventions to promote timely care–seeking for neonatal illness.

In a study done in rural Bangladesh to assess the care–seeking patterns for neonatal morbidity, it was seen that although most of the mothers sought outside care, only a small percentage of those who sought care considered going to a qualified provider and that was true even for the neonatal mortalities [27]. A recent study in Mirzapur Bangladesh found that even for older children (1–59 months) caretakers prefer to visit unqualified providers or other sources as compared to a formal provider [28]. Another study from Sylhet Bangladesh showed that preventative or curative care was sought for only 30.9% of preterm newborns from qualified providers [29]. A study by Chowdhury et al. [30] in Matlab, Bangladesh found that 37% of 365 neonates who had a fatal illness had formal care sought for them while the rest either received traditional or no care. The authors highlighted the need to design programs that take into consideration the use of traditional care and formal care in order to promote timely care–seeking.

Verbal and Social Autopsy studies to study care–seeking behavior

In low– and middle–income countries (LMICs), death registries are often either poorly kept or non–existent and, as a result, verbal autopsies (VAs) are used to help determine the likely cause of death. These are questionnaire instruments that are used to collect reported illness symptoms and information on pregnancy and intra–partum complications, in the case of neonates. The data are then used to determine the likely cause of death using physician assessments or expert algorithms [31–35]. Social autopsies (SAs) are conducted to help determine social and health system factors associated with care–seeking behavior [36,37].

There is very limited literature on comparative studies of those who seek formal care vs those who do not, and the factors associated with that decision [8,38]. In this study, we used data from a verbal and social autopsy (VASA) study in the Sylhet district of Bangladesh. We conducted an exploratory analysis to determine the illness symptoms as well as social and demographic factors that were associated with the delay in seeking formal care for a neonatal illness that led to death. Our analysis accounted for the competing risks of death and seeking informal care before formal care could be sought.

METHODS

Setting and participants

The VASA survey data were collected from four unions of Zakiganj sub–district with an estimated population of 102 000; and four unions of Kanaighat sub–district with an estimated population of 100 000 in the Sylhet district of Bangladesh. Initially, data were collected with a 1 year recall period, and this was extended for up to 2.5 years in order to attempt to achieve the desired sample size of up to 500 neonatal deaths. This sample size was determined for estimating the cause of death distribution (the first objective of the study) with 5% precision for the main common causes of death. The current–study's objective (delay in care–seeking) utilized a subset of these neonatal deaths that fitted the inclusion criteria. Babies born to ever–married mothers of child–bearing age (15 to 49 years old) were included. Deaths that occurred between October 2007 and May 2011 were included. Respondents were selected among participants of other studies on community–based interventions for maternal and newborn care. These were the Healthy Fertility study, which was aimed at improving healthy birth–spacing [39], and the Chlorhexidine trial which was a three–arm trial comparing the effect of umbilical cord cleansing with chlorhexidine once, over a 7–day period, and the control arm which was dry cord–care [39,40].The verbal autopsy questionnaires were administered retrospectively by trained female data collectors. Social autopsy interviews were conducted by trained female interviewers. Mothers who had had multiple deaths were interviewed for each of the deaths separately. The questionnaires used were the World Health Organization's (WHO) standard verbal autopsy tool [41] and the WHO/UNICEF–supported Child Health Epidemiology Reference Group (CHERG) social autopsy tool [36,42].

Analysis

The main event of interest was the time to seeking formal care, reported in days since illness onset. Illness onset was defined as the time when the first symptoms were recognized. Formal care in this context was defined as care provided by one of the following: a trained community health worker (CHW), private doctor or NGO/Government health center/post or hospital. Informal care was defined as seeking care from a traditional healer or from a pharmacist/drug seller. Other than the main outcome of seeking formal care, there were other possible events that may have taken place before formal care was sought and they were important to consider because they could have altered the probability of seeking formal care. These were a) death before any care was sought (survival bias) and b) seeking informal care first or only, and they were referred to as *competing risk events* [43]. Failure to account for these competing events (that is treating them as censoring events) may lead to an over–estimation of the incidence of seeking formal care and the effect estimates of the potential predictors. We calculated and plotted the cumulative incidence functions (the probability of an event of type k before or up to time t) for each of the possible events [44].

We then built a semiparametric regression model to estimate the cumulative incidence of seeking formal care first in the presence of competing events, and we report corresponding sub–hazard ratios (SHRs) with respect to each predictor [45,46]. The effect estimate is referred to as a 'sub–hazard' ratio because it pertains to one event among all possible events in any given time point. This model is analogous to the Cox proportional hazards model except that hazard ratios for an event k (such as seeking formal care) are calculated conditional on an individual having had no other event up to time t. SHRs are interpreted as a reduction or increase in cumulative incidence of an event. Unadjusted regression models were fitted for each potential predictor separately, and the predictors with a P–value ≤0.2 were included in the multivariable model. We investigated possible multi–collinearity among some of the predictors and its effect on the interpretation of the results.

A subset of the mothers who had not taken their neonate to a formal health care provider said that they had 'concerns' that prevented them from doing so. A subset of those who had sought formal care said they had 'concerns' that they had to overcome in doing so. Furthermore, those mothers who reported that their neonates died immediately were never asked about any concerns they had. These concerns were potential barriers to care–seeking and since they were not applicable or answered by all respondents, they were not included in the main regression analysis and only summary statistics of these are given.

Inclusion/exclusion

In order to assess care–seeking from home, only participants whose baby was either born at home, or left the delivery facility alive, were included in the analysis.

Potential predictors

We considered the following classes of predictors as shown in **Table 1**: neonate's demographic factors, neonatal care variables, illness symptoms, mother's/father's factors, household factors, social and health system factors. We used the WHO's Integrated Management of Childhood Illnesses (IMCI) severity grading for the first symptoms they observed. For the illness symptoms that were in the VA instrument but not in the IMCI, two physician authors (HDK, AKK) assigned symptoms as severe (requiring referral to higher level formal care) or possibly severe (requiring for-

Table 1. Potential predictors for care–seeking behavior

Factor	Variable
Neonate's demographic factors	• Age of neonate at illness onset (days) • Gender
Neonates care, illness symptoms/conditions	• Neonate ever breastfed? • Whether the newborn received any liquids or solids other than breast milk (exclusive breastfeeding); • Severity of the observed first symptoms for which care was reportedly sought • Birth size of the baby • Whether the neonate was preterm or not • Whether the baby had any malformation at birth • Mother/care–giver's perception of the status at illness onset: – feeding status (feeding normally, poorly or not at all) – activity (normally active, less active than normal or not moving)
Mother/father's factors	• Mother's age • Mother's education • Father's education • Whether the mother sought any antenatal care at a health provider • Whether mother had any pre–pregnancy medical condition
Household factors	• Whether mother was the household breadwinner • Where the mother stayed during the last days of her pregnancy
Specific barriers	• Any specific concerns or problems caregiver had: Thought baby not sick enough; no one available to go with caregiver; too much time from regular duties; someone else had to decide; too far to travel; no transportation available; cost; not satisfied with available health care; problem required traditional care; thought too sick to travel; thought child will die anyway; it was late at night; no transport/provider; other
Health system factors	• Delivery place • Time in minutes to usual health provider

mal health care). The listing of the symptoms and their severity scoring are given in **Online Supplementary Document.**

Ethical considerations

The VASA study was approved by the Institutional Review Boards of the Johns Hopkins University and the Bangladesh Medical Research Council. All respondents provided informed consent before being interviewed.

RESULTS

A total of 378 death records were available for analysis. Of these, 46 died at the facility in which they were born and thus, per inclusion/exclusion criteria, were excluded from the analysis; one had an unknown place of birth; and the rest (n = 331) were included in the analysis. The 331 comprised 307 babies born at home and 24 born at a facility and discharged alive. Furthermore, 12 of these mothers had given interviews about 2 deaths each. The recall period was 1 year in the initial study plan and under this criterion 72.5% (240/331) of the verbal autopsies were conducted. As the recall period was extended to 2.5 years in order to increase sample size, an additional 27.5% (91/331) VAs were conducted.

Table 2 shows some basic characteristics of these 331 participants and the types of care sought for them. The outcome of interest was seeking formal care first; and this was done for 91 (27.5%) neonates while 26 (7.9%) sought informal care before seeking formal care. Informal care only was sought for 59 (17.8%) neonates and 155 (46.8%) died before any care could be sought for them. The cumulative incidence functions for each of these competing events are given in **Figure 1**. These show that dying without any care sought was the most likely outcome, followed by seeking formal care first and seeking informal care only, while seeking informal care first was the least likely event.

The results of the unadjusted regression analysis for each potential predictor are given in **Table 3**, and the predictors with a *P*–value ≤0.2 were included in the multivariable model. The results from the multivariable model (**Table 4**)

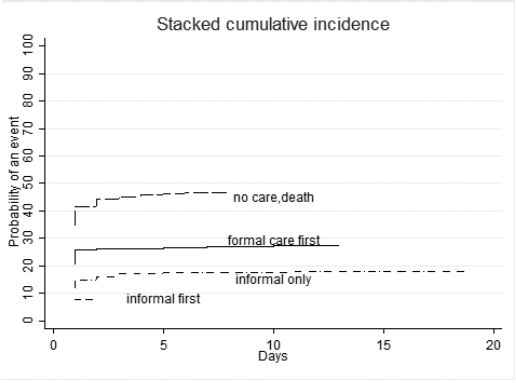


Figure 1. Cumulative incidence plots.

Table 2. Baseline characteristics of participants and care–seeking behavior

Variable	N = 331	%	Median (interquartile range)
Neonate's gender = female	147	44.4	
Age (days) at illness onset			0 (0,2)
Place of birth = hospital/facility	25	7.5	
Formal care sought first	91	27.5	
Informal care sought first	26	7.9	
Informal care only sought	59	17.8	
No care sought	155	46.8	
Median time (day of illness) to first seeking formal care			1 (1,2)
Median time (days) from illness onset to death			1 (0,2)
Median time(days) to informal care–seeking			1 (1,2)

show that there was a lower cumulative incidence of seeking care for preterm vs full–term neonates (SHR = 0.61, $P = 0.025$) and those who delivered at home (SHR = 0.52, $P < 0.010$). Those who reported less than normal activity in a neonate were more likely to seek formal care than those who reported normal activity (SHR = 1.95, $P = 0.025$).

Table 5 shows the summary data on the barriers to seeking care raised by respondents. A subset of 165 mothers reported having these concerns or barriers, and among those, 39 had taken their neonates to seek formal care despite the barriers. The most common barrier was cost, which was raised by 98 (60%) of the mothers, 67 of whom had not sought formal care. The next common barriers were: thinking that the baby would die anyway (n = 29, 18%), thinking the baby needed traditional care (n = 26, 16%), being too late at night to travel (n = 19, 12%), and the distance to the formal care facility (n = 18, 11%).

DISCUSSION

Care–seeking for neonatal and child illnesses in resource–limited settings is low. In this study, we aimed to determine factors associated with care–seeking behavior for fatal neonatal illness in Sylhet, Bangladesh, using data from verbal and social autopsy questionnaires. Our main outcome of interest was time to the first instance of seeking formal care. In order to conduct such an analysis, it was crucial to condition, at each time point during follow–up, on the neonate

Table 3. Characteristics of the sample and results of unadjusted regressions for the time in days to first instance of seeking formal care

Variable	N = 331	Sought formal care first (n, %)	Sub–hazard ratio for time to first seeking formal care	*P*–value
Age at illness onset–days:				
0	216	58 (26.9)	1	
1–3	26	8 (30.8)	1.2	0.64
3+	87	24 (27.6)	1.1	0.8
Missing	2			
Gender:				
Male	184	57 (30.9)	1	
Female	147	34 (23.1)	0.7	0.09
Newborn ever breastfed:				
No	170	38 (22.4)	1	
Yes	160	53 (33.3)	1.5	0.03
Don't know	1			
Were first symptoms observed severe/possibly severe?				
No	19	1 (20.3)	1	
Yes	267	78 (29.2)	2.74	0.15
Birth size:				
Very small/smaller than usual	214	58 (27.1)	1	
Average/larger than usual	117	33 (28.2)	0.99	0.96
Preterm birth:				
No	227	69 (30.4)	1	
Yes	103	22 (21.4)	0.66	0.08
Don't know	1			
Malformation at birth:				
No	325	89 (27.4)	1	
Yes	6	2 (33.3)	1.2	0.81
Feeding status at illness onset:				
Normal	43	11 (25.6)	1	
Poorly	109	41 (37.6)	1.5	0.21
Not at all	178	39 (21.9)	0.83	0.56
Activity status at illness onset:				
Normally active	53	10 (18.9)	1	
Less active than normal	182	64 (35.2)	1.9	0.04
Not moving	94	17 (18.1)	0.95	0.83
Missing	2	0		
Mother's age:				
18–20years	52	18 (34.6)	1	
21–25years	110	36 (32.7)	0.88	0.61

Table 3. Continued

Variable	N = 331	Sought formal care first (n, %)	Sub–hazard ratio for time to first seeking formal care	*P*–value
26–30years	97	18 (18.6)	0.47	0.02
>30years	72	19 (26.3)	0.72	0.27
Mother's education in years:				
0 years	113	26 (23.0)	1	
1–5 years	109	28 (25.7)	1.13	0.63
6–12 years	109	37 (33.9)	1.53	0.07
Any antenatal care for mother:				
No	142	90 (27.4)	1	
Yes	189	1 (33.3)	1.22	0.83
Mother had any pre–existing medical conditions:				
No	312	85 (27.2)	1	
Yes	19	6 (31.6	1.19	0.66
Breadwinner:				
Child's father	297	81 (27.3)	1	
Other	34	10 (29.4)	1.10	0.75
Place stayed during last days of pregnancy/during fatal illness:				
Own/current	303	83 (27.4)	1	
Parent's	25	7 (28.0)	1.02	0.96
Missing	3			
Mother able to turn to others for help?				
No	187	46 (24.6)	1	
Yes	144	45 (31.3)	1.30	0.17
Mother/her family ever been denied any community service:				
No	325	89 (27.4)	1	
Yes	6	2 (33.3)	1.24	0.73
Time to usual provider in minutes:				
<25	187	46 (24.6)	1	
≥25	144	45 (31.5)	1.28	0.19
Don't know	1			
Delivery place:				
Facility	25	15 (60.0)	1	
Home	306	76 (24.8)	0.38	<0.001

Table 4. Results from the multivariable regression of time to first instance of seeking formal care

Predictor	Sub–hazard ratio	P–value	95% CI lower limit	95% CI upper limit
Gender (male is reference)	0.72	0.094	0.46	1.06
Neonate ever breastfed by anyone? (no is reference)	1.45	0.106	0.92	2.28
Any severe symptom seen (no is reference)	2.54	0.181	0.65	9.92
Preterm (no is reference)	0.66	0.025	0.39	0.94
Activity at illness onset = less than normal (normal is reference)	1.95	0.048	1.01	3.78
Activity at illness onset = not moving (normal is reference)	1.12	0.796	0.48	2.59
Mother's age 21–25 y (18–20 y is reference)	1.07	0.759	0.67	1.73
Mother's age 26–30 y (18–20 y is reference)	0.56	0.057	0.31	1.02
Mother's age >30 (18–20 reference)	0.97	0.947	0.52	1.86
Mother's education 1–5 y (0 is reference)	1.19	0.486	0.72	1.99
Mother's education 6–12 y (0 is reference)	1.33	0.272	0.80	2.20
Mother could turn to others for help (no is reference)	1.14	0.486	0.79	1.64
Time to usual provider in minutes ≥25 (<25 is reference)	1.22	0.269	0.86	1.74
Delivery place = home (hospital/facility is reference)	0.52	0.010	0.32	0.85

CI – confidence interval, y – years

having survived a competing risk of death or seeking informal care first. Thus a competing risk time to event model was used.

In this Bangladesh cohort, the cumulative incidence analysis showed that neonates were more likely to die before any care could be sought, but that if care was sought, it was more likely to be formal vs informal care. Unadjusted analyses indicate that, formal care was sought first for only 27.5% of the sick neonates, and 47% never sought any care outside of the home. These results are similar to findings from other studies in rural Bangladesh with a level of care–seeking for fatal neonatal illnesses ranging from 35 to 37% [29,30].

Neonates whose activity level was reportedly less than normal had formal care sought first at a significantly higher rate than for those who were reported to have normal activity, while those who were reportedly not moving were neither more nor less likely to have formal care sought for them. This is evidence that perception of illness played a role in decision–making, and suggests that those who were reportedly not moving were thought to be less likely to survive even if care was sought. On the other hand, for those who were moving less than normal there was hope that they could survive if formal care was sought first. This corroborates the finding from other studies conducted in India

Table 5. Specific concerns/problems that were barriers to formal care–seeking–reported by mothers

	Total (N = 165)	Total N = 39 (of 165) sought formal care	Total N = 126 (of 165) who did not seek formal care
	n (%)	n (%)	n (%)
Mother's specific concern – thought not sick enough:			
No	151 (91.5)	39 (100)	112 (88.9)
Yes	14 (8.5)	0	14 (11.1)
Mother's specific concern – no one available to go with her:			
No	157 (95.2)	39 (100)	118 (93.7)
Yes	8 (4.8)	0	8 (6.3)
Mother's specific concern – cost:			
No	67 (40.6)	6 (15.4)	61 (48.4)
Yes	98 (59.4)	33 (84.6)	65 (51.6)
Mother's specific concern – too much time from regular duties:			
No	148 (89.7)	38 (97.4)	110 (87.3)
Yes	17 (10.3)	1 (2.6)	16 (12.7)
Mother's specific concern – someone else's decision:			
No	153 (92.7)	37 (94.9)	116 (92.1)
Yes	12 (7.3)	2 (5.13)	10 (7.9)
Mother's specific concern – too far to travel:			
No	147 (89.1)	36 (92.3)	111 (88.1)
Yes	18 (10.9)	3 (7.7)	15 (11.9)
Mother's specific concern – no transport:			
No	161 (97.6)	38 (97.4)	123 (97.6)
Yes	4 (2.4)	1 (2.6)	3 (2.4)
Mother's specific concern – thought needed traditional care:			
No	139 (84.2)	38 (97.4)	101 (80.2)
Yes	26 (15.8)	1 (2.6)	25 (19.8)
Mother's specific concern – too sick to travel:			
No	151 (91.5)	38 (97.4)	113 (89.7)
Yes	14 (8.5)	1 (2.6)	13 (10.3)
Mother's specific concern – thought would die anyway			
No	136 (82.4)	36 (92.3)	100 (79.4)
Yes	29 (17.6)	3 (7.7)	26 (20.6)
Mother's specific concern – late at night:			
No	146 (88.5)	37 (94.9)	109 (86.5)
Yes	19 (11.5)	2 (5.13)	17 (13.5)
Mother's specific concern – not satisfied with service at formal health facility:			
No	165 (100)	39 (100)	126 (100)
Yes	0	0	0
Mother's specific concern – other reasons:			
No	157 (95.2)	38 (97.4)	119 (94.4)
Yes	8 (4.9)	1 (2.6)	7 (5.6)

and Ghana that found that lethargy is one of the few signs of neonatal illness that mothers both recognize and understand to indicate the need for formal health care [15–17]. These studies distinguished such signs from others that mothers recognized but took to mean that traditional or no care was required and other illness signs that were poorly recognized. Thus, the current study adds to the evidence suggesting that focusing health messages on a few well–recognized and intrinsically motivating illness signs may be more effective in increasing formal health care–seeking for sick neonates than a strategy that urges care–seeking for all danger signs.

Preterm babies had care sought for them at a lower rate than full–term neonates. We found that among the subset that reported barriers to seeking formal care, 24% of the mothers who had preterm babies thought the babies would die anyway while only 14% of the mothers with full–term neonates felt that way. Thus, even though it is not possible to adjust for the reported barriers in the main model, it may be that being preterm led the mothers to think that there was no hope of survival.

We investigated whether those who were born at a health facility survived longer before illness onset and whether care–seeking decisions were different for them. The numbers that can be used to make any such inference were small. Of the total 331 participants, 25 were born at a facility. Of those 25, 15 (63%) reported that illness began on day 1 of life, which is comparable to 66% of the home-births. Furthermore, 17 (71%) of the facility births sought formal care on the same day as illness onset, which is also comparable to the 228 (75%) of home–births. Hence, we conclude that the small subset of facility births were not different in terms of their survival before illness onset, nor by the care–seeking decisions made for them.

Neonates who were delivered at home were less likely to have care sought for them. This was an expected finding because mothers who are less likely to seek care at facilities are also less likely to deliver there. In this Bangladesh cohort, the reported barriers to seeking care have also been found in similar low–income settings [10,11,22–25]. The main ones were the cost, having no hope of survival or believing that traditional care was more appropriate for some illnesses.

Strengths and limitations

The main strength of our study is that it was a comparative analysis of those caregivers who did vs those who did not seek formal care, accounting for competing risks of death or seeking informal care first. It thus allowed for a more definitive determination of the factors that constrain prompt formal health care–seeking fatally ill neonates. Data were collected by highly–trained and skilled female verbal autopsy interviewers working for the HFS and CHX studies as well as for the social autopsy component. Furthermore, this was a community–based study that included both home and hospital births, and thus it did not suffer selection bias that would occur if only hospital births had been included.

A limitation to our study is that there may have been bias in self–reported time to seeking care (days/h) given that the recall period was up to 2.5 years for some of the respondents.

Not all mothers responded to the questions of the specific barriers that prevented them from seeking care or that they

had to overcome. Thus it was not possible to include responses to these barriers in the main multivariable model. However, these reported barriers have also been found to apply in other similar settings.

The time to care–seeking data was not presented in hours for all types of events and thus we were not able to fully describe the pattern of events for the first day of life which is a crucial time when most deaths and actions are taken on. Time in hours was only asked of those who sought care at a formal provider and thus we could not give a comparative analysis of the events.

CONCLUSION

Our analysis has shown that in this rural Bangladesh cohort, the majority of the neonates died before formal care could be sought for them, but when care was sought it was more likely to be formal vs informal care. There were economic and social belief barriers that delayed or prevented care–seeking. There is a need for programs that address such barriers and educate caregivers about danger signs that require formal health care and the importance of timely care–seeking, particularly for preterm neonates and those who deliver at home.

Acknowledgements: The authors thank the study participants and field staff for their efforts in implementing the study, and the Ministry of Health and Family Welfare, Government of Bangladesh.

Funding information: Funding for the Verbal and Social Autopsy field work was provided by U.S. Agency for International Development/Bangladesh and the Office of Population and Reproductive Health, U.S. Agency for International Development/Washington D.C., under the terms of Award No GHS–A–00–08–00002–00 (Maternal and Child Health Integrated Program (MCHIP) Leader with Associates Cooperative Agreement), No. GPO–AA–05–00025–00 (Associate Cooperative Agreement with the Access Program), No. GHS–A–00–04–00002–00 (Reference Leader Cooperative Agreement with the Access Program), and No. GHS–A–00–03–00019–00 (Global Research Activity Cooperative Agreement with the Johns Hopkins Bloomberg School of Public Health). Funding for data analysis was provided by the Bill and Melinda Gates Foundation through a grant to the U.S. Fund for UNICEF for the Child Health Epidemiology Reference Group. The authors' views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development, the United States Government, and/or the decisions, policy, or views of their respective organizations.

Authors' contributions: NB, SA, AHB, HDK contributed to the management and conduct of the field work. NB and SA conducted data management. HDK and AKK supervised data management and analysis. NK conducted data coding and literature review. BASN conducted data analysis and drafted the manuscript. All authors critically reviewed and commented on the draft paper and approved the final manuscript.

Competing interests: The author completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). None of the authors has any competing interests to declare.

REFERENCES

1 UN Inter–agency Group for Child Mortality. Child Mortality Estimates September 2014. Available at: http://www.childmortality.org/. Accessed: 4 January 2016.

2 Lawn JE, Blencowe H, Oza S, You D, Lee AC, Waiswa P, et al. Every Newborn: progress, priorities, and potential beyond survival. Lancet. 2014;384:189-205. Medline:24853593 doi:10.1016/S0140-6736(14)60496-7

3 Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis. Lancet. 2010;375:1969-87. Medline:20466419 doi:10.1016/S0140-6736(10)60549-1

4 You D, Jones G, Hill K, Wardlaw T, Chopra M. Levels and trends in child mortality, 1990–2009. Lancet. 2010;376:931-3. Medline:20851244 doi:10.1016/S0140-6736(10)61429-8

5 UNICEF. Levels and trends in child mortality: report 2013. New York, 2013.

6 Lawn JE, Cousens S, Zupan J. 4 million neonatal deaths: when? Where? Why? Lancet. 2005;365:891-900. Medline:15752534 doi:10.1016/S0140-6736(05)71048-5

7 Liu L, Oza S, Hogan D, Perin J, Rudan I, Lawn JE, et al. Global, regional, and national causes of child mortality in 2000–13, with projections to inform post–2015 priorities: an updated systematic analysis. Lancet. 2015;385:430-40. Medline:25280870 doi:10.1016/S0140-6736(14)61698-6

REFERENCES

8 Herbert HK, Lee AC, Chandran A, Rudan I, Baqui AH. Care seeking for neonatal illness in low– and middle–income countries: a systematic review. PLoS Med. 2012;9:e1001183. Medline:22412355 doi:10.1371/journal.pmed.1001183

9 Thaddeus S, Maine D. Too far to walk: maternal mortality in context. Soc Sci Med. 1994;38:1091-110. Medline:8042057

10 Källander K, Hildenwall H, Waiswa P, Galiwango E, Peterson S, Pariyo G. Delayed care seeking for fatal pneumonia in children aged under five years in Uganda: a case–series study. Bull World Health Organ. 2008;86:332-8. Medline:18545734 doi:10.2471/BLT.07.049353

11 Waiswa P, Kallander K, Peterson S, Tomson G, Pariyo GW. Using the three delays model to understand why newborn babies die in eastern Uganda. Trop Med Int Health. 2010;15:964-72. Medline:20636527 doi:10.1111/j.1365-3156.2010.02557.x

12 Samuelsen H, Tersbol BP, Mbuyita SS. Do health systems delay the treatment of poor children? A qualitative study of child deaths in rural Tanzania. BMC Health Serv Res. 2013;13:67. Medline:23421705 doi:10.1186/1472-6963-13-67

13 Upadhyay RP, Rai SK, Krishnan A. Using three delays model to understand the social factors responsible for neonatal deaths in rural Haryana, India. J Trop Pediatr. 2013;59:100-5. Medline:23174990 doi:10.1093/tropej/fms060

14 Applying the Three Delays Model: improving access to care for newborns with danger signs. Save The Children, 2013. Available at: http://www.healthynewbornnetwork.org/hnn-content/uploads/Applying-the-three-delays-model_Final.pdf. Accessed: 4 January 2016.

15 Awasthi S, Srivastava NM, Pant S. Symptom–specific care–seeking behavior for sick neonates among urban poor in Lucknow, Northern India. J Perinatol. 2008;28 Suppl 2:S69-75. Medline:19057571 doi:10.1038/jp.2008.169

16 Hill Z, Kendall C, Arthur P, Kirkwood B, Adjei E. Recognizing childhood illnesses and their traditional explanations: exploring options for care–seeking interventions in the context of the IMCI strategy in rural Ghana. Trop Med Int Health. 2003;8:668-76. Medline:12828551 doi:10.1046/j.1365-3156.2003.01058.x

17 Mohan P, Iyengar SD, Agarwal K, Martines JC, Sen K. Care–seeking practices in rural Rajasthan: barriers and facilitating factors. J Perinatol. 2008;28 Suppl 2:S31-7. Medline:19057566 doi:10.1038/jp.2008.167

18 Sreeramareddy CT, Shankar RP, Sreekumaran BV, Subba SH, Joshi HS, Ramachandran U. Care seeking behaviour for childhood illness—a questionnaire survey in western Nepal. BMC Int Health Hum Rights. 2006;6:7. Medline:16719911 doi:10.1186/1472-698X-6-7

19 Syed U, Khadka N, Khan A, Wall S. Care–seeking practices in South Asia: using formative research to design program interventions to save newborn lives. J Perinatol. 2008;28 Suppl 2:S9-13. Medline:19057572 doi:10.1038/jp.2008.165

20 de Zoysa I, Bhandari N, Akhtari N, Bhan MK. Careseeking for illness in young infants in an urban slum in India. Soc Sci Med. 1998;47:2101-11. Medline:10075250 doi:10.1016/S0277-9536(98)00275-5

21 Ogunlesi TA, Ogunlesi FB. Family Socio–Demographic Factors and Maternal Obstetric Factors Influencing Appropriate Health–Care Seeking Behaviours for Newborn Jaundice in Sagamu, Nigeria. Matern Child Health J. 2012;16:677-84. Medline:21365297 doi:10.1007/s10995-011-0765-1

22 Assefa T, Belachew T, Tegegn A. Mothers' health care seeking behavior for childhood illnesses in Derra district, Northshoa zone, Oromia regional state, Ethiopia. Ethiopian Journal of Health Sciences. 2008;18:87-94.

23 Hildenwall H, Tomson G, Kaija J, Pariyo G, Peterson S. "I never had the money for blood testing"–caretakers' experiences of care–seeking for fatal childhood fevers in rural Uganda–a mixed methods study. BMC Int Health Hum Rights. 2008;8:12. Medline:19055738 doi:10.1186/1472-698X-8-12

24 Ogunlesi TA, Olanrewaju DM. Socio–demographic factors and appropriate health care–seeking behavior for childhood illnesses. J Trop Pediatr. 2010;56:379-85. Medline:20167633 doi:10.1093/tropej/fmq009

25 Terra de Souza AC, Peterson KE, Andrade FM, Gardner J, Ascherio A. Circumstances of post–neonatal deaths in Ceara, Northeast Brazil: mothers' health care–seeking behaviors during their infants' fatal illness. Soc Sci Med. 2000;51:1675-93. Medline:11072887 doi:10.1016/S0277-9536(00)00100-3

26 Bangladesh Demographic and Health Survey, 2011. Available at: http://pdf.usaid.gov/pdf_docs/pnady754.pdf. Accessed: 4 January 2016.

27 Ahmed S, Sobhan F, Islam A. Neonatal Morbidity and care–seeking behaviour in rural areas of Bangladesh. J Trop Pediatr. 2001;47:98-105. Medline:11336143 doi:10.1093/tropej/47.2.98

28 Das SK, Nasrin D, Ahmed S, Wu Y, Ferdous F, Farzana FD, et al. Health care–seeking behavior for childhood diarrhea in Mirzapur, rural Bangladesh. Am J Trop Med Hyg. 2013;89(Suppl):62-8. Medline:23629937 doi:10.4269/ajtmh.13-0107

29 Shah R, Mullany LC, Darmstadt GL, Talukder RR, Rahman SM, Mannan I, et al. Determinants and pattern of care seeking for preterm newborns in a rural Bangladeshi cohort. BMC Health Serv Res. 2014;14:417. Medline:25242278 doi:10.1186/1472-6963-14-417

30 Chowdhury HR, Thompson SC, Ali M, Alam N, Yunus M, Streatfield PK. Care seeking for fatal illness episodes in neonates: a population–based study in rural Bangladesh. BMC Pediatr. 2011;11:88. Medline:21999253 doi:10.1186/1471-2431-11-88

31 Murray CJ, Lopez AD, Black R, Ahuja R, Ali SM, Baqui A, et al. Population Health Metrics Research Consortium gold standard verbal autopsy validation study: design, implementation, and development of analysis datasets. Popul Health Metr. 2011;9:27. Medline:21816095 doi:10.1186/1478-7954-9-27

REFERENCES

32 Murray CJ, Lozano R, Flaxman AD, Serina P, Phillips D, Stewart A, et al. Using verbal autopsy to measure causes of death: the comparative performance of existing methods. BMC Med. 2014;12:5. Medline:24405531 doi:10.1186/1741-7015-12-5

33 Baqui AH, Black RE, Arifeen SE, Hill K, Mitra SN, al Sabir A. Causes of childhood deaths in Bangladesh: results of a nationwide verbal autopsy study. Bull World Health Organ. 1998;76:161-71. Medline:9648357

34 Baqui AH, Darmstadt GL, Williams EK, Kumar V, Kiran TU, Panwar D, et al. Rates, timing and causes of neonatal deaths in rural India: implications for neonatal health programmes. Bull World Health Organ. 2006;84:706-13. Medline:17128340 doi:10.2471/BLT.05.026443

35 Engmann C, Garces A, Jehan I, Ditekemena J, Phiri M, Mazariegos M, et al. Causes of community stillbirths and early neonatal deaths in low–income countries using verbal autopsy: an International, Multicenter Study. J Perinatol. 2012;32:585-92. Medline:22076413 doi:10.1038/jp.2011.154

36 Kalter HD, Salgado R, Babille M, Koffi AK, Black RE. Social autopsy for maternal and child deaths: a comprehensive literature review to examine the concept and the development of the method. Popul Health Metr. 2011;9:45. Medline:21819605 doi:10.1186/1478-7954-9-45

37 Waiswa P, Kalter HD, Jakob R, Black RE, Social Autopsy Working Group. Increased use of social autopsy is needed to improve maternal, neonatal and child health programmes in low–income countries. Bull World Health Organ. 2012;90:403–A.

38 Kalter HD, Salgado R, Moulton LH, Nieto P, Contreras A, Egas ML, et al. Factors constraining adherence to referral advice for severely ill children managed by the Integrated Management of Childhood Illness approach in Imbabura Province, Ecuador. Acta Paediatr. 2003;92:103-10. Medline:12650309 doi:10.1111/j.1651-2227.2003.tb00478.x

39 Ahmed S, Norton M, Williams E, Ahmed S, Shah R, Begum N. Operations research to add postpartum family planning to maternal and neonatal health to improve birth spacing in Sylhet District, Bangladesh. Glob Health Sci Pract. 2013;1:262-76. Medline:25276538 doi:10.9745/GHSP-D-13-00002

40 Mullany LC, El Arifeen S, Winch PJ, Shah R, Mannan I, Rahman SM, et al. Impact of 4.0% chlorhexidine cleansing of the umbilical cord on mortality and omphalitis among newborns of Sylhet, Bangladesh: design of a community–based cluster randomized trial. BMC Pediatr. 2009;9:67. Medline:19845951 doi:10.1186/1471-2431-9-67

41 World Health Organization. Standard neonatal verbal autopsy questionnaire, revised version. World Health Organization. Geneva, WHO, 2003.

42 CHERG Verbal and Social Autopsy Survey Tools and Information 2014. Available at: http://cherg.org/projects/pathways.html#socialautopsy. Accessed on: 4 January 2016.

43 Satagopan JM, Ben–Porat L, Berwick M, Robson M, Kutler D, Auerbach AD. A note on competing risks in survival data analysis. Br J Cancer. 2004;91:1229-35. Medline:15305188 doi:10.1038/sj.bjc.6602102

44 Coviello V, Bogess M. Cumulative incidence estimation in the presence of competing risks. Stata J. 2004;4:103-12.

45 Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc. 1999;94:496-509. doi:10.1080/01621459.1999.10474144

46 Cleves M, Gutierres RG, Gould W, Marchencko YV. An introduction to survival analysis using STATA. College Station, Texas: Stata Press Publication, StataCorp LP; 2010.

journal of global health

Electronic supplementary material:
The online version of this article contains supplementary material.

PAPERS

What is kangaroo mother care? Systematic review of the literature

Grace J Chan[1,2,3], Bina Valsangkar[3], Sandhya Kajeepeta[2], Ellen O Boundy[2], Stephen Wall[3]

[1] Department of Pediatrics, Harvard Medical School, Boston, MA, USA
[2] Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Boston, MA, USA
[3] Saving Newborn Lives, Save the Children, Washington, DC, USA

Background Kangaroo mother care (KMC), often defined as skin–to–skin contact between a mother and her newborn, frequent or exclusive breastfeeding, and early discharge from the hospital has been effective in reducing the risk of mortality among preterm and low birth weight infants. Research studies and program implementation of KMC have used various definitions.

Objectives To describe the current definitions of KMC in various settings, analyze the presence or absence of KMC components in each definition, and present a core definition of KMC based on common components that are present in KMC literature.

Methods We conducted a systematic review and searched PubMed, Embase, Scopus, Web of Science, and the World Health Organization Regional Databases for studies with key words "kangaroo mother care", "kangaroo care" or "skin to skin care" from 1 January 1960 to 24 April 2014. Two independent reviewers screened articles and abstracted data.

Findings We screened 1035 articles and reports; 299 contained data on KMC and neonatal outcomes or qualitative information on KMC implementation. Eighty–eight of the studies (29%) did not define KMC. Two hundred and eleven studies (71%) included skin–to–skin contact (SSC) in their KMC definition, 49 (16%) included exclusive or nearly exclusive breastfeeding, 22 (7%) included early discharge criteria, and 36 (12%) included follow–up after discharge. One hundred and sixty–seven studies (56%) described the hours per day of SSC.

Conclusions There exists significant heterogeneity in the definition of KMC. A large number of studies did not report definitions of KMC. Skin–to–skin contact is the core component of KMC, whereas components such as breastfeeding, early discharge, and follow–up care are context specific. To implement KMC effectively development of a global standardized definition of KMC is needed.

Globally, 44% of under–five deaths occur during the neonatal period, and the proportion of under–five deaths due to neonatal causes continues to rise [1,2]. Preterm birth (before 37 weeks gestation) accounts for 35% of neonatal deaths. Low birth weight (defined as <2500 g) is commonly used as a surrogate measure of preterm birth [3]. Preterm and low birth weight infants who survive the neonatal period are more likely to experience neonatal morbidities including acute respiratory, gastrointestinal, immunologic, central nervous system, hearing and vision problems than both term and normal weight infants [4].

Correspondence to:
Grace J Chan
677 Huntington Ave
Building 1, #1108
Boston, MA
USA 02115
grace.chan@hsph.harvard.edu

A significant proportion of deaths among preterm and low birth weight infants is preventable. There is evidence that kangaroo mother care (KMC), when compared to conventional neonatal care in resource–limited settings, significantly reduces the risk of mortality in infants born in facilities who are clinically stable and weighing less than 2000 g [5]. KMC also reduces the risk of hypothermia, severe illness, nosocomial infection, and length of hospital stay, and improves growth, breastfeeding, and maternal–infant attachment [5,6].

Despite strong evidence for mortality and morbidity reduction in low– and middle–income settings and endorsement from the World Health Organization (WHO), country–level adoption and implementation of KMC has been limited. In a systematic assessment of health system bottlenecks among countries with a high burden of neonatal deaths, KMC was identified as an intervention with significant health systems barriers to scale–up including leadership and governance, health financing, health workforce, health service delivery, health information systems, and community ownership and partnership [7]. Health intervention priority–setting tools, such as the Lives Saved Tool and Child Health and Nutrition Research Initiative methodology, have identified KMC as a high priority intervention based on criteria such as mortality benefit and equity [8,9].

In response to limited global uptake of KMC, in 2013, a group of newborn health stakeholders led by the Bill and Melinda Gates Foundation and Save the Children's Saving Newborn Lives Program launched a global KMC Acceleration Convening. The goal was to address barriers to implementation, increase uptake of KMC as part of an integrated Reproductive Maternal Newborn and Child Health package, and identify research priorities [10]. In addition to implementation barriers, a lack of a clear definition of KMC has made *effective* coverage at scale of KMC challenging. A multi–country study in Africa found variation in KMC implementation across facilities in countries with national commitment to KMC [11]. Regional, country, and facility differences in health worker capacity, financial resources, leadership, health information systems, and cultural and community structures create challenges to developing and adopting a global definition of KMC.

The WHO has defined KMC as early, continuous, and prolonged skin–to–skin contact (SSC) between the mother and preterm babies; exclusive breastfeeding or breast milk feeding; early discharge after hospital–initiated KMC with continuation at home; and adequate support and follow–up for mothers at home [12]. While the WHO provides guidance on the components of KMC, guidance on the operationalization and clinical implementation of KMC are needed. There are significant variations in the timing of initiation, duration of SSC, positioning, necessary equipment and supplies, discharge criteria, follow–up frequency, indicators and measurement, and health workforce needs. The variations in these components have differential effects on preterm and low birth weight outcomes. As the global newborn health community begins to accelerate implementation of KMC, a standardized operational definition is needed. We conducted a systematic review of the KMC literature to 1) describe the current definitions of KMC in various settings, 2) analyze the presence or absence of WHO KMC components in each definition, and 3) present a core definition of KMC–common components that are present in at least 70% of all studies and programs–and describe how KMC definitions vary by context. This review provides a basis for development of an operational definition and clinical standards to accelerate the uptake of KMC globally.

METHODS

We searched PubMed, Embase, Web of Science, Scopus, and WHO regional databases: AIM, LILACS, IMEMR, IMSEAR, and WPRIM using the search terms "kangaroo mother care", "kangaroo care", and "skin to skin care" with no language restrictions from 1 January 1960 to 24 April 2014 for original reports including case–control studies, cohort studies, randomized control trials, and case series with 10 or more participants (see **Online Supplementary Document** for the review protocol and full search strategy). Following PRISMA guidelines, studies were included if they contained at least one of the following: the amount of time KMC was practiced, an association between KMC (as an isolated exposure, not part of a larger package) add any outcome, barriers to implementing KMC or factors necessary for successful implementation of KMC. Exclusion criteria were non–human subjects, case series or descriptive studies with fewer than 10 participants, and non–primary data collection or analysis (eg, reviews, meeting abstracts, editorials). Our population of interest included mothers, newborns, or mother–newborn dyads (not restricted to any specific ages) who have practiced KMC as well as health care providers, health facilities, communities, and health systems that have implemented KMC.

We also conducted hand–searches through the reference lists of the articles included in our review and published systematic reviews. Cochrane reviews were searched for relevant articles. To search the "grey literature" for unpublished studies, we explored programmatic reports and requested data from programs implementing KMC to obtain programmatic perspectives in addition to those provided by research studies. Reports were included following the same criteria as above.

Two independent reviewers examined titles, abstracts and full–text articles for inclusion into the review using a

screening form based on our inclusion criteria. Using standardized data abstraction forms, two reviewers abstracted data independently from all included articles and reports. At each stage, reviewers compared results to ensure agreement. In the case of disagreement between the two reviewers, a third party acted as a tiebreaker. Native speakers abstracted data from articles in foreign languages. Languages for which a native speaker was not identified (ie, German, Finnish, Korean, Thai and Polish) were translated using an online translation software to assist with data abstraction. If an article or report were missing any information, we contacted the authors to request the data.

Using standardized forms, data were abstracted on study characteristics such as study design, country, sample size, location, and duration of follow–up. We abstracted data on KMC definitions including data on SSC, exclusive breastfeeding, early discharge from the facility, and follow–up and as well as other components [12]. We generated categorical variables for each component and calculated descriptive frequencies, means, medians and ranges for quantitative data.

RESULTS

Study selection and characteristics

Our search strategy yielded 1035 records of which 299 were included in our review (**Figure 1**). Details of each included study are found in Table S1 in **Online Supplementary Document**. Summary characteristics of the included studies are presented in **Table 1**. In the last five years, as KMC research gaps have gained growing attention, the number of studies conducted has increased. One hundred and thirty–four studies (45%) were published in the last five years between 2010 and 2014, 134 (45%) between 2000 and 2009, and 31 (10%) between 1988 and 1999. Common study types were randomized control trials (n = 85, 28%), surveys or interviews (n = 58, 19%), and cohorts (n = 43, 14%). Other study types included pre–post studies, facility–level evaluations, non–randomized intervention studies, and randomized crossover trials. One hundred and forty–four studies (48%) had less than 50 participants and 47 (16%) had 200 or more participants. Geographically, 115 (38%) of the studies took place in the Americas, 64 (21%) in Europe, 44 (15%) in Africa, 29 (10%) in Southeast Asia, 20 (7%) in Western Pacific, and 16 (5%) in Eastern Mediterranean regions. More studies were in countries with low neonatal mortality rates (NMRs), ie, less than 5 per 100 live births (n = 130, 43%), than in countries with high NMRs, ie, 30 or higher (n = 10, 3%) [13]. The majority of studies, 192 (64%), were in an urban setting. One hundred and seventy–five studies (59%) took place in health facilities, 107 (36%) in neonatal intensive

Table 1. Characteristics of included studies

	N = 299	%
Year:		
2010 to 2014	134	44.82
2000 to 2009	134	44.82
1988 to 1999	31	10.36
Sample size:		
<50	144	48.16
50 to <100	61	20.40
100 to <200	47	15.72
≥200	47	15.72
Study type:		
Randomized control trial	85	28.43
Surveys or interview	58	19.40
Cohort study	43	14.38
Pre–post intervention study	33	11.04
Facilities evaluation	23	7.69
Intervention trial	15	5.02
Randomized cross over	14	4.68
Other (chart review, case–control, cross over, surveillance)	28	9.36
World Health Organization region:		
Americas	115	38.46
Europe	64	21.40
Africa	44	14.72
Southeast Asia	29	9.70
Western Pacific	20	6.69
Eastern Mediterranean	16	5.35
Multiple regions	4	1.34
Missing	7	2.34
Neonatal mortality rate (death per 1000 live birth):		
<5	130	43.48
5 to <15	84	28.09
15 to <30	66	22.07
≥30	10	3.34
Missing	9	3.01
Setting (rural or urban):		
Urban	192	64.21
Urban and rural	23	7.69
Rural	10	3.34
Missing	74	24.75
Population source:		
Health facility	175	58.53
Neonatal intensive care unit or stepdown unit	107	35.79
Community or population–based surveillance	11	3.68
Missing	6	2.01
Gestational age:		
Preterm 34 to <37 weeks	57	19.06
Very preterm <34 weeks	51	17.06
Full term ≥37 weeks	33	11.04
Mixed preterm and very preterm <37 weeks	26	8.70
All gestational ages	28	9.36
Missing	104	34.78
Birth weight:		
Low birth weight 1500 to <2500 g	52	17.39
Mixed low <2500 g and very low birth weight <1500 g	45	15.05
All birth weights	25	8.36
Very low birth weight <1500 g	21	7.02
Non low birth weight ≥2500 g	9	3.01
Low birth weight vs non–low birth weight	1	0.33
Missing	146	48.83

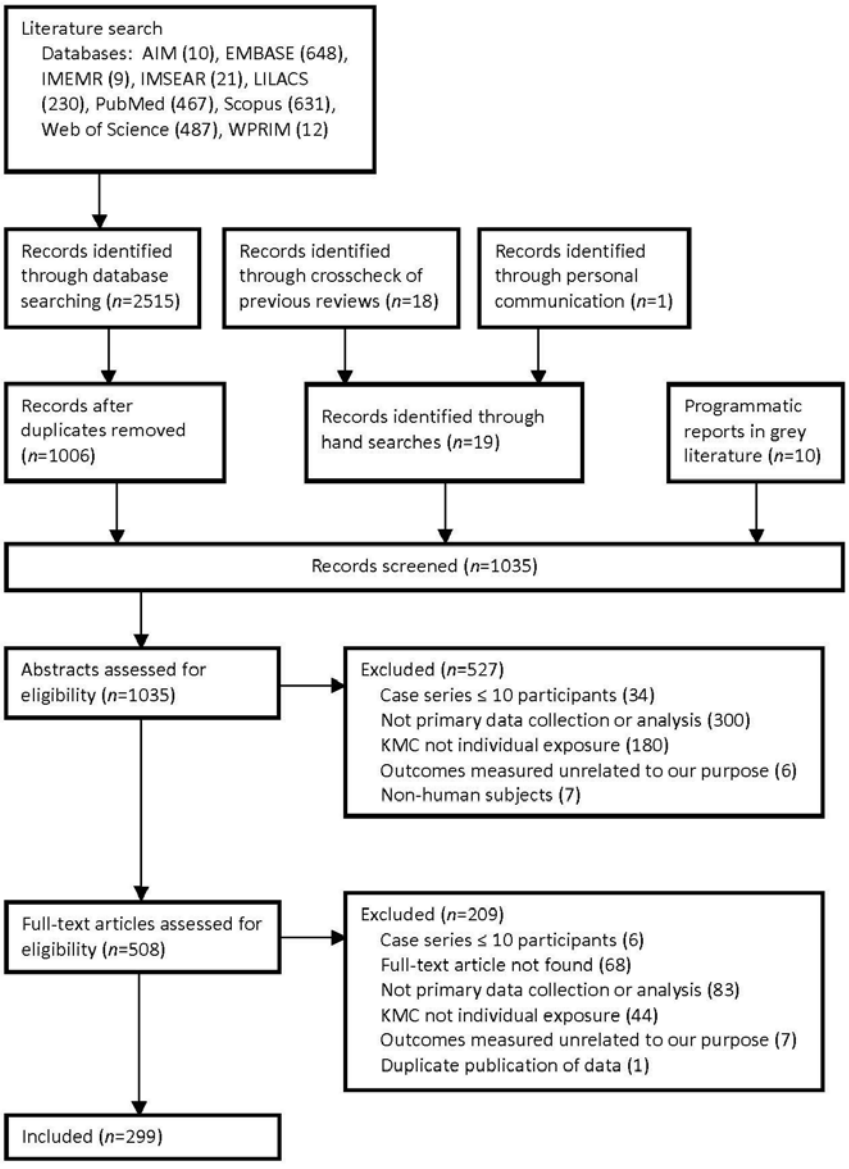


Figure 1. Flow diagram of study selection.

care units or stepdown units, and 11 (4%) were community or population–based.

Most studies included preterm newborns less than 37 weeks gestation (n = 134, 45%), 33 studies (11%) included only full term infants 37 weeks gestation or greater, 28 studies (9%) included newborns of all gestational ages, and 104 studies (35%) did not report gestational ages of the study participants. Similarly, 73 studies (24%) were among low birth weight infants less than 2500 g; 52 studies (17%) included infants less than 2500 g to 1500 g, and 21 (7%) studies were among very low birth weight infants less than 1500 g. Forty–five studies (15%) included a mix of low and very low birth weight newborns. Nine studies (3%) were among newborns weighing 2500 g or greater and 25 studies (8%) included newborns of all birth weights. One hundred forty–six studies (49%) did not describe birth weight characteristics. Forty three studies (14%) reported neither gestational age nor birth weight.

KMC components

The individual components of KMC varied across studies (**Table 2**). Kangaroo mother care was not defined in 88 studies (29%). All 211 studies (71%) with KMC definitions included SSC as a component. One–hundred forty–eight studies (50%) included SSC only. For the additional components, 49 studies (16%) included SSC and exclusive or near–exclusive breastfeeding, 36 (12%) included SSC and follow–up after discharge from the health facility, and 22 (7%) included early discharge from the health facility.

Table 2. Description of kangaroo mother care components in studies

Kangaroo mother care components	N = 299	%
Skin–to–skin contact only	148	49.50
Skin–to–skin contact, breastfeeding	25	8.36
Skin–to–skin contact, breastfeeding, follow–up	16	5.35
Skin–to–skin contact, early discharge, follow–up	13	4.35
Skin–to–skin contact, breastfeeding, early discharge, follow up	7	2.34
Skin–to–skin contact, breastfeeding, early discharge	1	0.33
Skin–to–skin contact, early discharge	1	0.33
Undefined kangaroo mother care	88	29.43

Skin–to–skin contact

Among the studies that defined SSC as part of the KMC package, criteria for SSC initiation, SSC ending, and SSC duration were not well described (**Table 3** and **Table 4**). In 43 studies (14%), SSC was initiated after non–stability criteria were met, 27 studies (9%) promoted immediate initiation of SSC within 60 minutes of birth, 76 studies (25%) encouraged SSC after stability criteria were met, 18 studies (6%) encouraged SSC after a painful procedure, and 135 (45%) did not describe SSC initiation criteria. Forty–three studies observed initiation of SSC of which 4 (9%) observed immediate initiation of SSC. Criteria for stability were non–specific including the terms "clinically stable," "adapted to extra–uterine life," "can tolerate handling," and "without serious illness". More defined criteria included "satisfactory APGAR score," "stable weight," and "stable respiratory and hemodynamic parameters." Criteria to end SSC were largely non–specific with terms "one day or less," "until baby no longer accepts," or "until parent no longer accepts." More specific terms included "until reaches satisfactory weight [2000 grams or 2500 grams]". We compared descriptions of SSC with observations of SSC to differentiate promotion vs practice. Most studies (>85%) did not include data on observations of SSC practiced (**Table 3**).

Data on the duration of SSC are needed to understand the benefits of SSC as well as the feasibility to scale KMC; however this was missing from most studies (**Table 4**). One hundred thirty–two studies (44%) did not describe the number of hours per day SSC was promoted. Seventy–eight studies (26%) encouraged SSC for less than two hours per day, 15 of these studies examined the effect of SSC on painful procedures. Otherwise, the most common duration of SSC promoted was 22 hours or more (n = 46, 15%). Only 37 studies (12%) observed duration of SSC practiced, of which six (2%) observed at least 22 hours per day SSC practiced. SSC duration was also categorized inconsistently as continuous, intermittent, number of hours per session, number of sessions per day, and number of days. Definitions of the term continuous included 24 hours per day, continuous within sessions, or one continuous session but less than 24 hours a day.

Table 3. Promoted skin–to–skin contact characteristics compared to observed skin–to–skin contact characteristics

	Promoted skin–to–skin contact		Observed skin–to–skin contact	
	N	%	N	%
Skin–to–skin contact initiation:				
After stability criteria were met	76	25.42	11	3.68
After non–stability criteria were met	43	14.38	28	9.36
Immediately, regardless of stability	27	9.03	4	1.34
Prior to painful procedure	18	6.02	0	0.00
Undefined or not applicable	135	45.15	256	85.62
Skin–to–skin contact stability criteria:				
Respiratory and/or hemodynamically stable	28	9.36	2	0.67
Clinically stable–not specified further	20	6.69	5	1.67
Adapted to extra–uterine life	8	2.68	0	0.00
Without serious illness	7	2.34	2	0.67
Can tolerate handling	6	2.01	1	0.33
Stable weight	4	1.34	1	0.33
Satisfactory APGAR score	2	0.67	0	0.00
Term	1	0.33	0	0.00
Undefined or not applicable	223	74.58	288	96.32
When was skin–to–skin contact instructed to end?				
One day or less	48	16.05	5	1.67
Until baby no longer accepts	22	7.36	1	0.33
Shortly after painful procedure	13	4.35	0	0.00
After one day and up to two weeks	11	3.68	5	1.67
Until reaches satisfactory weight (2000;3000 g)	10	3.34	5	1.67
After two weeks	8	2.68	5	1.67
Until parent or baby no longer accepts	7	2.34	0	0.00
Until discharge	4	1.34	3.00	1.00
Until parent no longer accepts	4	1.34	0	0.00
Until reached satisfactory health status	3	1.00	0	0.00
Undefined or not applicable	169	56.52	275	91.97

Breastfeeding

Breastfeeding habits were reported in 105 (35%) studies: 38 (13%) reported exclusive breastfeeding, 22 (7%) nearly–exclusive breastfeeding, and 35 (12%) breastfeeding and supplemental feeding (**Table 5**). In most studies, breastfeeding initiation time was not reported (n = 261, 87%). Breastfeeding was started immediately or within one hour of birth in 15 studies (5%), between one and 24 hours after birth in two studies (1%), and 24 hours or longer after birth in five studies (2%). In nine studies (3%) breastfeeding was started at KMC initiation, and seven studies (2%) included physical maturity criteria for initiation of breastfeeding. Seventeen studies (6%) described breastfeeding frequency in their patient population, 13 (4%) studies reported women breastfeeding every two to three hours and four studies (1%) reported women breastfeeding whenever possible.

Table 4. Promoted skin–to–skin contact duration compared to observed skin–to–skin contact duration

	Promoted skin–to–skin contact duration		Observed skin–to–skin contact duration	
	N = 299	%	N = 299	%
Skin–to–skin contact continuous or intermittent within session:				
Continuous within one session	117	39.13	16	5.35
Continuous (24 h per day)	44	14.72	7	2.34
Intermittent (multiple sessions)	26	8.70	17	5.69
Undefined or not applicable	112	37.46	259	86.62
Skin–to–skin contact duration (hours per session):				
1 to 2 sessions	90	30.10	13	4.35
3 to 4 sessions	11	3.68	0	0.00
5 to 8 sessions	2	0.67	0	0.00
≥8 sessions	0	0.00	1	0.33
Undefined or not applicable	196	65.55	285	95.32
Skin–to–skin contact duration (number hours per day):				
<2 h	78	26.09	13	4.35
2 to <4 h	28	9.36	3	1.00
4 to <9 h	13	4.35	8	2.68
9 to <12 h	1	0.33	3	1.00
12 to <22 h	1	0.33	4	1.34
≥22 h	46	15.38	6	2.01
Undefined or not applicable	132	44.15	262	87.63
Skin–to–skin contact duration (number days):				
1 to 5 d	74	24.75	11	3.68
6 to <30 d	19	6.35	8	2.68
≥30 d	5	1.67	1	0.33
Dependent on hospital stay	7	2.34	1	0.33
Undefined or not applicable	194	64.88	278	92.98

Discharge criteria from facility

Fourteen percent of studies (n = 42) described the criteria used for hospital discharge in their study populations (**Table 6**). The most common criteria were clinical stability (n = 19, 6%) or meeting a specified weight gain or weight minimum cutoff (n = 15, 5%). Seven studies (2%) required a combination of adequate weight gain and exclusive breastfeeding prior to discharge. Most studies did not report when infants were discharged (n = 285, 95%). Six studies (2%) reported discharge within seven days of life and eight studies (3%) reported discharge after seven days of life.

Follow–up

Sixty–one studies (20%) described follow–up of infants after discharge, of which 29 studies (48%) followed–up with newborns in health facilities, 22 studies (36%) in homes, and 9 studies (15%) in both facilities and homes (**Table 6**). Follow–up time varied from one month or less (n = 8, 3%) to six to 18 months (n = 13, 4%). Most studies (n = 270, 90%) did not report compliance with follow–up, 11 (4%) reported 90% or higher compliance.

Table 5. Description of breastfeeding characteristics

	N = 299	%
Breastfeeding habits:		
Exclusive	38	12.71
Mixed with other food	35	11.71
Nearly exclusive	22	7.36
Combination	8	2.68
No breastfeeding	2	0.67
Undefined or not applicable	194	64.88
When did breastfeeding start?:		
Immediately or within one hour of delivery	15	5.02
When kangaroo mother care started	9	3.01
Once reached satisfactory degree of physical maturity	7	2.34
One day or more after birth	5	1.67
After one hour but within 24 h of birth	2	0.67
Undefined or not applicable	261	87.29
Breastfeeding frequency:		
Every two to three hours	13	4.35
Whenever possible	4	1.34
Undefined or not applicable	282	94.31

Table 6. Description of discharge and follow–up characteristics

	N = 299	%
Discharge criteria:		
Clinically stable	19	6.35
Adequate weight gain	10	3.34
Exclusively breastfeeding and consistently gaining weight	7	2.34
Absolute weight cutoff	5	1.67
Neonatologist approval	1	0.33
Within time of birth	0	0.00
Undefined or not applicable	257	85.95
Discharge timing:		
After seven days of life	8	2.68
Within seven days of life	6	2.01
Undefined or not applicable	285	95.32
Follow–up location:		
Facility	29	9.70
Home	22	7.36
Facility and home	9	3.01
Phone call or letter	1	0.33
Undefined or not applicable	238	79.60
Follow–up time:		
>3 months to 6 months	11	3.68
>6 months to 12 months	11	3.68
Dependent on adequate weight gain	10	3.34
≤1 months	8	2.68
>1 months to 3 months	8	2.68
Until 40 weeks gestational age	4	1.34
>12 months to 18 months	2	0.67
Undefined or not applicable	245	81.49
Compliance with follow–up:		
70 to <90%	11	3.68
90 to <100%	9	3.01
<70%	7	2.34
100%	2	0.67
Undefined or not applicable	270	90.30

Other components

Studies also described clothing recommendations, newborn positioning, and temperature monitoring during KMC. In 64 studies (21%) participants were instructed to clothe their infant in only a diaper during kangaroo care, an additional 64 studies (21%) encouraged use of a diaper, cap, and socks, and 17 (6%) promoted having the infant naked during SSC contact (**Table 7**). The majority of studies (n = 179, 60%) instructed participants to position the infant prone on the care provider's chest during SSC, while five studies (2%) encouraged a side–lying or breastfeeding position. In 59 studies (20%), the kangaroo care provider was instructed to be in a reclined position, while an upright position was encouraged in 48 studies (16%). Temperature of the infant was monitored during SSC in 71 studies (24%).

DISCUSSION

There is significant heterogeneity in the definition of KMC and a large number of studies did not report a definition of KMC. Of the studies that defined KMC, SSC was present in all studies. Additional KMC components–breastfeeding, early discharge, and follow–up–were missing in the majority of studies. These findings suggest that SSC is accepted in research and programmatic settings as an essential component of KMC, but the other components vary by context, defined as demographic, economic, social, and cultural factors, and newborn characteristics.

The lack of a clear KMC definition and guidance for implementing KMC is a reflection of incomplete evidence. Evidence for KMC is largely based on meta–analyses that combine studies with heterogeneous definitions of KMC and occur in different settings [5,6]. Attempts to stratify the association of KMC on outcomes by KMC components, newborn characteristics (birth weight, gestational age), and high NMR vs low NMR often do not yield statistically significant results because of the limited data available. We do not know the effect of different combinations of KMC components, nor do we understand the feasibility with which each component can be implemented effectively in different contexts. Our study was limited by the lack of data on the duration of SSC. Furthermore, measurement of SSC duration was based on mothers' report of time with minimal observational data. Studies where SSC duration was measured by an independent observer may be biased by the Hawthorn effect.

To define the optimal duration of SSC, we need additional data on the dose response of SSC duration on mortality and morbidity outcomes. The benefits of SSC are likely dependent on the duration of SSC, however the duration of SSC must also be balanced with the feasibility of practicing SSC for extended periods of time. In most settings promoting SSC 24 hours a day is not feasible. Understanding the minimal duration of SSC that provides the maximal benefits will provide more specific recommendations. Most studies initiated KMC after stabilization of the newborn and the effect of KMC on mortality and morbidity is generalizable to the population of newborns who survive to be stabilized. The effect of KMC immediately after birth before stabilization is unclear due to inconclusive evidence [14–17]. Additional efforts to test the effect of KMC prior to stabilization and to define stability is needed through further studies or by consulting experts at each level of care (primary, secondary, or tertiary care) through a Delphi method.

To operationalize KMC, the simpler the intervention the more likely it is to scale [18]. A simple and clear operational definition for KMC is needed. Evidence suggests benefits for newborns less than 2000 g, who are stabilized in facilities with SSC as the primary component. More work is needed to improve the measurement of gestational age and improving the recording of birth weights in facilities to better understand the impact of KMC and for whom there are benefits. Our review suggests that skin–to–skin contact is the core minimal component of KMC and variations depend on context and individual clinical needs of the newborn. For example, extremely preterm newborns who are unable to coordinate their suck and swallow will need feeding support such as nasogastric feeding or intravenous fluid. In high resource settings with space and infection precautions, a provider may recommend SSC for a preterm infant but choose not to discharge early from the facility. To operationalize KMC, a simple matrix that lists newborn characteristics in columns and KMC components in rows for different settings, ie, tertiary, secondary, primary or community levels, can take into account the core SSC com-

Table 7. Description of clothing and positioning during kangaroo mother care

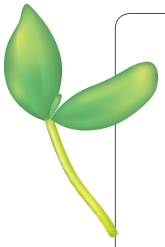
	Promoted clothing and positioning		Observed clothing and positioning	
	N	%	N	%
Clothing of kangarooed baby:				
Diaper or nappy	64	21.40	8	2.68
Diaper, cap, and socks	64	21.40	6	2.01
Naked	17	5.69	2	0.67
Undefined or not applicable	154	51.51	283	94.65
Position of kangarooed baby:				
Prone on mother's chest	179	59.87	17	5.69
On side or next to mother	3	1.00	1	0.33
Breastfeeding position	2	0.67	0	0.00
Undefined or not applicable	115	38.46	281	93.98
Position of provider:				
Inclined or reclined	59	19.73	8	2.68
Upright	48	16.05	5	1.67
Variation of inclined and upright	12	4.01	2	0.67
Undefined or not applicable	180	60.20	284	94.98

ponents with variations based on differences in the newborn and context.

As implementation of KMC begins to accelerate globally, data on the context, individual newborn factors, and KMC components can be collected and harmonized to generate a model that will best define KMC for a set of individual newborn characteristics in specific settings. Research and programmatic agendas to advance KMC should include a standardized set of indicators and measurement tools that document SSC initiation criteria, SSC duration as number of hours per day promoted and ideally observed, feeding protocols, discharge criteria from a facility to community and follow–up standards, and discharge criteria from KMC. To track progress, indicators and standard measurement tools are needed to measure coverage of key newborn interventions including KMC [19]. The release of the new preterm guidelines by the World Health Organization, where KMC is recommended for all newborns less than 2000 g, will provide an opportunity for programs and researchers to start addressing definition gaps, establish global recommendations of operational definitions and core components of KMC, and accelerate KMC within care of preterm babies.

CONCLUSION

Developing a standardized operational definition of KMC and employing indicators and measurement tools to measure and evaluate KMC acceleration efforts is needed. More than half of the studies equate KMC with SSC. Moving forward, careful distinction between KMC and SSC is needed. While SSC is beneficial for all newborns, KMC should be clearly defined, at the bare minimum, as a package of interventions including SSC, exclusive breastfeeding, and close monitoring for preterm and/or low birthweight babies. Researchers and program implementers can contribute to building a more solid evidence base for KMC by measuring and reporting how KMC is defined–the components implemented and the feasibility of implementation based on the context–and the outcomes measured. A central and accessible database to share knowledge should contain this data in addition to standardized indicators, such as the proportion of eligible newborns who receive KMC and the barriers and facilitators to implementation of KMC.



Acknowledgments: We thank Stacie Constantian, Roya Dastjerdi, and Tobi Skotnes for reviewing and abstracting data. Rodrigo Kuromoto and Eduardo Toledo reviewed non–English articles. We acknowledge Kate Lobner for developing and running the search strategy. We would like to thank the mothers and newborns who participated in these studies to better understand how research and programs define KMC.

Funding: Funding for this systematic review was provided by Saving Newborn Lives program of Save the Children Federation, Inc.

Authorship contributions: All authors listed have participated in the concept, design, analysis and interpretation of the data, drafting, or editing of the manuscript. All authors have approved the manuscript as submitted.

Competing interests: The authors completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). None of the authors has any competing interests to declare. There are no competing interests to disclose.

REFERENCES

1 Bhutta ZA, Das JK, Bahl R, Lawn JE, Salam RA, Paul VK, et al. Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? Lancet. 2014;384:347-70. Medline:24853604 doi:10.1016/S0140-6736(14)60792-3

2 Liu L, Oza S, Hogan D, Perin J, Rudan I, Lawn JE, et al. Global, regional, and national causes of child mortality in 2000–13, with projections to inform post–2015 priorities: an updated systematic analysis. Lancet. 2015;385:430-40. Medline:25280870 doi:10.1016/S0140-6736(14)61698-6

3 Birthweight L. Country, Regional, and Global Estimate. New York: UNICEF, 2004.

4 Institute of Medicine Committee on Understanding Premature B. Assuring Healthy O. The National Academies Collection: Reports funded by National Institutes of Health. In: Behrman RE, Butler AS, editors. Preterm Birth: Causes, Consequences, and Prevention. Washington (DC): National Academies Press (US). National Academy of Sciences, 2007.

5 Lawn JE, Mwansa-Kambafwile J, Barros FC, Horta BL, Cousens S. Kangaroo mother care to prevent neonatal deaths due to preterm birth complications. Int J Epidemiol. 2010;39:i144-54. Medline:20348117 doi:10.1093/ije/dyq031

REFERENCES

6 Conde–Agudelo A, Diaz–Rossello JL. Kangaroo mother care to reduce morbidity and mortality in low birthweight infants. Cochrane Database Syst Rev. 2014;4:CD002771. Medline:24752403

7 Dickson KE, Simen–Kapeu A, Kinney MV, Huicho L, Vesel L, Lackritz E, et al. Every Newborn: health–systems bottlenecks and strategies to accelerate scale–up in countries. Lancet. 2014;384:438-54. Medline:24853600 doi:10.1016/S0140-6736(14)60582-1

8 Bahl R, Martines J, Bhandari N, Biloglav Z, Edmond K, Iyengar S, et al. Setting research priorities to reduce global mortality from preterm birth and low birth weight by 2015. J Glob Health. 2012;2:010403. Medline:23198132 doi:10.7189/jogh.01.010403

9 Rudan I, Kapiriri L, Tomlinson M, Balliet M, Cohen B, Chopra M. Evidence–based priority setting for health care and research: tools to support policy in maternal, neonatal, and child health in Africa. PLoS Med. 2010;7:e1000308. Medline:20644640 doi:10.1371/journal.pmed.1000308

10 Engmann C, Wall S, Darmstadt G, Valsangkar B, Claeson M. Consensus on kangaroo mother care acceleration. Lancet. 2013;382:e26-7. Medline:24246562 doi:10.1016/S0140-6736(13)62293-X

11 Bergh AM, Kerber K, Abwao S, de–Graft Johnson J, Aliganyira P, Davy K, et al. Implementing facility–based kangaroo mother care services: lessons from a multi–country study in Africa. BMC Health Serv Res. 2014;14:293. Medline:25001366 doi:10.1186/1472-6963-14-293

12 World Health Organization DoRHaR. Kangaroo mother care: a practical guide. WHO Library. Geneva: World Health Organization, 2003.

13 Bank W. Data: Mortality rate, neonatal (per 1000 live births) 2014. Available: http://data.worldbank.org/indicator/SH.DYN.NMRT. Accessed: 1 October 2014.

14 Ahmed S, Mitra SN, Chowdhury AM, Camacho LL, Winikoff B, Sloan NL. Community Kangaroo Mother Care: implementation and potential for neonatal survival and health in very low–income settings. J Perinatol. 2011;31:361-7. Medline:21311502 doi:10.1038/jp.2010.131

15 Worku B, Kassie A. Kangaroo mother care: a randomized controlled trial on effectiveness of early kangaroo mother care for the low birthweight infants in Addis Ababa, Ethiopia. J Trop Pediatr. 2005;51:93-7. Medline:15840760 doi:10.1093/tropej/fmh085

16 Cattaneo A, Davanzo R, Worku B, Surjono A, Echeverria M, Bedri A, et al. Kangaroo mother care for low birthweight infants: a randomized controlled trial in different settings. Acta Paediatr. 1998;87:976-85. Medline:9764894 doi:10.1111/j.1651-2227.1998.tb01769.x

17 Lincetto O, Nazir AI, Cattaneo A. Kangaroo mother care with limited resources. J Trop Pediatr. 2000;46:293-5. Medline:11077939 doi:10.1093/tropej/46.5.293

18 Atun R, de Jongh T, Secci F, Ohiri K, Adeyi O. Integration of targeted health interventions into health systems: a conceptual framework for analysis. Health Policy Plan. 2010;25:104-11. Medline:19917651 doi:10.1093/heapol/czp055

19 Moran AC, Kerber K, Sitrin D, Guenther T, Morrissey CS, Newby H, et al. Measuring coverage in MNCH: indicators for global tracking of newborn care. PLoS Med. 2013;10:e1001415. Medline:23667335 doi:10.1371/journal.pmed.1001415

(*continued from the outside cover*)

PAPERS

RESEARCH THEME: GLOBAL HEALTH RESEARCH PRIORITIES

PAPERS

Journal of Global Health is published biannually by the Edinburgh University Global Health Society (EUGHS). The aim of the EUGHS is to inspire and educate individuals, groups, organizations and communities about global health issues. Its key objectives are to provide a platform for University of Edinburgh students to share ideas and/or experiences in Global Health with each other; organize meetings and other events to raise awareness of Global Health issues within the University and more widely; and to give students opportunities to present their work. The society is affiliated to the Centre for Population Health Sciences of the University of Edinburgh, and it works closely with the Global Health Academy of the University of Edinburgh.

Subscription (including postage): GBP 30 per any single issue, GBP 50 for annual subscription, GBP 25 for students, GBP 20 for EUGHS members. Please contact JoGH office for details.

Contact for subscription information:
Journal of Global Health
Rachael Atherton
Centre for Global Health Research;
College of Medicine and Veterinary Medicine,
The University of Edinburgh, Edinburgh, Scotland, UK
Edinburgh EH8 9AG, Scotland, UK
Phone +44 131 651 1386 Fax +44 131 650 6909
Rachael.Atherton@ed.ac.uk

Journal design: Snježana Engelman Džafić for LASER*plus*, Zagreb, Croatia
Realisation: LASER*plus,* Zagreb, Croatia, www.laser-plus.hr

(*continued on the inside*)

ISSN 2047-2978