

Vol. 7 No. 2/2  
December 2017  
ISSN 2047-2978

journal of  
**gl****bal**  
health

**Journal of Global Health (JoGH)** is a peer-review 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. All submissions and editorial inquiries should be directed to editor[at]jogh.org.

**Publisher:** 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. 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, while also maintaining collaboration with academic partners throughout the world.

The views expressed by the authors do not reflect the views of JoGH or its publisher.

**Journal design:** Snježana Engelman Džafić for LASERplus, Zagreb, Croatia

**Realisation:** LASERplus, Zagreb, Croatia, www.laser-plus.hr

**Creative Commons Licence:** CC BY 4.0.

Print ISSN: 2047-2978

Online ISSN: 2047-2986

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.

## EDITORS-IN-CHIEF

**Professor Ana Marušić** – Chair, Department of Research in Biomedicine and Health, The University of Split School of Medicine, Split, Croatia

**Professor Igor Rudan** – Chair in International Health and Molecular Medicine; Co-Director, WHO Collaborating Centre for Population Health Research and Training; Co-Director, Centre for Global Health Research, College of Medicine and Veterinary Medicine, The University of Edinburgh, Edinburgh, Scotland, UK

**Professor Harry Campbell** – Professor of Genetic Epidemiology and Public Health; Co-Director, WHO Collaborating Centre for Population Health Research and Training; Co-Director, Centre for Global Health Research, College of Medicine and Veterinary Medicine, The University of Edinburgh, Edinburgh, Scotland, UK

## REGIONAL EDITORS

**North America:** **Professor Robert E. Black** – Director for the Institute of International Programs; Johns Hopkins Bloomberg School of Public Health; Johns Hopkins University, Baltimore, MD, USA • **Professor Rifat Atun** – Professor of Global Health Systems; Director of Global Health Systems Cluster; Harvard T. H. Chan School of Public Health, Harvard University, Cambridge, MA, USA

**South America:** **Professor Claudio F. Lanata** – Nutritional Research Institute; Lima, Peru • **Professor Cesar G. Victora** – Professor of Epidemiology; Federal University of Pelotas; Pelotas, Brazil

**Europe:** **Professor Sue C. Welburn** – Vice-Principal, Global Access; Professor, Medical and Veterinary Molecular Epidemiology; The University of Edinburgh, Edinburgh, Scotland, UK • **Dr Evropi Theodoratou** – Centre for Global Health Research, College of Medicine and Veterinary Medicine, The University of Edinburgh, Edinburgh, Scotland, UK

**Africa:** **Professor Shabir A. Madhi** – Executive Director, National Institute for Communicable Diseases; DST/NRF South African Research Chair in Vaccine Preventable Diseases; University of the Witwatersrand, Johannesburg, South Africa • **Professor Mark R. Tomlinson** – Department of Psychology; Stellenbosch University, Matieland, South Africa

**Asia:** **Professor Zulfiqar A. Bhutta** – The Noordin Noormahomed Sheriff Endowed Professor and Founding Chair; Department of Paediatrics and Child Health; The Aga Khan University, Karachi, Pakistan • **Professor Shams El Arifeen** – Professor of Epidemiology; Director, Centre for Child and Adolescent Health; International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR, B); James P. Grant School of Public Health, BRAC University, Dhaka, Bangladesh

**China:** **Professor Yan Guo** – Professor of Public Health; Vice President, Peking University Health Science Center; Vice Chairman of the Chinese Rural Health Association; Vice Director of the China Academy of Health Policy (PKU); School of Public Health, Peking University, Beijing, China • **Dr Kit Yee Chan** – Visiting Scholar, School of Public Health, Peking University Health Science Center; Peking University, Beijing, China and Reader, Center for Population Health Sciences; College of Medicine and Veterinary Medicine; The University of Edinburgh, Scotland, UK

**India:** **Professor K. Srinath Reddy** – President, Public Health Foundation of India; President, World Heart Federation; Chair, Core Advisory Group on Health and Human Rights, National Human Rights Commission, India; President, National Board of Examinations, India; Bernard Lown Visiting Professor of Cardiovascular Health, Harvard School of Public Health, Boston, MA, USA • **Dr Harish Nair** – Public Health Foundation of India; Indian Institute of Public Health, New Delhi, India; Centre for Global Health Research, College of Medicine and Veterinary Medicine, The University of Edinburgh, Edinburgh, Scotland, UK

**Australia and Western Pacific:** **Professor Trevor Duke** – Director, Centre for International Child Health; Director, WHO Collaborating Centre for Research and Training in Child and

Neonatal Health; Deputy Director of the Intensive Care Unit and Clinical Director of the General Intensive Care Unit; The University of Melbourne and Royal Children's Hospital, Melbourne, Australia • **Professor Wei Wang** – Professor of Postgraduate Medicine; School of Medical Sciences, Edith Cowan University, Perth, Australia

## EDITORIAL COUNCIL

**Professor James Smith** – Vice-Principal International; Professor of African and Development Studies; School of Social and Political Science; The University of Edinburgh, Edinburgh, Scotland, UK • **Jake Broadhurst** – Head, Global Academies Directorate; The University of Edinburgh, Edinburgh, Scotland, UK • **Professor Mickey Chopra** – Honorary Professor of Global Health; Centre for Global Health Research, College of Medicine and Veterinary Medicine, The University of Edinburgh, Edinburgh, Scotland, UK; Lead health Specialist, The World Bank, Washington, DC, USA • **Dr Elizabeth Grant** – Director, Global Health Academy; Senior Lecturer in Global Health and Development; Programme Director, Distance Learning Masters Programmes; Centre for Population Health Sciences; College of Medicine and Veterinary Medicine; The University of Edinburgh, Edinburgh, Scotland, UK • **Professor Devi Sridhar** – Chair in Global Public Health; Centre for Global Health Research, College of Medicine and Veterinary Medicine, The University of Edinburgh, Edinburgh, Scotland, UK and Co-Director, Centre for AIDS Interdisciplinary Research (CAIRO), Blavatnik School of Government, Oxford University, Oxford, UK • **Professor Sarah Cunningham-Burley** – Dean of Molecular, Genetic and Population Health Sciences, Professor of Medical and Family Sociology, The University of Edinburgh, Edinburgh, Scotland, UK • **Professor Aziz Sheikh** – Co-Director of the Centre of Medical Informatics and Professor of Primary Care Research and Development, The University of Edinburgh, Edinburgh, Scotland, UK

## EDITORIAL BOARD

**Dr Davies Adeloje** – Covenant University Ota Ogun State, Nigeria; and Centre for Global Health Research, The University of Edinburgh, Edinburgh, Scotland, UK • **Professor Abdullah H. Baqui** – Director, International Center for Maternal and Newborn Health; The Johns Hopkins Center for Global Health; Johns Hopkins University, Baltimore, MD, USA • **Dr Luciana Brondi** – Centre for Global Health Research, College of Medicine and Veterinary Medicine, The University of Edinburgh, Edinburgh, Scotland, UK • **Dr W. Abdullah Brooks** – Director, Kamalapur Field Site; Head, Infectious Diseases Unit; The Johns Hopkins Center for Global Health; Johns Hopkins University, Baltimore, MD, USA • **Professor Graham Brown** – Foundation Director; Nossal Institute for Global Health; The University of Melbourne, Melbourne, Australia • **Ruth M. Campbell** – Honorary Associate/Researcher; Harvard Global Equity Initiative, Cambridge, MA, USA • **Associate Professor Josip Car** – Director of Health Services and Outcomes Research Programme; Lee Kong Chian School of Medicine; Imperial College Gnanayang Technological University, Singapore and Director, Global eHealth Unit, School of Public Health, The Imperial College London, London, UK • **Professor Simon Cousens** – Professor of Epidemiology and Medical Statistics; Director of Research Degrees; Department of Infectious Disease Epidemiology; London School of Hygiene and Tropical Medicine, London, UK • **Professor Karen Edmond** – Winthrop Professor; School of Paediatrics and Child Health; The University of Western Australia, Perth, Australia • **Dr Christa L. Fischer Walker** – Associate Scientist; Division of Global Disease Epidemiology and Control; The Johns Hopkins Center for Global Health; Johns Hopkins University, Baltimore, MD, USA • **Professor Wendy J. Graham** – Chair in Obstetrics and Gynaecology; The Institute of Applied Sciences; The University of Aberdeen, Aberdeen, UK • **Dr David B. Hipgrave** – Honorary Senior Researcher; Nossal Institute of Global Health; University of Melbourne, Melbourne, Australia and UNICEF, New York, NY, USA • **Dr Hope L. Johnson** – Center for Global Health; Johns Hopkins University Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA and Head, Programme Outcomes and Impact, GAVI Alliance, Geneva, Switzerland • **Professor Keith P. Klugman** – William H. Foege Chair of Global Health; Hubert Department of Global Health, Rollins School of Public Health; The Emory University, Atlanta, GA, USA • **Assistant Professor Ivana Kolčić** – Assistant Professor

of Epidemiology; Croatian Centre for Global Health; Department of Public Health; The University of Split School of Medicine, Split, Croatia • **Assistant Professor Tomislav Lauc** – Assistant Professor of Orthodontics and Dental Anthropology; Editor-in-Chief, South European Journal of Orthodontics and Dentofacial Research; Department of Anthropology, School of Social Sciences and Humanities; University of Zagreb, Zagreb, Croatia • **Professor Joy E. Lawn** – Director, Global Evidence and Policy, Saving Newborn Lives programme, Save the Children; Director, MARCH Centre; The London School of Hygiene and Tropical Medicine, London, UK • **Clinical Assistant Professor James A. Litch** – Director, Perinatal Interventions Program; Global Alliance to Prevent Prematurity and Stillbirth (GAPPS), Seattle Children's; Department of Global Health and Department of Epidemiology; University of Washington, Seattle, WA, USA • **Dr Li Liu** – Assistant Scientist; Division of Global Disease Epidemiology and Control; The Johns Hopkins Center for Global Health; Johns Hopkins University, Baltimore, MD, USA • **Professor Azeem Majeed** – Head, Department of Primary Care & Public Health; Imperial College London, UK • **Professor Julie Meeks Gardner** – Head, Caribbean Child Development Centre; Consortium for Social Research and Development; The University of the West Indies, Mona, Jamaica • **Professor Katherine L. O'Brien** – Executive Director, The International Vaccine Access Center (IVAC); Global Disease Epidemiology and Control, Center for Global Health; Johns Hopkins University Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA • **Associate Professor Ozren Poljšak** – Associate Professor of Biostatistics; Director, Croatian Centre for Global Health; Department of Public Health; The University of Split School of Medicine, Split, Croatia • **Professor Carine Ronsmans** – Professor of Epidemiology; Head, Department of Infectious Diseases Epidemiology, The London School of Hygiene and Tropical Medicine, London, UK • **Professor Craig E. Rubens** – Affiliate Professor of Global Health and Affiliate Professor of Pediatrics; Executive Director, Global Alliance to Prevent Prematurity and Stillbirth (GAPPS); Seattle Children's and University of Washington School of Medicine, Seattle, WA, USA • **Professor Samir K. Saha** – Head, Department of Microbiology; Executive Director, The Child Health Research Foundation; Bangladesh Institute of Child Health, Dhaka Shishu Hospital, Dhaka, Bangladesh • **Professor Joshua A. Salomon** – Professor of Global Health; Department of Global Health and Population; Harvard School of Public Health and Harvard Global Health Institute; Harvard University, Boston, MA, USA

• **Dr Giorgio Tamburlini** – Research Director, Institute of Child Health IRCCS Burlo Garofolo; Scientific Director, European School for Maternal, Newborn, Child and Adolescent Health; Director, WHO Collaborating Centre on Maternal and Child Health; The Institute Burlo Garofolo, Trieste, Italy • **Assistant Professor Alexander C. Tsai** – Assistant Professor of Psychiatry; Department of Psychiatry; Massachusetts General Hospital and Harvard Medical School; Boston, USA • **Professor Peter Tugwell** – Professor of Medicine and Epidemiology & Community Medicine; co-Director, WHO Collaborating Centre for Knowledge Translation and Health Technology Assessment in Health Equity; Director, Centre for Global Health; Institute of Population Health and Department of Medicine; Faculty of Medicine, The University of Ottawa, Ottawa, Canada • **Associate Professor Anita K. Zaidi** – Director of Research, Department of Paediatrics and Microbiology; The Aga Khan University, Karachi, Pakistan and Director, Enteric and Diarrhoeal Diseases, Bill and Melinda Gates Foundation, Seattle, WA, USA • **Associate Professor Lina Zgaga** – Associate Professor of Epidemiology; Department of Public Health and Primary Care; Trinity College Centre for Health Sciences, Dublin, Ireland • **Professor Stanley H. Zlotkin** – Professor, Departments of Paediatrics, Nutritional Sciences and Dalla Lana School of Public Health Vice President, Medical and Academic Affairs, Department of Pediatrics and Surgery, Division of Clinical Medicine; The Hospital for Sick Children, The University of Toronto, Toronto, Canada • **Dr Henry Kalter** – Institute of International Programs; Johns Hopkins Bloomberg School of Public Health; Johns Hopkins University, Baltimore, MD, USA

## MANAGING EDITOR/NEWS SECTION EDITOR

**Rachael Atherton** – Centre for Global Health Research; College of Medicine and Veterinary Medicine, The University of Edinburgh, Edinburgh, Scotland, UK

# Journal of Global Health: The Mission Statement

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, policy-makers 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*



Children suffering from diarrhoea in intensive care unit at the hospital in Dhaka adjunct to International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR, B). At the beginning of the 21<sup>st</sup> century diarrhoea was a major cause of child mortality, causing nearly 2 million child deaths each year world-wide. Since then, a broad distribution of oral rehydration sachets (ORS), much improved school enrolment for girls, and widespread maternal education on symptom recognition and intervention reduced this burden to under 500 000 deaths per year. Still, more should be done and it is expected that a roll-out of rotavirus vaccine may further reduce this unacceptable death toll.

# “Crowdsourcing” ten years in: A review

Kerri Wazny

Centre for Global Health, Usher Institute for Informatics and Population Sciences, University of Edinburgh, Edinburgh, Scotland, UK

**Background** First coined by Howe in 2006, the field of crowdsourcing has grown exponentially. Despite its growth and its transcendence across many fields, the definition of crowdsourcing has still not been agreed upon, and examples are poorly indexed in peer-reviewed literature. Many examples of crowdsourcing have not been scaled-up past the pilot phase. In spite of this, crowdsourcing has great potential, especially in global health where resources are lacking. This narrative review seeks to review both indexed and grey crowdsourcing literature broadly in order to explore the current state of the field.

**Methods** This is a review of reviews of crowdsourcing. Semantic searches were conducted using Google Scholar rather than indexed databases due to poor indexing of the topic. 996 articles were retrieved, of which 69 were initially identified as being reviews or theoretically-based. 21 of these were found to be irrelevant and 48 articles were reviewed.

**Results** This narrative review focuses on defining crowdsourcing, taxonomies of crowdsourcing, who constitutes the crowd, research that is amenable to crowdsourcing, regulatory and ethical aspects of crowdsourcing and some notable examples of crowdsourcing.

**Conclusions** Crowdsourcing has the potential to be hugely promising, especially in global health, due to its ability to collect information rapidly, inexpensively and accurately. Rigorous ethical and regulatory controls are needed to ensure data are collected and analysed appropriately and crowdsourcing should be considered complementary to traditional research methods.

“No one knows everything, everyone knows something [and] all knowledge resides in humanity; digitalisation and communication technologies must become central in this coordination of far flung genius” [1]. Although examples of crowdsourcing and “wisdom of the crowds” have been reported hundreds of years ago [2,3], the term “crowdsourcing” was coined in 2006 by Howe in his *Wired* magazine article [4]. In the article, Howe defines crowdsourcing as “the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call” [4] and he adds that “crowdsourcing is the mechanism by which talent and knowledge is matched to those in need of it” to the definition in a later article [5]. Since Howe’s article and partially due to the availability of modern technology [6,7], use of crowdsourcing has skyrocketed [8]. Although research in this area has grown exponentially in the last decade, many authors feel that the potential of crowdsourcing is still underutilised and underexploited [5,9–11].

## Correspondence to:

Kerri Wazny  
 Centre for Global Health Research  
 Usher Institute of Population Health Sciences and Informatics  
 University of Edinburgh  
 Old Medical School  
 Teviot Place  
 Edinburgh EH8 9AG  
 United Kingdom  
 kerri.wazny@ed.ac.uk

As crowdsourcing requires, depending on the definition, ‘outsourcing’ a task or tasks to a large crowd [12], advances in technology have facilitated the efficiency of this method [2,6,13–15]. Indeed, research that was previously inconceivable due to the scale is now achievable through crowdsourcing [6]. Kama-jian states that 35% of smart phone users check their phones prior to getting out of bed and, as of 2013, over 5 billion people worldwide had access to mobile phones [14,16]. Prior to Howe’s *Wired* article, Louis van Ahn introduced the idea of human computing, where humans are used to solve complex problems that computers are not capable of [17]. While machine learning has made great strides, computers are poor at perception; humans can conceptualise, discriminate and filter, learn, adapt using their background knowledge and apply common sense and experience that machine are unable to do [18]. In addition to humans actively crowdsourcing data, ubiquitous computing, where computers exist through the physical environment, are virtually invisible to the user, and act as passive sensors has great potential for generating large amounts of data [16,18]. Cell phones, for example, can collect photo, video, acoustic, gyroscopic, accelerometer and proximal information and can also be used to add pairing devices to collect additional information, such as pollution sensors [16]. Crowdsourced spatial analysis from GIS data can be very useful, especially for providing resources in emergency situations, for delivering logistics and for efficient targeting of interventions [19].

As individuals are biased towards the correct answer, Buecheler et al. estimate that if a million individuals were to contribute towards answering a problem via crowdsourcing, there would be a 97.7% likelihood that the crowd would arrive at the correct answer [20]. While pilot studies have not reached sample sizes close to that scale, many have had great success in achieving extremely promising results. For example, crowdsourcing has been demonstrated to produce accurate results across a range of medical diagnostic studies, including malaria, grading images for glaucoma and diabetic retinopathy, skin self-examination for skin cancers, and images for cancer polyps [21–26].

Despite the interest in the area of crowdsourcing exploding in the past decade, many authors do not agree on its definition or on what counts as crowdsourcing, with some academics considering Wikipedia a “classic” example of crowdsourcing, for example, while others insist it is not crowdsourcing [12]. Text and data mining is another example that is on the fringe of crowdsourcing’s definition.

In addition to there being many definitions of crowdsourcing, many authors have offered different taxonomies of crowdsourcing, some focusing on types of crowdsourcing while others focus on its production model. Furthermore, there are debates on who participates in crowdsourcing – whether it is laypersons, amateurs, professionals, experts, or a combination.

Although crowdsourcing has existed for decades, it is agreed upon that technology has facilitated its growth. Platforms such as Amazon Mechanical Turk and Crowdfunder enable companies to hire workers to perform crowdsourced exercises for extremely low prices. Other crowdsourcing platforms, such as Innocentive or Crowdmed, offer a competitive winner-takes-all model. Sensors in wearable technology have also facilitated the ability to collect mass amounts of information.

Crowdsourcing can increase the accuracy of computer automated tasks, lower costs, increase the scale of research, transcend boundaries and borders, produce novel discoveries and increase the speed of research progression, among other benefits. However, there are concerns with the generalisability of the samples, as the crowd is self-selected, security and data protection issues of sensitive data, and the possibility of malicious workers. Some studies have added quality protection measures to weed out malicious workers, such as adding cut-offs for scores on previous tasks and screening questions. Additional regulation is needed for ethical issues, such as obtaining informed consent and data use policies.

Crowdsourcing has considerable benefits in research, as it has the potential to substantially lower costs while massively increasing the sample size and researchers can receive the data in real-time [7,16,19,27–29]. Because of these qualities, crowdsourcing has potential to improve global health research. Indeed, crowdsourcing is used frequently to set research priorities in global health, most often in maternal, newborn and child health, due to the popularity of the Child Health and Nutrition Research Initiative’s (CHNRI) method of research priority setting which uses collective opinion to identify and score research priorities against a set list of criteria [30]. The CHNRI method is becoming the most frequently used research priority setting method due to its transparent, systematic nature; it was designed to capitalise on the principles of Surowiecki’s “Wisdom of the Crowd,” which will be described in the further in the paper [31]. Furthermore, research in global health faces an even larger burden than research in high-income countries with regards to funding, logistics, poor existing health care systems, health care workers to collect data, equipment, and patient access to health care, especially in rural or conflict areas [21,32–37]. As access to mobile phones in low- and middle-income countries is still increasing, crowdsourcing may

provide a complementary route of data collection to traditional sources, capitalising on structures and knowledge already in place in the countries [38].

## METHODS

As previous authors had reported few search results in indexed journals [10,27,39] and crowdsourcing is a new method, semantic searches in Google Scholar were used to retrieve both peer-reviewed and grey literature published on crowdsourcing. “Crowdsourcing” as well as ‘crowdsourcing’ joint with health terms, such as genetics, diagnosis, epidemiology, surveillance, public health and disease were searched in August, 2015. Crowdsourcing and global health was searched initially, as well, but the results overlapped entirely with crowdsourcing and health and crowdsourcing and public health. The titles of results were scanned until it was clear that results appearing were no longer relevant. Full details of the searches, as well as the number of pages of Google Scholar results scanned, can be found in **Box 1**. In total, 995 results were identified through the Google Scholar search, which is substantially more than any other reviews have identified. 375 results were discarded as duplicates or irrelevant once abstracts were read.

Results were organised within Endnote into categories, including reviews, theory of crowdsourcing, health, public planning, GPS-related, translation, robotics, visual perception, logistics of crowdsourcing, which was broken down into motivations, quality, reliability, stability, and others. This review reports on the papers reporting on reviews and theory as well as a portion of the health-related papers, as there were 285 health papers and many of their interventions overlapped. Further reviews can be conducted with the results of the search and organised Endnote library, but are outside the scope of the current review.

The reviews and theoretical papers generally covered the varying definitions of crowdsourcing, taxonomies of crowdsourcing, participants, modes of participation, when research is suitable for crowdsourcing, benefits and concerns with crowdsourcing, recommendations for regulation and quality control, including ethical regulations and examples of crowdsourcing.

## DISCUSSION

### Defining crowdsourcing

The definition of crowdsourcing as well as some ‘traditional’ examples of crowdsourcing, such as Wikipedia, are highly debated; this is likely due to both the relative newness of the term and the flexibility and adaptability of the method [1,5,7,8,10–12,20,40–43]. To further complicate authors’ attempts to define crowdsourcing, there are a variety of related concepts that have been used synonymously, including: citizen science, health 2.0, wisdom of the crowds, peer production, open sourcing, expert sourcing, collective intelligence, human computation, community-based participatory research, participatory epidemiology, outsourcing and open sourcing [1,3,7,12,43]. While some, like expert-sourcing, are easy to understand as crowdsourcing with experts, the differences between crowdsourcing and others are more nuanced.

Three terms, specifically, are used abundantly in literature and often interchangeably with crowdsourcing: health 2.0, wisdom of the crowds, and citizen science. While applications of crowdsourcing are often a combination of these, especially in the field of health, there are important distinctions between them [5,8,11].

Swan (2012) defines citizen science as non-professionals conducting science-related activities [8]. Non-professionals can include scientists of professionals who are conducting activities outside their own fields (so that they are amateurs in that field). All of the examples given by Swan include citizen science at a mass-scale, and thus are all citizen-science activities that are also using crowds [8]. It may be possible to imagine an activity in which citizens are acting as scientists, collecting data or participating in an experiment that is not at a mass scale, however, such as if citizens provide feedback in the design of a study at a small-scale. Therefore, not all citizen science must be crowdsourcing, but much of it will be.

Health 2.0 is defined, also by Swan, as active participation in one’s health care using web 2.0 technologies [8]. This could include using m-Health applications to track diet and exercise, for example. Using these applications itself would not be considered crowdsourcing, as data are not necessarily collected and there is no unified output. However, if data were collected, the act of collecting data from this could be considered crowdsourcing. Thus, health 2.0 technology can contribute towards crowdsourcing but is not necessarily crowdsourcing.

“Wisdom of the crowds” is another related term. This refers to the use of *knowledge* of a large crowd of people and also requires an intelligent crowd. This also differs, slightly, from the term crowdsourcing, as not all crowdsourcing tasks require knowledge or intelligence. Unlike citizen science and health 2.0, all ‘wisdom of the crowds’ tasks are forms of crowdsourcing, but not all crowdsourcing are necessarily applications of a ‘wisdom of the crowds.’ An example of a task requiring intelligence would be using a crowd to diagnose malaria cells in blood smears. In this, each participant needs to use their knowledge or intelligence to consider which blood smears contain or do not contain malaria parasites. Some, perhaps arguable, examples of crowdsourcing that would *not* be considered requiring knowledge could be RECAPTCHA, passive surveillance such as environmental surveillance using ubiquitous computing and mobile phones, reporting systems, or text mining. In his book, Surowiecki lists four requirements for an intelligent crowd that are particularly important for crowdsourcing tasks that require knowledge (ie, are ‘wisdom of the crowds’ tasks). They are: (i) diversity, which adds perspectives that would otherwise be absent; (ii) independence, limiting the influence of one person’s opinions on other’s; (iii) decentralisation, to develop tacit, specialised knowledge; and (iv) aggregation, to combine the diverse, independent, knowledgeable opinions of the crowd [31].

In addition to these three terms, crowdsourcing is often contrasted to open sourcing or outsourcing. Although some authors believe that crowdsourcing is a special form of outsourcing [3], many authors conclude that the major difference between crowdsourcing and outsourcing is the presence of a contract [10]. In addition, in a crowdsourcing exercise, the organisation or crowdsourcing initiator has the rights to whatever is produced and the crowd is aware of this [10]. Intellectual property rights are also one of the major differences between crowdsourcing and open sourcing or peer-production, along with the hierarchical structure of crowdsourcing [1,10]. In open-sourcing or peer-production, the product that is being worked on is free, will remain free and the crowd that is working on it volunteers their labour to make the free product better. In crowdsourcing, the crowd is volunteering but, if they are contributing to a product, it is unlikely to be available for free [1]. Furthermore, with open-sourced and peer-production models, which are usually software, the software and its code are released and coders work and submit bug fixers as they come up, with no hierarchy. With crowdsourcing, there is a clear call for work.

Crowdsourcing has other key features including a clear, open call for participants and a large crowd. Since there are many different definitions, Estelles-Arolas et al. reviewed definitions of crowdsourcing and developed an integrative definition using Tatarikiewicz’s approach, which is based on developing a global definition of the concept of art. In their review, the authors found 8 key qualities of a crowdsourcing definition, namely: a) who forms the crowd; b) what the crowd has to do; c) how the crowd is reimbursed; d) who initiates the crowdsourcing process; e) what the product of crowdsourcing is; f) what type of process is used; g) what type of call is used; and h) by what medium the call is made [12].

The integrative definition that the authors devise from their review is [12]:

“Crowdsourcing is a type of participative online activity in which an individual, and institution, a non-profit organisation, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and utilize to their advantage what the user has brought to the venture, whose form will depend on the type of activity undertaken.”

Although he describes it as a taxonomy rather than a definition, the features of Geiger et al.’s description of crowdsourcing is similar to Estelles-Arolas et al.’s integrative definition. The key features Geiger describes are: (i) pre-selection of contributors (how ‘open’ the call is, but usually the authors state there are no limits); (ii) accessibility of peer contributors (whether they can access each other’s contributions); (iii) aggregation (to what extent the input is used); and (iv) remuneration (fixed, success-based or none) [12,44].

Each feature of Estelles-Arolas et al.’s integrative definition and Geiger et al.’s taxonomy is discussed below.

## Estelles-Arolas et al.’s and Geiger et al.’s

### *Who forms the crowd (corresponds to Geiger et al.’s (i) pre-selection of contributors)*

The majority of authors reviewed by Estelles-Arolas et al. did not provide a distinct definition for their crowds, instead describing a crowd as a large group of people or individuals, consumers, or volunteers

[12]. The authors found that the crowds size could vary from a few thousand to several hundred thousand and their skill levels could also vary from being very unskilled, in the case of Amazon Mechanical Turk (AMT) workers to extremely skilled InnoCentive submitters, who are often hold PhDs [12]. However, in a study in business management, ideas generated by professionals and laypeople through crowdsourcing were compared and those by laypeople were more novel and offered the customer more benefits; however they were less feasible [14].

In contrast, Brabham specifically examined how authors refer to crowds and found that the majority of articles refer to crowds as being composed of amateurs [7]. However, he argues that ‘amateurism’ in crowdsourcing is a myth, and blames this partially on Howe’s original definition of crowdsourcing. In Brabham’s review, he found that most crowds were comprised of self-selected professionals, such as InnoCentive’s submitters being extremely well-educated, those who submitted advertisements for Doritos’ SuperBowl advertisement contest were mostly film school students and the majority of iStock Photos’ submitters are professional photographers [7]. Is amateurism not being paid or lacking access to tools? Brabham cites Stebbins’ definition of amateurs: “amateurs are guided by standards of excellence set by professionals and not necessarily inferior, feel an obligation to their pursuit, restrain professions from over-emphasising technique and from stressing superficialities instead of meaningful or profound work or products” [7]. He contrasts this definition with a definition of amateurs as “one lacking experience” and further argues that professionalism is a class about status and linked to capitalism. Crowdsourcing, then, represents the ‘race to the bottom’ to allow greater profit margins by falsely positioning who should be described as professionals as amateurs and underpaying them for their work [7].

With regards to the demographics of the crowd, Ranard et al.’s review found that few articles reported on demographics and for those that did, the level of demographics reported varied [27]. However, Khare et al. state that the crowd should be poorly defined and diverse [3]. Brabham believes that there are three types of diversity necessary: (i) identity; (ii) skills; and (iii) political investment. However, his vision of identity includes national, sex, gender, race, economic class, disability, religion, among other things [7]. As Surowiecki stated, diversity is important to having a wise crowd [14,31]. Kamajian found that technical and ‘social marginality’ were beneficial for success in InnoCentive submissions; social marginality was defined as being female [14].

Geiger et al. aim to classify different types of crowdsourcing processes, and in doing so describe the ‘openness’ of their calls. The authors found that most crowdsourcing processes have a completely open call but some restrict contributions from participants by using either qualification-based (ie, the contributors need to have demonstrated a certain level of qualification or skills prior to participating) or context-based (ie, the participants need to be in a certain demographic) limitations [44].

### ***What the crowd has to do***

Estelles-Arolas et al.’s review came across a dichotomy regarding the purpose of the crowd; one group of authors believed that the purpose of the crowd was to complete tasks and the other, to solve problems [12]. Some authors believe that tasks must be divisible into lower-level tasks in order to be suitable for crowdsourcing [5,12,28]. Estelles-Arolas et al. conclude that “any non-trivial problem can benefit from crowdsourcing” [12].

In this review, various authors attempted to make classifications of *what* crowdsourcing should aim to do. These are found in **Table 1**. As one can see, some authors disagree that open innovation and peer production fall outside the realm of crowdsourcing. The authors also differ with regards to the level of detail of their classifications, ranging from a dichotomous classification of microtasks and megatasks [3] to Geiger et al.’s and Saxton et al.’s more detailed classifications of types of crowdsourcing processes [11,44]. However, at its heart, many of the classifications can be conflated to combination of Geiger et al.’s second and Aitamurto et al.’s classifications: crowd creation; crowd voting (including prediction markets); crowd processing; crowd rating; crowd solving; and crowd funding. However, crowd funding is the mobilisation of monetary funds for a common goal and thus is not covered by this review [44,45].

### ***How the crowd is reimbursed (corresponds to Geiger et al.’s remuneration)***

Many of the authors in Estelles-Arola et al.’s review identified reimbursement as monetary reimbursement. The range of monetary reimbursement is large, varying from US\$ 0.01 for each human intelligence task (HIT) performed on the AMT platform to millions of dollars for the successful solution chosen from InnoCentive’s competitions [12]. Geiger et al. look at whether reimbursement is fixed, varied or voluntary as a means to classify crowdsourcing projects. AMT projects would have fixed reimbursements, where

**Table 1.** Classifications of what crowdsourcing should aim to do

Prpic [39]	Public health; health promotion; health maintenance; and health research
Kamajian [14]	Collection; collaboration; individual decisions; and group decisions
Brabham et al. [9]	Knowledge discovery and management (gathering, organising and reporting); distributed human intelligence tasking problems (ie, AMT); broadcast search (ie, locating a needle in a haystack, Innocentive); and, peer-vetted creative problem production (ie, Threadless)
Aitamurto et al. [45] (citing Howe)	Crowd wisdom; crowd creation; crowd voting (including prediction markets); and crowd funding
Saxton et al. [11]	Intermediate (find, finish and earn through the web, ie, AMT); citizen media production; collaborative software development; digital goods sales (ie, iStock Photo); peer-to-peer social financing (ie, Kickstarter); product design (ie, Threadless); consumer reporting; knowledge base building; collaborative science projects
Khare et al. [3]	Microtasks (dissegregated then joined); and mega-tasks (“open innovation”)
Parvanta et al. [46]	Crowdfunding; crowd labour; and crowd research
Pedersen et al. [47]	Co-creation; crowd creation; crowd voting; crowd wisdom; and crowd funding
Yuen et al. [17]	Voting systems; information sharing; games; and creative
Gaiger et al. (a) [44]	Integrative sourcing without remuneration (ie, Wikis); selective sourcing without crowd assessment (ie, private contributors, public design or innovation contests); selective sourcing with crowd assessment (contests where the public assesses contributions); integrative sourcing with success-based remuneration (ie, iStock Photo); and integrative sourcing with fixed remuneration
Geiger et al. (b) [44]	Crowd processing (ie, GalaxyZoo); crowd rating (ie, Trip Advisor); crowd solving (ie, FoldIt); and crowd creation (ie, Threadless)

all members of the crowd are remunerated the same amount for their participation, whereas InnoCentive employs a success-based remuneration plan [44]. However, both Estelles-Arolas et al. and Geiger et al. acknowledge that not all crowdsourcing projects pay monetarily, and that monetary remuneration is not necessarily the primary motivation for the participants. Estelles-Arolas et al.'s review suggests that participant motivations mirror Maslow's hierarchy of individual needs: economic reward, social recognition, self-esteem and development of individual skills. In addition to or in lieu of financial rewards, individuals participating in crowdsourcing are able to develop their skills through freelancing, contribute to their community, have fun, share knowledge and be recognised through their contributions, Parvanta et al. describe the motivations as the 'four fs': fun, fulfilment, fame, and fortune [46]. In addition to these, crowdsourcing activities such as RECAPTCHA have capitalised on task being integral to another task the user is trying to access and have been wildly successful [6]. An additional, similar, motivation that Swan identifies is biocitizenry, in which the crowd participates in order to gain access to studies [8]. Doan and colleagues suggest that, in addition to those listed above, making users pay for a service, providing ownership situations or requiring contribution to crowdsourcing through their employment, having instant gratification or providing an enjoyable experience of a necessary service will motivate a crowd [43]. In their review, Zhao and Zhu found that only 2/55 studies used motivational theories in designing their interventions [10]. Zhao and Zhu, Kostkova, and Kittur call for further research into crowd motivation, specifically the use of serious gaming, auction bidding and understanding crowd behaviour in task selection [10,38,48].

Some authors reviewed mentioned inequities regarding crowd contributions. Parvanta et al. describe a 90%/9%/1% rule for participation, in which 90% of the crowd observes, 9% participates from time-to-time and 1% participates regularly [46]. This breakdown would be more amenable to a service such as YouTube or Wikipedia, where observing or viewing a product is an option. Zhao and Zhu, describe super contributors, contributors and outliers but do not give a percentage of contributions between the three categories [10]. Holley states that the majority of work is completed by 10% of the crowd and these super contributors are often retirees or young, dynamic professionals [49].

### **Who initiates the crowdsourcing process**

Generally, an institution or organisation initiates the crowdsourcing process with an open call [12]. However, there have also been instances where the crowdsourcer has been a governmental department, such as in Iceland [42].

### **What the product of crowdsourcing is**

Many authors reviewed by Estelles-Arolas et al. felt that the initiator receives the result sought for the task advertised, which was usually the result for a given problem. Others believed the product was either

knowledge, ideas, or some type of added value [12]. The exact type product of crowdsourcing can be very diverse and has not been agreed upon, but generally is some type of result that is requested by and has value to the initiator.

### ***What type of process is used***

Estelles–Arolas et al.'s review found many authors who identified crowdsourcing as an outsourcing process, specifically referring to AMT while others referred to it as a problem–solving process or a production model [12]. As described previously, crowdsourcing differs from open sourcing, outsourcing and peer–production. Many articles reviewed in this review specifically mentioned the use of online, outsourcing–like mediums, such as AMT [3,6,48] and CrowdFlower [6]. In AMT and CrowdFlower, the initiator (or crowdsourcer) posts a task and the ‘crowd’ responds and are paid in small quantities for completing small HITs. Other online platforms use distributed online processes to compete for the best solution, such as InnoCentive or CrowdMed [27,32]. Advances in mHealth, such as wearable technologies and sensors, could enable real–time data collection and monitoring from mass amounts of people [38]. Kostkova estimates that 75 million wearable technological devices will have been shipped by 2018 and calls for behavioural research using these devices [38]. The data from these devices could be considered crowdsourcing if there is a specific call for data. Gamification has also been used to enhance the crowd's experience while crowdsourcing and encourage participation [21,50,51]. Finally, another debatable form of crowdsourcing could be data mining, using Twitter posts or Google Flu Trends [32,52,53]. However, according to the definitions of crowdsourcing by both Estelles–Arolas et al. and Geiger et al., data mining would not be in the realm of crowdsourcing.

### ***What type of call is used***

The majority of authors reviewed by Estelles–Arolas et al. refer to an open call as the form of call that must be made in order to satisfy a crowdsourcing criterion. However, Estelles–Arolas et al. disagree and use the term ‘flexible open call’ meaning that participation is non–discriminatory but the call is tailored to the specific initiative and thus, can be limited to a community where there is specific knowledge or expertise (but anyone in this community can answer) [12].

### ***By what medium the call is made***

Estelles–Arolas et al.'s state that of the authors they reviewed, the medium the call was made through was unanimously agreed upon to be the Internet and Estelles–Arolas et al. agreed [12]. However, as stated previously, crowdsourcing has existed prior to the Internet, as has wisdom of the crowds. Thus, while the Internet has enabled crowdsourcing to be used much more effectively and efficiently, it is not necessarily reliant on the Internet as a medium and could be used over a different medium, though this would be less efficient.

#### **Box 1. Crowdsourcing Semantic Searches Conducted in Google Scholar**

1. “Crowdsourcing”  
a. Up to 25 pages
2. “Crowdsourcing” and “Health”  
a. Up to 15 pages
3. “Crowdsourcing” and “Immunology”  
a. Up to 5 pages
4. “Crowdsourcing” and “Genetics”  
a. Up to 9 pages
5. “Crowdsourcing” and “Public Health”  
a. Up to 20 pages
6. “Crowdsourcing” and “Disease”  
a. Up to 25 pages
7. “Crowdsourcing” and “Surveillance”  
a. To 20 pages
8. “Crowdsourcing” and “Diagnosis”  
a. Up to page 14

### **Geiger et al.'s taxonomy/features of crowdsourcing**

#### ***Accessibility of peer contributors***

Geiger et al. discuss the degree of which the crowd is able to access each other's contributions to the product as a feature of the crowdsourcing process and have four categories: none, view, assess or modify [44]. In some crowdsourcing activities, members of the crowd cannot view each other's contributions at all, while others use a crowd not only to for submissions but also to judge which submissions are the best (ie, Threadless). In other crowdsourcing exercises, participants can modify each other's submissions. For example, Kittur posted a Spanish poem for translation through crowdsourcing and the crowd was able to interact with each other, discuss possible translations and together, the crowd submitted a final, translated poem. The authors found this translation to be better than the commonly accepted English translation [48]. Finally, Geiger et al. found that some crowdsourcing projects allow the crowd to view other submissions prior to submitting their own [44].

## Aggregation

Aggregation refers to how the responses of the crowd are used by the crowdsourcer. The two major ways the responses can be used are to be combined or to be selected [44]. InnoCentive, Threadless, and Crowd-Med, for example, are *selective* crowdsourcing companies, which choose the best solution or design to a particular problem. Crowdsourcing projects run on AMT often aggregate or combine solutions from the crowd as a whole.

The definition from Estelles–Arolas et al. excludes Wikipedia, YouTube and Flickr. Wikipedia is excluded on the grounds that there is no initiator (crowdsourcing organisation), that the authors do not feel that the initiator receives benefit, and that there is no open call. YouTube is excluded on the grounds that there is no clear goal, that the crowdsourcer's benefit is not clear, which is arguable as YouTube 'stars' receive compensation for views, that there is no clear initiator, the initiator's benefit is not clearly defined, the crowdsourcing process is not participative and there is no open call. Finally, Flickr, which is a photo sharing website, also fails due to lack of a clear goal, lack of clear benefit to the crowd, lack of clear benefit to the initiator, not being participative and not using an open call [12].

Despite Estelles–Arola et al.'s integrative definition, some authors strongly believe websites such as Wikipedia are not only examples of crowdsourcing, but are *the classic* examples of crowdsourcing [10]. Indeed, Howe, who 'coined' crowdsourcing considers Wikipedia as a classic crowdsourcing example, as do others [4,10,20]. Osella's review found that some authors' definitions of crowdsourcing are so expansive that they consider the entire Internet a form of 'crowdsourcing,' citing O'Reilly and Batelle: "the Web as a whole is a marvel of crowdsourcing, as are marketplaces such as those on eBay and Craigslist, mixed media collections such as YouTube and Flickr, and the cast personal lifestream collections on Twitter, MySpace, and Facebook" [5].

## When to use crowdsourcing

Many authors reviewed discussed situations that were amenable to crowdsourcing (see **Box 2**). First, crowdsourcing should be used in tasks that require humans, ie, where technology either cannot complete the task or where people can do it better [3,6] and where crowds are better than individuals or experts [6]. But, what specific features would a task need to have to satisfy these broad conditions? Authors have suggested a wide range of conditions which are laid out in **Box 2**. These features are a combination of theoretical and application-based conditions and are, at times, conflicting. Kamajian reviewed crowdsourcing in medicine and his suggestions mirror Surowiecki's wisdom of the crowds conditions – he believes that the crowd must have tacit knowledge, be diverse but the problem itself must not be tacit, that the firm must not have the knowledge (otherwise why is it seeking the crowd?), and he focuses on the likelihood of the crowd's expertise and its diversity [14]. In comparison, Kittur describes applications of crowdsourcing, describing

those typically conducted through AMT, that are verifiable, have an objectifiably 'right' answer, low cognitive load, and require little expertise are most conducive to crowdsourcing [48]. Kamajian's and Kittur's images of ideal crowdsourcing are in direct opposition to one another. One feature of 'when to use crowdsourcing' that has some agreement is that the task is divisible into lower-level tasks, though this is not a necessary condition [5,48].

As opposed to focusing on characteristics of projects amenable to crowdsourcing, Buecheler and colleagues describe the characteristics of a principal investigator who would be amenable to taking on a crowdsourcing project. They state that the career age, job satisfaction, cosmopolitan scale, tenure, funding, apparatus and time must be considered; however, the authors do not give an estimate of which features within these characteristics are ideal for crowdsourcing [20].

Other authors gave specific tasks that they felt crowdsourcing was most suitable for, such as solv-

**Box 2.** Conditions for when to use Crowdsourcing found in review.

- Tasks that require humans (i.e. where technology either cannot complete the task or where people can do it better) [3,6]
- Crowds are better than individuals or experts [6]
- Firm expertise is low [14]
- Likelihood of crowd expertise is high [14]
- Firm expertise is distant from solution [14]
- Problem is not tacit, immobile, unique or complex [14]
- Relevant experience is diverse [14]
- Problem is modular [14]
- Expertise is tacit, immobile, unique or diverse [14]
- IP for problem is protect, problem is not legally protected [14,45]
- No problems with ownership or usage of solution [48]
- Problem does not contain sensitive information [5]
- Problem divisible into lower-level tasks [5,48]
- Low cognitive load [48]
- Problem is fast to complete [48]
- Task requires little expertise [48]
- Solution is objective and verifiable [48]
- Low barriers to entry [48,49]
- Low interaction required [5]

ing problems, completing tasks, being creative, developing products or ideas [5]. Castillo believed that crowdsourcing was ideal for medical imaging research, in particular, while Thawrani and colleagues suggested that researchers should use crowdsourcing to capitalise off medical data to find more specific causes of illnesses and also to bring processes up-to-date, such as handwritten medical records in India [13,32].

Finally, some authors reviewed gave tips for using crowdsourcing in research. Most importantly, selecting a clear and appropriate research question was emphasised [2,45,49]. Having a big challenge, and clear, measurable goals that are communicated to participants was seen as important as this helps motivate the participants, along with as providing options regarding levels and modes of participation [49]. Finally, the importance of acknowledging participation was highlighted [49].

### Benefits of using crowdsourcing

Benefits identified in the literature review are divided into process-based benefits and results-based benefits, and are displayed in **Table 2**. Several of these benefits could have fit into both categories. Benefits include the speed of research progression, low cost, increased accuracy of results, ability to coordinate with machine-learning and improve algorithms, act as a public advocacy tool, work in emergency situations, and transcend boundaries and borders. Crowdsourcing is a powerful, flexible tool that can be used in many situations as a supplement to traditional research. Its mobility and low cost make it ideal for global health, where barriers such as lack of human resources, funding, conflict areas and baseline epidemiological data can create barriers to targeting interventions.

### Concerns with crowdsourcing

In spite of its benefits, crowdsourcing is still subject to numerous challenges, regulatory, and ethical issues that need to be addressed, considered, and anticipated prior to designing a crowdsourcing study or intervention.

Quality assurance issues were the most commonly identified by the articles reviewed. In instances where a crowd is asked to answer questions where there is no 'right' answer, it becomes difficult to verify if responses are true and not malicious [32,48]. Additionally, there is a debate regarding having untrained laypersons complete scientific activities that are normally reserved for experts; experts may protest these activities [8,32]. Finally, concerns were voiced regarding a potential so-called "Hawthorne observer-expected effect," wherein which members of the 'crowd' acts in a way they feel the researcher may want them to [56]. Possible solutions for these issues were proposed, including having multi-level reviews. Here, there are multiple stages to each crowdsourcing task and each task is reviewed multiple times and aggregated [6] having objectifiable tasks to 'weed out' malicious workers or having standards by which

**Table 2.** Benefits of crowdsourcing listed by articles reviewed, divided into process-based benefits and results-based benefits

<b>Process-based benefits</b>	<ul style="list-style-type: none"> <li>• Low-cost alternative to traditional behavioural, epidemiological and sensory research [7,19,27,41,48]</li> <li>• Large potential scale of participants involved [27]</li> <li>• Large scale of coverage of potential intervention [16]</li> <li>• Can raise public awareness [27,32,54,55]</li> <li>• Transcends borders and boundaries [13]</li> <li>• Can be democratic [7]</li> <li>• High social robustness [29]</li> <li>• High mobility [16]</li> <li>• Able to 'tap into' untapped expertise [27]</li> <li>• Ability to cover unpredictable events [16]</li> <li>• Widespread software available to enable feasibility [16]</li> <li>• Some benefits difficult to quantify, such as "value of enthusiastic user" [45]</li> </ul>
<b>Results-based benefits</b>	<ul style="list-style-type: none"> <li>• Increased accuracy over or when results combined with machine learning tasks [27]</li> <li>• Enables high speed of research progression [27,29,48]</li> <li>• Novel discoveries [7,27,29,32,48]</li> <li>• Data produced previously unattainable [19]</li> <li>• Can complete tasks otherwise not possible, including digitizes medical artefacts or notes [32]</li> <li>• Rewards may accrue more directly [8]</li> <li>• Possible to detect and respond to disease outbreaks earlier [19]</li> <li>• Result accuracy has been shown to be equal to or more accurate than traditional research [8,47]</li> <li>• Results can improve users' lives [16]</li> </ul>

workers must fulfil prior to be considered for the task [6,27]. For example, in AMT, workers may have obtained certain scores in previous tasks.

Regarding sampling, the denominator is rarely known in crowdsourcing tasks and this can pose problems for analysis [56]. Sampling bias can occur due to inverted sampling [6,8,56] and due to self-reported data [8]. Luan and Law reported cultural and geographical biases in GIS data reviewed [19]. Additionally, there is likely to be biased samples in comparison to the general population with regards to income, literacy, age, access to technology and values [19,56].

Other authors cited concerns for security, citing potential loss of data due to a rise of cyber-attacks [38] or mishandling of sensitive information [32]. Logistical issues cited were specific to platforms or types of crowdsourcing and included troubles with languages and file formats when data mining, trouble with battery life usage, competing with prioritisation of other application on mobile devices, and privacy for ubiquitous computing (sensors in mobile devices) [19] and for AMT, not having proof of payment for work completed and institutional issues gaining approval [3]. In addition, funding being non-traditional was identified as a barrier for all crowdsourcing research [8].

### Regulatory and ethical issues

Despite Thawrani et al.'s and other's concerns that crowdsourcing could compromise anonymity, other authors were concerned that the anonymity of crowdsourcing could raise ethical concerns [6]. Williams identified instances in which crowdsourcing may have resulted in the deaths of bloggers and could be used to falsely identify (or fail to identify) weapons of mass destruction (WMD) in Iran [6]. As crowdsourcing is a nascent field, there is no Review Ethics Board (REB) or Institutional Review Board (IRB) process specific to it, to the author's knowledge, despite it being quite different from other methodologies. Exploitation of both the crowdsourcing worker and of the industries the crowdsourcing is taking place in are possible, thus REB/IRB review is very important [7,9,29,56]. Informed consent procedures will differ from general research, as researchers will not have in-person interaction with the participants and will not necessarily be aware of their levels of reading comprehension. The data use policies could represent a unique challenge to informed consent if products are used commercially.

Brabham reports that, while currently it is difficult for crowds to organise themselves against unfair labour practices, "crowdslapping" does happen [7]. This is when a crowd 'rebels' against the competition and is, essentially, a crowd of malicious workers, rallying against the project. A recent example of "crowdslapping" is a United Kingdom contest to name an RSS vessel, and the Natural Environment Research Council intended the boat to be named after an inspiring figure. The winning name was "Boaty McBoatface," which was ultimately rejected in favour of "David Attenborough." However, a remote undersea vessel was named "Boaty" in memory of the competition [57].

While not considered crowdsourcing by the working definition in this article, text/data mining has unique ethical issues, especially regarding consent, anonymity and researchers planning to use this method must consider this, through community engagement or other methods.

### Notable (non-medical) examples of crowdsourcing

A second paper [58] will review health-related examples of crowdsourcing. Aside from health-related examples, there were over 50 examples of crowdsourcing named in the reviews, with purposes ranging from public policy [42] to mapping isolationist states [6], assisting with or reporting on human rights issues [6,18], mapping or reporting on the environment [6,27], designing t-shirts [1] or linking families [49]. Some notable, interesting and successful examples of crowdsourcing in the non-scientific or medical world are described below:

#### *Guardian's MP expenses*

The UK newspaper, the Guardian, utilised crowdsourcing and freedom of information request to have the crowd comb through Members of Parliament's (MPs) expense claims to look for fraudulent claims. There were over 500 000 expense claims uploaded and over 170 000 documents were analysed within 80 hours alone [6]. As a result of this activity, British MPs were convicted of fraud, forced to resign or had to issue apologies.

#### *Ushahidi*

Ushahidi is a SMS- and web-based platform that was created after the Kenyan election in 2007 to report on election violence [6]. It is an open-sourced platform that combines GIS information with time, allow-

ing the crowdsourcing initiator to filter by place and time, which makes it ideal in disaster situations [18]. It has been used for elections, violence, corruption and disasters, including reporting cholera after the Haitian earthquake and in Kenya, Uganda, Nigeria, Haiti, Libya and Egypt [6,53].

### **GalaxyZoo**

GalaxyZoo is a crowdsourcing project that uses volunteers from around the globe to classify galaxies visually. As of 2013, it had successfully classified nearly 900 000 galaxies using hundreds of thousands of volunteers [27].

### **Transcribe Bentham**

Transcribe Bentham is a project which aims to transcribe works of Jeremy Bentham, a famous utilitarian philosopher, in order for them to be available to all. There were over 12 000 un-transcribed manuscripts and the project is based at University College London (UCL) [59].

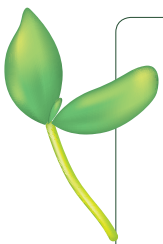
### **RECAPTCHA**

Captcha stands for “Completely Automated Public Turing test to tell Computers and Humans Apart.” Louis van Ahn, the father of human computing, extended CAPTCHA, adding an additional word so people would need to translate two words; the first was a known ‘anti-bot’ word but the second was from an archive that needed to be digitalised [60]. In 2009, RECAPTCHA was able to digitalise 20 years of the New York Times’ archives and 110 years of archives were projected to be completed by the end of 2010 [60].

## CONCLUSION

Crowdsourcing is a field that is relatively nascent, yet blossoming. Because of its infancy, researchers have not yet agreed on its definition or what does or does not constitute its practice. Despite this, several key qualities have emerged. In order to be considered crowdsourcing, a task must be distributed by an organisation via a flexible open call for the purpose of obtaining some knowledge, idea or added value, through a medium that’s similar but not an outsourced model. Usually, crowdsourcing employs the Internet, though this is not necessary. A crowd can be formed by both experts and amateurs, and the crowd can be rewarded monetarily or through recognition or skill-development. Sometimes the results are aggregated, but in other exercises, the best solution is chosen. In this way, applications of crowdsourcing are themselves very diverse and it is not surprising that authors have struggled to provide an all-encompassing definition.

Despite the difficulties defining it, crowdsourcing is beneficial both in the process and in the results. It is often low-cost, rapid, and has the possibility to transcend fields, borders, can coordinate with machine-learning, raise public awareness and produce novel discoveries. Crowdsourcing could be hugely promising in global health where resources are low and there is a paucity of data if a concerted effort is made to bring it to scale, especially through marrying the global health community with crowdsourcing and computer science researchers.



**Acknowledgments:** I would like to thank Igor Rudan for his support in writing this paper.

**Funding:** None.

**Authorship contribution:** KW conceived of the paper, conducted the searches, analysed the information and drafted the manuscript.

**Competing interests:** The author has completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author), and declares no conflicts of interest.

- 1 Brabham DC. Crowdsourcing as a model for problem solving an introduction and cases. *Convergence* (London). 2008;14:75-90. doi:10.1177/1354856507084420
- 2 Nickoloff S. Capsule Commentary on Ranard et al., Crowdsourcing-harnessing the masses to advance health and medicine, a systematic review. *J Gen Intern Med*. 2014;29:186. Medline:24197628 doi:10.1007/s11606-013-2620-0
- 3 Khare R, Good BM, Leaman R, Su AI, Lu Z. Crowdsourcing in biomedicine: challenges and opportunities. *Brief Bioinform*. 2016;17:23-32. Medline:25888696 doi:10.1093/bib/bbv021
- 4 Howe J. The rise of crowdsourcing. *Wired magazine*. 2006;14:1-4.
- 5 Osella M. A Multi-dimensional approach for framing crowdsourcing archetypes. PORTO – Publications Open Repository Torino, Politecnico di Torino; 2014. Available: <http://porto.polito.it/2535900/>. Accessed: 14 October 2017.
- 6 Williams C. Crowdsourcing research: A methodology for investigating state crime. *State Crime Journal*. 2013;2:30-51. doi:10.13169/statecrime.2.1.0030
- 7 Brabham DC. THE MYTH OF AMATEUR CROWDS. A critical discourse analysis of crowdsourcing coverage. *Inf Commun Soc*. 2012;15:394-410. doi:10.1080/1369118X.2011.641991
- 8 Swan M. Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem. *J Med Internet Res*. 2012;14:e46. Medline:22397809 doi:10.2196/jmir.1988
- 9 Brabham DC, Ribisl KM, Kirchner TR, Bernhardt JM. Crowdsourcing applications for public health. *Am J Prev Med*. 2014;46:179-87. Medline:24439353 doi:10.1016/j.amepre.2013.10.016
- 10 Zhao YX, Zhu QH. Evaluation on crowdsourcing research: Current status and future direction. *Inf Syst Front*. 2014;16:417-34. doi:10.1007/s10796-012-9350-4
- 11 Saxton GD, Oh O, Kishore R. Rules of crowdsourcing: Models, issues, and systems of control. *Inf Syst Manage*. 2013;30:2-20. doi:10.1080/10580530.2013.739883
- 12 Estellés-Arolas E, González-Ladrón-de-Guevara F. Towards an integrated crowdsourcing definition. *J Inf Sci*. 2012;38:189-200. doi:10.1177/0165551512437638
- 13 Castillo M. The wisdom of crowds. *AJNR Am J Neuroradiol*. 2013;34:1863-5. Medline:23292527 doi:10.3174/ajnr.A3417
- 14 Kamajian SD. How Crowdsourcing & Crowdfunding Are Fueling Health Care Innovation. *Osteopathic Family Physician*. 2015;7:26-30.
- 15 Wiggins A, Crowston K, editors. From conservation to crowdsourcing: A typology of citizen science. System Sciences (HICSS), 2011 44th Hawaii international conference on; 2011: IEEE.
- 16 Kanhere SS. Participatory sensing: Crowdsourcing data from mobile smartphones in urban spaces. *Distributed computing and internet technology*: Springer; 2013. p. 19-26.
- 17 Yuen M-C, King I, Leung K-S, editors. A survey of crowdsourcing systems. Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on; 2011: IEEE.
- 18 Boulos MNK, Resch B, Crowley DN, Breslin JG, Sohn G, Burtner R, et al. Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: trends, OGC standards and application examples. *Int J Health Geogr*. 2011;10:67. Medline:22188675 doi:10.1186/1476-072X-10-67
- 19 Luan H, Law J. Web GIS-based public health surveillance systems: A systematic review. *ISPRS Int J Geoinf*. 2014;3:481-506. doi:10.3390/ijgi3020481
- 20 Buecheler T, Sieg JH, Fuchsli RM, Pfeifer R, editors. Crowdsourcing, open innovation and collective intelligence in the scientific method-a research agenda and operational framework. ALIFE; 2010.
- 21 Mavandadi S, Dimitrov S, Feng S, Yu F, Sikora U, Yaglidere O, et al. Distributed medical image analysis and diagnosis through crowd-sourced games: A malaria case study. *PLoS One*. 2012;7:e37245. Medline:22606353 doi:10.1371/journal.pone.0037245
- 22 Ozcan A. Educational games for malaria diagnosis. *Sci Transl Med*. 2014;6:233ed9. Medline:24760185 doi:10.1126/scitranslmed.3009172
- 23 Mitry D, Peto T, Hayat S, Blows P, Morgan J, Khaw T, et al. Crowdsourcing as a screening tool to detect clinical features of glaucomatous optic neuropathy from digital photography. *PLoS One*. 2015;10: e0117401. Medline:25692287 doi:10.1371/journal.pone.0117401
- 24 Brady CJ, Villanti AC, Pearson JL, Kirchner TR, Gupta OP, Shah CP. Rapid grading of fundus photographs for diabetic retinopathy using crowdsourcing. *J Med Internet Res*. 2014;16:175-84. Medline:25356929 doi:10.2196/jmir.3807
- 25 Gehl R, King AJ, Grossman D, Jensen JD. Crowdsourcing as an innovative communication strategy in early melanoma detection. Available: <https://pdfs.semanticscholar.org/97e3/7c55a93a43774aa64c33a8e94edc6c9ef465.pdf> . Accessed: 15 November 2017.
- 26 Chan C, Holosko MJ. An overview of the use of mechanical turk in behavioral sciences implications for social work. *Res Soc Work Pract*. 2016;26:441-8. doi:10.1177/1049731515594024
- 27 Ranard BL, Ha YP, Meisel ZF, Asch DA, Hill SS, Becker LB, et al. Crowdsourcing-harnessing the masses to advance health and medicine, a systematic review. *J Gen Intern Med*. 2014;29:187-203. Medline:23843021 doi:10.1007/s11606-013-2536-8

- 28 Kittur A, Chi EH, Suh B, editors. Crowdsourcing user studies with Mechanical Turk. Proceedings of the SIGCHI conference on human factors in computing systems; 2008: ACM.
- 29 Prainsack B, Prainsack B, Werner-Felmayer G, Schickel S. Understanding participation: the 'citizen science' of genetics. *Genetics as Social Practice*. 2014:147-64.
- 30 Rudan I, Yoshida S, Chan KY, Sridhar D, Wazny K, Nair H, et al. Setting health research priorities using the CHNRI method: VII. A review of the first 50 applications of the CHNRI method. *J Glob Health*. 2017;7:011004. Medline:28685049 doi:10.7189/jogh.07.011004
- 31 Surowiecki J. *The wisdom of crowds*. New York: Random House; 2004.
- 32 Thawrani V, Londhe ND, Singh R. Crowdsourcing of Medical Data. *IETE Tech Rev*. 2014;31:249-53. doi:10.1080/02564602.2014.906971
- 33 Chunara R, Chhaya V, Bane S, Mekaru SR, Chan EH, Freifeld CC, et al. Online reporting for malaria surveillance using micro-monetary incentives, in urban India 2010-2011. *Malar J*. 2012;11:43. Medline:22330227 doi:10.1186/1475-2875-11-43
- 34 Chunara R, Smolinski MS, Brownstein JS. Why We Need Crowdsourced Data in Infectious Disease Surveillance. *Curr Infect Dis Rep*. 2013;15:316-9. Medline:23689991 doi:10.1007/s11908-013-0341-5
- 35 Mutale W, Bond V, Mwanamwenge MT, Mlewa S, Balabanova D, Spicer N, et al. Systems thinking in practice: the current status of the six WHO building blocks for health system strengthening in three BHOMA intervention districts of Zambia: a baseline qualitative study. *BMC Health Serv Res*. 2013;13:291. Medline:23902601 doi:10.1186/1472-6963-13-291
- 36 Leung NH, Chen A, Yadav P, Gallien J. The Impact of Inventory Management on Stock-Outs of Essential Drugs in Sub-Saharan Africa: Secondary Analysis of a Field Experiment in Zambia. *PLoS One*. 2016;11:e0156026. Medline:27227412 doi:10.1371/journal.pone.0156026
- 37 Baker U, Peterson S, Marchant T, Mbaruku G, Temu S, Manzi F, et al. Identifying implementation bottlenecks for maternal and newborn health interventions in rural districts of the United Republic of Tanzania. *Bull World Health Organ*. 2015;93:380-9. Medline:26240459 doi:10.2471/BLT.14.141879
- 38 Kostkova P. Grand challenges in digital health. *Front Public Health*. 2015;3:134. Medline:26000272 doi:10.3389/fpubh.2015.00134
- 39 Prpic J. Health care crowds: Collective intelligence in public health. 2015. Available: <https://sites.lsa.umich.edu/collectiveintelligence/wp-content/uploads/sites/176/2015/02/Prpic-CI-2015-Abstract.pdf>. Accessed: 14 October 2017.
- 40 Brabham DC. Crowdsourcing the public participation process for planning projects. *Plann Theory*. 2009;8:242-62. doi:10.1177/1473095209104824
- 41 Brabham DC. Motivations for participation in a crowdsourcing application to improve public engagement in transit planning. *J Appl Commun Res*. 2012;40:307-28. doi:10.1080/00909882.2012.693940
- 42 Aitamurto T. Crowdsourcing for democracy: A new era in policy-making. 2012. Available: [https://cddrl.fsi.stanford.edu/sites/default/files/Crowdsourcing\\_for\\_DemocracyF\\_www.pdf](https://cddrl.fsi.stanford.edu/sites/default/files/Crowdsourcing_for_DemocracyF_www.pdf). Accessed: 1 July 2017.
- 43 Doan A, Ramakrishnan R, Halevy AY. Crowdsourcing systems on the World-Wide Web. *Commun ACM*. 2011;54:86-96. doi:10.1145/1924421.1924442
- 44 Geiger D, Seedorf S, Schulze T, Nickerson RC, Schader M, editors. *Managing the crowd: Towards a taxonomy of crowdsourcing processes*. AMCIS; 2011.
- 45 Aitamurto T, Leiponen A, Tee R. The promise of idea crowdsourcing—benefits, contexts, limitations. Nokia Ideasproject White Paper. 2011.
- 46 Parvanta C, Roth Y, Keller H. Crowdsourcing 101 a few basics to make you the leader of the pack. *Health Promot Pract*. 2013;14:163-7. Medline:23299912 doi:10.1177/1524839912470654
- 47 Pedersen J, Kocsis D, Tripathi A, Tarrell A, Weerakoon A, Tahmasbi N, et al, editors. Conceptual foundations of crowdsourcing: a review of IS research. System Sciences (HICSS), 2013 46th Hawaii International Conference on; 2013: IEEE.
- 48 Kittur A. Crowdsourcing, collaboration and creativity. *ACM Crossroads*. 2010;17:22-6. doi:10.1145/1869086.1869096
- 49 Holley R. Crowdsourcing: How and why should libraries do it? *D-Lib*. 2010;16:1-21. doi:10.1045/march2010-holley
- 50 Susumpow P, Pansuwan P, Sajda N, Crawley AW, editors. Participatory disease detection through digital volunteerism: how the doctorme application aims to capture data for faster disease detection in Thailand. Proceedings of the companion publication of the 23rd international conference on World wide web companion; 2014: International World Wide Web Conferences Steering Committee.
- 51 Freeman MK. *The use of crowdsourcing and the role of game mechanics in identifying erroneous disease burden estimates*: University of Washington; 2013.
- 52 Chunara R, Aman S, Smolinski M, Brownstein JS. Flu near you: an online self-reported influenza surveillance system in the USA. *Online J Public Health Inform*. 2013. doi:10.5210/ojphi.v5i1.4456
- 53 Freifeld CC, Chunara R, Mekaru SR, Chan EH, Kass-Hout T, Iacucci AA, et al. Participatory epidemiology: Use of mobile phones for community-based health reporting. *PLoS Med*. 2010;7:e1000376. Medline:21151888 doi:10.1371/journal.pmed.1000376

- 54 Merchant RM, Asch DA, Hershey JC, Griffis H, Hill S, Saynisch O, et al. Abstract 57: A crowdsourcing, mobile media, challenge to locate automated external defibrillators. *Circulation*. 2012;126:A57.
- 55 Merchant RM, Asch DA, Hershey JC, Griffis HM, Hill S, Saynisch O, et al. A crowdsourcing innovation challenge to locate and map automated external defibrillators. *Circ Cardiovasc Qual Outcomes*. 2013;6:229-36. Medline:23481522 doi:10.1161/CIRCOUTCOMES.113.000140
- 56 Weiner M. The potential of crowdsourcing to improve patient-centered care. *Patient*. 2014;7:123-7. Medline:24671920 doi:10.1007/s40271-014-0051-6
- 57 Friedman U. The lessons of Boaty McBoatface. *Atlantic*. 2016. Available: <https://www.theatlantic.com/international/archive/2016/05/boaty-mcboatface-parliament-lessons/482046/>. Accessed: 1 July 2017.
- 58 Wazny K. Applications of Crowdsourcing in Health: An Overview. *J Glob Health*. 2017 in press.
- 59 Oomen J, Aroyo L, editors. Crowdsourcing in the cultural heritage domain: opportunities and challenges. Proceedings of the 5th International Conference on Communities and Technologies; 2011: ACM.
- 60 Schenk E, Guittard C, editors. Crowdsourcing: What can be Outsourced to the Crowd, and Why. Workshop on Open Source Innovation, Strasbourg, France; 2009.

# Distributional benefits of tobacco tax and smoke-free workplaces in China: A modeling study

Stéphane Verguet<sup>1</sup>, Gillian Tarr<sup>2</sup>,  
Cindy L Gauvreau<sup>3</sup>, Sujata Mishra<sup>4</sup>,  
Prabhat Jha<sup>4</sup>, Lingrui Liu<sup>1</sup>, Yue Xiao<sup>5</sup>,  
Yingpeng Qiu<sup>5</sup>, Kun Zhao<sup>5</sup>

<sup>1</sup> Department of Global Health and Population,  
Harvard T.H. Chan School of Public Health,  
Boston, Massachusetts, USA

<sup>2</sup> Department of Epidemiology, University of  
Washington, Seattle, Washington, USA

<sup>3</sup> Canadian Partnership Against Cancer, Toronto,  
Ontario, Canada

<sup>4</sup> Center for Global Health Research, Saint  
Michael's Hospital and University of Toronto,  
Toronto, Ontario, Canada

<sup>5</sup> China National Health Development Research  
Center, Beijing, China

**Background** Tobacco taxation and smoke-free workplaces reduce smoking, tobacco-related premature deaths and associated out-of-pocket health care expenditures. We examine the distributional consequences of a price increase in tobacco products through an excise tax hike, and of an implementation of smoke-free workplaces, in China.

**Methods** We use extended cost-effectiveness analysis (ECEA) to evaluate, across income quintiles of the male population (the large majority of Chinese smokers), the premature deaths averted, the change in tax revenues generated, and the financial risk protection procured (eg, poverty cases averted, defined as the number of individuals no longer facing tobacco-related out-of-pocket expenditures for disease treatment, that would otherwise impoverish them), that would follow a 75% increase in cigarette prices through substantial increments in excise tax fully passed onto consumers, and a nationwide total implementation of workplace smoking bans.

**Results** A 75% increase in cigarette prices would avert about 24 million premature deaths among the current Chinese male population, with a third among the bottom income quintile, increase additional tax revenues by US\$ 46 billion annually, and prevent around 9 million poverty cases, 19% of which among the bottom income quintile. Implementation of smoking bans in workplaces would avert about 12 million premature deaths, with a fifth among the bottom income quintile, decrease tax revenues by US\$ 7 billion annually, and prevent around 4 million poverty cases, 12% of which among the bottom income quintile.

**Conclusions** Increased excise taxes on tobacco products and workplace smoking bans can procure large health and economic benefits to the Chinese population, especially among the poor.

China, with its 300 million plus smoking population ie, nearly one-third of the world's total, can alleviate much of the global burden of tobacco-related disease by effectively implementing tobacco control policies. Since ratification of the World Health Organization's (WHO) Framework Convention on Tobacco Control (FCTC) in 2005, China has shown poor compliance to the FCTC's Monitor, Protect, Offer, Warn, Enforce, and Raise (MPOWER) measures [1-3]. For example, tobacco taxes remain low, contributing only 56% of final cigarette prices, and nationwide smoking bans are yet to be comprehensively implemented [3,4].

## Correspondence to:

Stéphane Verguet  
Department of Global Health and Population  
Harvard T.H. Chan School of Public Health  
665 Huntington Avenue  
Boston, MA 02115, USA  
verguet@hsph.harvard.edu



Besides significantly contributing to premature mortality, tobacco use can impose severe financial consequences for households. Out-of-pocket (OOP) health care expenditures associated with the treatment of tobacco-related disease can be impoverishing. Cigarette expenditures also form a large proportion of all household expenditures for the poor, and together with associated health care expenses arising from smoking can contribute to increased poverty rates [5-8].

China's slow response to smoking can be attributed to a deeply engrained tobacco culture along with structural and political obstructions [3]. Nevertheless, public support for tobacco control is growing [9,10] indicating potential for scaling up policies such as tobacco tax and smoke-free workplaces. Tobacco taxation is the most effective control policy [11,12], and modeling studies that assumed taxes fully passed onto consumers found substantial health and financial gains with the lowest income groups largely benefiting [13-15]. Yet, tobacco taxation has been so far underused in China [3]. Furthermore, the country may count up to 740 million individuals exposed to secondhand smoke (SHS) [3], which causes cancers and cardiovascular diseases [16]. Though policies against SHS are formulated primarily to protect non-smokers, they can create smoke-free areas and encourage smokers to quit or smoke less. For instance, a meta-analysis showed that workplace smoking bans in four high-income countries led to an average absolute 4% reduction in smoking prevalence [17]. However, the evidence on the effectiveness of smoke-free policies in reducing smoking in general is mixed, with variable effect sizes [17-19].

In China smoke-free policies have been differentially implemented at municipal and regional levels, and prominently enforced only during major events like the Beijing Olympics [20]. Overall, the proportion of workplaces having bans could range in 2010 from 60% in Shanghai to 20% in Jiangxi [21], and would be higher than the proportion of indoor public places having bans [21]. In 2014, nationwide workplace bans were officially proposed and relayed with large media coverage. Encouragingly, the "coming into effect" of comprehensive bans in Beijing in June 2015 was hailed domestically and internationally [22]. Other Chinese cities have adopted smoke-free laws. However, many municipal regulations are not effective due to weak enforcement including partial bans allowing for example smoking in some public places [20].

Using extended cost-effectiveness analysis (ECEA) methods [14,15,23,24], the objective of this paper is to examine and compare the distributional impact of expanding two critical tobacco control policies in China: aggressive increase in the excise tax on tobacco products; and enforcement of smoking bans in workplaces. In doing so, we update a previously validated ECEA framework and analysis [14] that estimated the health benefits, change in tax revenues, and financial risk protection, by socio-economic group, in China.

## METHODS

We utilize an existing ECEA analytical framework of tobacco taxation in China [14] and develop it further in simulating and comparing two key policies: (1) a large increase in excise taxes, raising the share of all applicable taxes of the retail price of tobacco products to 75%; and (2) an implementation of total smoking bans in workplaces. Taxes currently only contribute to about 56% (39% for excise taxes) of retail prices of cigarettes in China [4], which is far from the 70% excise tax contribution to the final consumer price recommended by WHO to have a large impact on cigarette consumption [1,25]. On the other hand, worksite total bans represent an essential step forward on the way to comprehensive smoke-free environments, which have been weakly enforced in China so far [20,26].

This paper builds on a former ECEA of tobacco taxation in China [14], and extends it in three important ways. First, it uses an updated set of parameters (eg, price of cigarettes, tax share as a percentage of cigarette price; Table 1). Second, it estimates smoking-related premature deaths, and cases of impoverishment and catastrophic expenditures due to OOP treatment costs of tobacco-related diseases, both critical measures of lack of financial protection commonly used by policymakers [48]. Third, it adds the examination of another key policy among the MPOWER measures [2], the enactment of smoking bans in workplaces.

We selected the policies of increases in excise taxes (i) and smoke-free workplaces (ii) as they were two essential MPOWER measures [2]. Policy (i) was chosen because China recently passed in 2015 a tobacco tax reform increasing the retail price of tobacco with tax then representing 56% of cigarette prices [4]; yet, this is far from WHO's 70% recommendation [1,25]. Policy (ii) was chosen because of China's recent advancements in adopting smoke-free laws. Notably, Beijing adopted in 2015 a comprehensive 100%

smoke-free law for all indoor public places with high compliance rates [22,49]. More than 15 other cities since enacted similar legislations or enforced smoke-free policies, such as Shanghai and Shenzhen in 2017; nevertheless, such regulations are not always fully enforced, and China has yet to implement a nationwide smoke-free legislation. Our choice of policies highlights the potentially large benefits to be reaped as China makes nascent but important, concrete steps towards reducing its own burden and, in turn impacts the global burden of smoking.

### Study population

We examined the current Chinese male population [27] as a whole (excluding Hong Kong and Macao Special Administrative Regions). Policy impact was estimated for males only, as they disproportionately engage in smoking [28,50]: nationwide, about 53% of males smoke compared to only 2% of females at ages 15 years and above [28]. The population was structured using five-year age groups from age 0 to age 84 and a single age group for all men above age 85; and further divided into income quintiles.

National smoking prevalence of manufactured cigarettes by age group was obtained from the Global Adult Tobacco Survey (GATS) China Report 2010 for ages 15–69 [28] (Table 1). A study of smoking among elderly in Hong Kong was used to estimate the smoking prevalence in men above age 70 [29]. The future smoking prevalence of those under age 15 was assumed to be the prevalence among 15- to 19-year-olds; and no additional smoking initiation would take place among those above age 15. These were two conservative assumptions as the prevalence among those aged older (eg, 25- to 49-year-olds) was higher (Table 1).

Price elasticity of demand for cigarettes varied per income quintile. We assumed an average price elasticity of  $-0.38$  [51], with greater elasticity among poorer smokers. We also assumed that price elasticity was twice as large in younger smokers (15- to 24-year-olds), who are more responsive to changes in prices (two-three times more) than older smokers [11,25]. Thus, we assigned 15- to 24-year-olds a 2-fold elasticity modifier across all quintiles, which we also applied to the elasticity that would affect smoking initiation for current 0- to 14-year-olds. Half of all price elasticity was apportioned to participation and half to consumption [11,25].

### Policy scenarios

The two policies, excise tax increase and workplace total bans, were independently modeled and applied uniformly nationwide.

#### *Excise tax increase*

Excise taxes currently only contribute to about 39% of retail prices of cigarettes in China [4]. An excise tax hike passed fully onto the consumer and resulting in a 75% increase of the cigarette pack retail price was projected, which affected smoking participation, consumption, and initiation. This led to a 75% rate of all applicable taxes (a 65% excise tax rate) on the retail price. All changes in smoking behavior except initiation were modeled as occurring in the first year of the increase. The proportion of smokers quitting and “averted” future smokers (ie, current 0- to 14-year-olds that would not initiate smoking) was calculated for each income quintile as the product of: the quintile-specific price elasticity, the proportion of the price elasticity affecting participation (one half), the youth modifier if appropriate, and the relative price increase. This proportion was further multiplied by the number of baseline smokers per age group to obtain the number of smokers who would quit or be averted.

#### *Workplace total bans*

The reduction in smoking following a total workplace ban was modeled as a one-time reduction in the number of baseline smokers based on a relative reduction in smoking prevalence of 9% (as used in SimSmoke China) [30]. We calculated the numbers of current smokers projected to quit and of future smokers averted by multiplying the number of baseline smokers per age group and the assumed relative reduction in prevalence.

### Policy outcomes

Impact was assessed for: averted premature deaths; additional revenues generated through excise tax hike; averted OOP expenditures due to tobacco-related disease treatment costs; prevented cases of medical impoverishment (hereafter referred as poverty cases); and prevented cases of catastrophic expenditures.

**Table 1.** Inputs used in the modeling of the expansion of tobacco control policies in China.

INPUT	VALUE	SOURCE
Male population by age group	<ul style="list-style-type: none"> <li>• 0–4 y-olds: 46 223 844</li> <li>• 5–9 y-olds: 42 116 819</li> <li>• 10–14 y-olds: 44 333 255</li> <li>• 15–19 y-olds: 57 372 413</li> <li>• 20–24 y-olds: 69 787 588</li> <li>• 25–29 y-olds: 54 148 396</li> <li>• 30–34 y-olds: 48 300 078</li> <li>• 35–39 y-olds: 60 477 911</li> <li>• 40–44 y-olds: 62 353 282</li> <li>• 45–49 y-olds: 52 513 698</li> <li>• 50–54 y-olds: 41 888 301</li> <li>• 55–59 y-olds: 41 743 573</li> <li>• 60–64 y-olds: 28 223 579</li> <li>• 65–69 y-olds: 19 966 448</li> <li>• 70–74 y-olds: 15 697 892</li> <li>• 75–79 y-olds: 10 754 066</li> <li>• 80–84 y-olds: 5 524 515</li> <li>• ≥85 y-olds: 2 757 397</li> </ul>	[27]
Smoking prevalence per age group (% of male population)	<ul style="list-style-type: none"> <li>• 15–19 y-olds: 14.0%</li> <li>• 20–24 y-olds: 48.8%</li> <li>• 25–29 y-olds: 53.0%</li> <li>• 30–34 y-olds: 52.2%</li> <li>• 35–39 y-olds: 57.5%</li> <li>• 40–44 y-olds: 68.0%</li> <li>• 45–49 y-olds: 66.7%</li> <li>• 50–54 y-olds: 58.0%</li> <li>• 55–59 y-olds: 57.7%</li> <li>• 60–64 y-olds: 47.3%</li> <li>• 65–69 y-olds: 37.6%</li> <li>• 70–74 y-olds: 21.0%</li> <li>• 75–79 y-olds: 19.0%</li> <li>• 80–84 y-olds: 17.0%</li> <li>• ≥85 y-olds: 13.0%</li> </ul>	Authors' calculations based on [28,29]
Relative smoking prevalence per income quintile	<ul style="list-style-type: none"> <li>• Income quintiles I to IV: 1.14 times average per age group</li> <li>• Income quintile V: 0.86 times average per age group</li> </ul>	[14,28]
Cigarette consumption (cigarettes per day) per income quintile	<ul style="list-style-type: none"> <li>• Income quintile I to V: {15.6, 15.5, 13.8, 12.7, 12.7}</li> </ul>	[14,28]
Price per pack (20 cigarettes) (2015 US\$)	<ul style="list-style-type: none"> <li>• US\$ 2.00 (before excise tax increase)</li> <li>• US\$ 3.50 (after excise tax increase)</li> </ul>	[4]
Taxes per pack (20 cigarettes) (2015 US\$)	<ul style="list-style-type: none"> <li>• US\$ 1.12 (before: 56% of retail price)</li> <li>• US\$ 2.63 (after: 75% of retail price)</li> </ul>	[4]
Relative smoking prevalence reduction among workers after workplace smoking ban	<ul style="list-style-type: none"> <li>• 9%</li> </ul>	[30]
Proportion of deaths among smokers attributable to smoking	<ul style="list-style-type: none"> <li>• 0.50</li> </ul>	[31]
Reduction of smoking-attributable death risk by age at quitting	<ul style="list-style-type: none"> <li>• 15–19 y-olds: 96.9%</li> <li>• 20–24 y-olds: 94.8%</li> <li>• 25–29 y-olds: 92.1%</li> <li>• 30–34 y-olds: 89.2%</li> <li>• 35–39 y-olds: 86.6%</li> <li>• 40–44 y-olds: 83.7%</li> <li>• 45–49 y-olds: 79.5%</li> <li>• 50–54 y-olds: 72.9%</li> <li>• 55–59 y-olds: 62.8%</li> <li>• 60–64 y-olds: 49.9%</li> <li>• 65–69 y-olds: 36.4%</li> <li>• 70–74 y-olds: 24.7%</li> <li>• 75–79 y-olds: 15.7%</li> <li>• 80–84 y-olds: 9.1%</li> <li>• ≥85 y-olds: 4.5%</li> </ul>	Authors' derivations based on [31] ( <b>Online Supplementary Document</b> , section 1)

**Table 1.** Continued

INPUT	VALUE	SOURCE
Proportion of smoking-attributable deaths per cause of death	<ul style="list-style-type: none"> <li>• COPD: 11.3%</li> <li>• Stroke: 45.5%</li> <li>• Heart disease: 22.8%</li> <li>• Neoplasm: 20.4%</li> </ul>	[32]
Tobacco-related disease treatment costs (2015 US\$)	<ul style="list-style-type: none"> <li>• COPD: US\$ 2256</li> <li>• Stroke: US\$ 2197</li> <li>• Heart disease: US\$ 11 774</li> <li>• Neoplasm: US\$ 14 794</li> </ul>	[14] and based on [33-40]
Utilization of health care by tobacco-related disease (%)	<ul style="list-style-type: none"> <li>• COPD: 33%</li> <li>• Stroke: 80%</li> <li>• Heart disease: 81%</li> <li>• Neoplasm: 50%</li> </ul>	[14] and based on [41-43]
Relative utilization of health care per income quintile	<ul style="list-style-type: none"> <li>• Income quintile I to V: (0.79, 0.98, 1.00, 1.08, 1.15) times average (applies to % above)</li> </ul>	[14] and based on [44]
Fraction of health care costs reimbursed by insurance schemes	<ul style="list-style-type: none"> <li>• 48%</li> </ul>	Authors' derivation based on [45] ( <b>Online Supplementary Document</b> , section 1)
Annual income per capita (2015 US\$)	<ul style="list-style-type: none"> <li>• Income quintile I: 0 to US\$ 992</li> <li>• Income quintile II: US\$ 992 to 1870</li> <li>• Income quintile III: US\$ 1870 to 2973</li> <li>• Income quintile IV: US\$ 2973 to 4718</li> <li>• Income quintile V: &gt; US\$ 4718</li> </ul>	Income distribution based on average per capita income of US\$ 3039 and Gini coefficient of 0.43 [46,47]
Assumed price elasticity of demand for cigarette by age group ( $\geq 25$ y-olds; 15–24 y-olds; future smokers ie, under 15 y-olds) and income quintile	<ul style="list-style-type: none"> <li>• Income quintile I: -0.64; -1.28; -1.28</li> <li>• Income quintile II: -0.51; -1.02; -1.02</li> <li>• Income quintile III: -0.38; -0.76; -0.76</li> <li>• Income quintile IV: -0.25; -0.50; -0.50</li> <li>• Income quintile V: -0.12; -0.24; -0.24</li> </ul>	

COPD, chronic obstructive pulmonary disease, y – year, d – day

Additional tax revenues were based on changes in smoking prevalence, consumption, and tax increase. Averted premature deaths among quitters were the primary health outcome. These averted deaths were then used to calculate the number of cases of poverty and catastrophic expenditures averted. All outcomes were examined by income quintile.

Tax revenues prior to policy change were calculated based on the average number of cigarettes consumed per day, the price per pack, the current tax rate (56%), and the baseline number of current smokers. After policy change, revenues were based on non-quitting smokers, reductions in consumption, and tax increases through excise tax hike (from 56% to 75% of retail price).

Smoking-related premature deaths were calculated for all male smokers currently alive. We assumed that half of all deaths among smokers were attributable to smoking [31] and that the risk of death was attenuated among former smokers based on age at quitting. Under the simplifying assumption that no current smokers would quit in the absence of policy, we calculated the number of premature deaths without policy as half (ie, 50% premature mortality rate) [31] of the baseline smoker population and of those currently under age 15 anticipated to initiate smoking. The attenuation of excess mortality risk among former smokers was modeled by age at cessation using cubic splines based on age-specific risk reductions [31]. After policy, premature deaths were calculated as 50% of continuing smokers and the age-attenuated reduction of baseline smokers quitting (Table 1). Averted premature deaths were apportioned among the four main causes of smoking-related death [32]: stroke, heart disease, neoplasms, and chronic obstructive pulmonary disease. Healthcare utilization for each cause and an adjustment per quintile were used to determine how many of those with an averted death would have incurred medical expenses. OOP expenditures were calculated by subtracting average inpatient reimbursement covered by insurance [45] from cause-specific treatment costs (see Section 1 in **Online Supplementary Document**).

Per capita income [46] and Gini coefficient [47] were used to create gamma distributions for income [52,53]. Simulations generated income for each averted premature death that would have incurred medical expenses, at the quintile level. Averted cases of poverty were calculated as individuals for whom the simulated income was above US\$ 1.90 per day but whose annual net income would have decreased to

less than US\$ 1.90 per day after paying out of pocket for disease treatment. Averted cases of catastrophic expenditures were calculated as individuals for whom OOP tobacco-related disease treatment costs would have exceeded 10% of their simulated annual income.

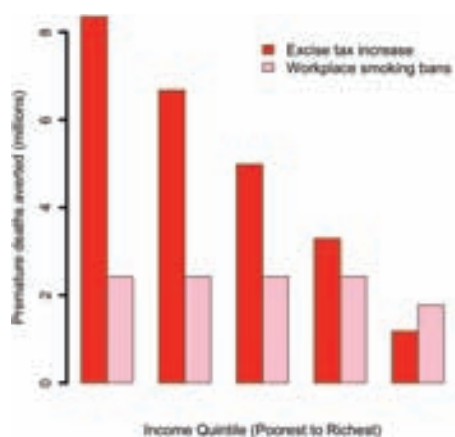
### Sensitivity analyses

A number of sensitivity analyses were conducted to test key scenarios and parameters (see Section 2 in **Online Supplementary Document**). First, for excise tax hike, the model was run with a flat price elasticity across income quintiles (eg,  $-0.38$ ). Second, we tested the impact of “brand switching” by incorporating a parameter which could capture the proportion of smokers (proportions of 0.33 and 0.75 were tested) who would respond to tobacco price increases by switching to a cheaper cigarette brand instead of quitting or decreasing consumption. This “brand switching” effect could capture for instance substitution to off-market (black market) products. Third, for workplace bans, we used an alternative effect size in assuming both an absolute reduction in smoking prevalence of 3.8% and a decrease in consumption of 3.1 cigarettes per day among continuing smokers based on a meta-analysis from four countries [17]. In this case, the absolute reduction in prevalence was further adjusted to 2.2% accounting for the proportion of worksites already having full bans (31%) [28] and restricting impact to men under age 60 that were employed (82%) [54]. Insufficient evidence however prevented from testing the differential responsiveness to worksite bans by income quintile or by age group. Yet, we tested an alternative relative prevalence reduction of 4.5% among the bottom income quintile (keeping 9% in all the other quintiles), to capture the possibility that smokers in the bottom quintile may not be employed in the formal sector where such smoking bans could be implemented. Fourth, for each policy, we selected two alternative poverty thresholds, US\$ 1 and US\$ 3 per day, respectively, to estimate the number of poverty cases averted.

Complete details of the mathematical derivations used and of the sensitivity analyses implemented are given in **Online Supplementary Document**. All simulations used the R statistical software (<http://www.r-project.org>).

## RESULTS

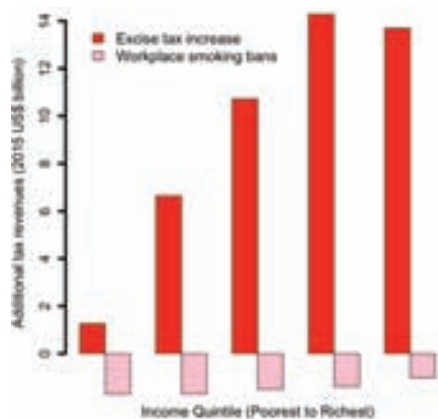
Increasing the retail price of cigarettes by 75% (or raising the tax share as a proportion of the retail price for all applicable taxes from 56% to 75%) would reduce the number of tobacco-related premature deaths by 24 million: 61% of averted deaths would be among the bottom two income quintiles, compared to 18% among the top two quintiles (**Figure 1**). This distribution is a consequence of poorer individuals being more responsive to the significant relative change (75%) in cigarette prices, and thus quitting in larger numbers. Additional annual tax revenues raised through excise tax hike would be US\$ 47 billion: 17% and 60% of these additional revenues would come from smokers from the bottom 40% and top 40% of the population income distribution, respectively (**Figure 2**). These findings are largely driven by poorer individuals being more responsive to the relative change in cigarette prices, hence quitting and smoking less cigarettes in larger numbers. The lower smoking prevalence among the top income quintile explains the slight difference in revenues between quintiles IV and V (**Figure 2**).



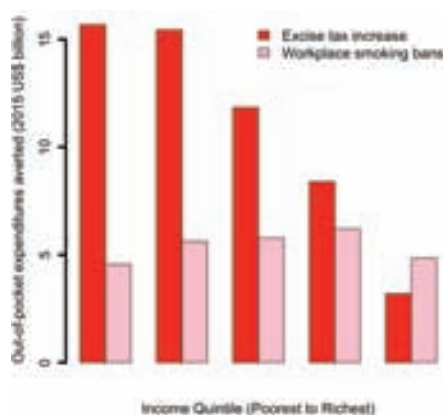
**Figure 1.** Impact of tobacco control policies (75% increase in the retail price of cigarettes through excise tax; workplace total smoking bans) in China, per income quintile, on the number of tobacco-related premature deaths averted.

OOP expenditures averted would be US\$ 55 billion: 57% and 21% of these would be among smokers from the bottom 40% and top 40% of the income distribution, respectively (**Figure 3**). Nine million cases of poverty would be prevented, primarily among the bottom two quintiles (69%; **Figure 4**); and 16 million cases of catastrophic expenditures would be prevented, primarily among the bottom two quintiles (57%; **Figure 5**). This is a consequence of: poorer individuals being more responsive to the relative change in cigarette prices, quitting in larger numbers and facing less tobacco-related disease OOP treatment costs; a lower income of poorer individuals; and the choice of the poverty threshold (see variations in the poverty cases findings when distinct poverty thresholds are used; **Figure S3 in Online Supplementary Document**).

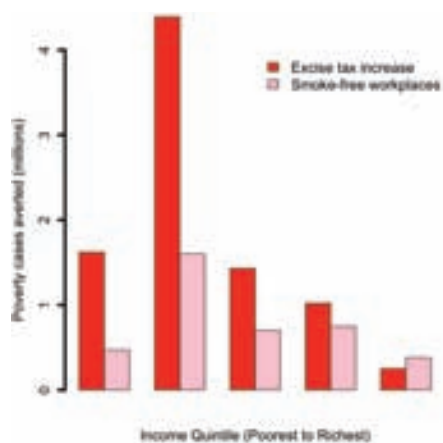
Comparatively, implementing workplace bans would yield smaller reductions in tobacco-related premature mortality and poverty cases averted (**Figures 1 and 4**). This is due to the smaller effect size (eg, among adults, 9% relative reduction in smoking prevalence for workplace bans compared with an estimated average reduction of 0.40 by 0.75 by 1/2 ~ 15% for price hike).



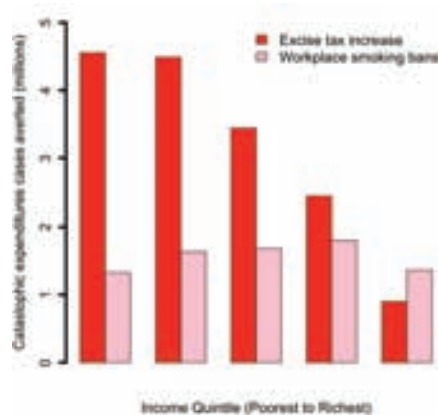
**Figure 2.** Impact of tobacco control policies (75% increase in the retail price of cigarettes through excise tax; workplace total smoking bans) in China, per income quintile, on the net change in annual tax revenues collected on cigarette sales among current smokers (15 years of age and above).



**Figure 3.** Impact of tobacco control policies (75% increase in the retail price of cigarettes through excise tax; workplace total smoking bans) in China, per income quintile, on the amount of out-of-pocket tobacco-related disease treatment costs averted.



**Figure 4.** Impact of tobacco control policies (75% increase in the retail price of cigarettes through excise tax; workplace total smoking bans) in China, per income quintile, on the number of tobacco-related poverty cases averted due to the prevention of out-of-pocket tobacco-related disease treatment costs.

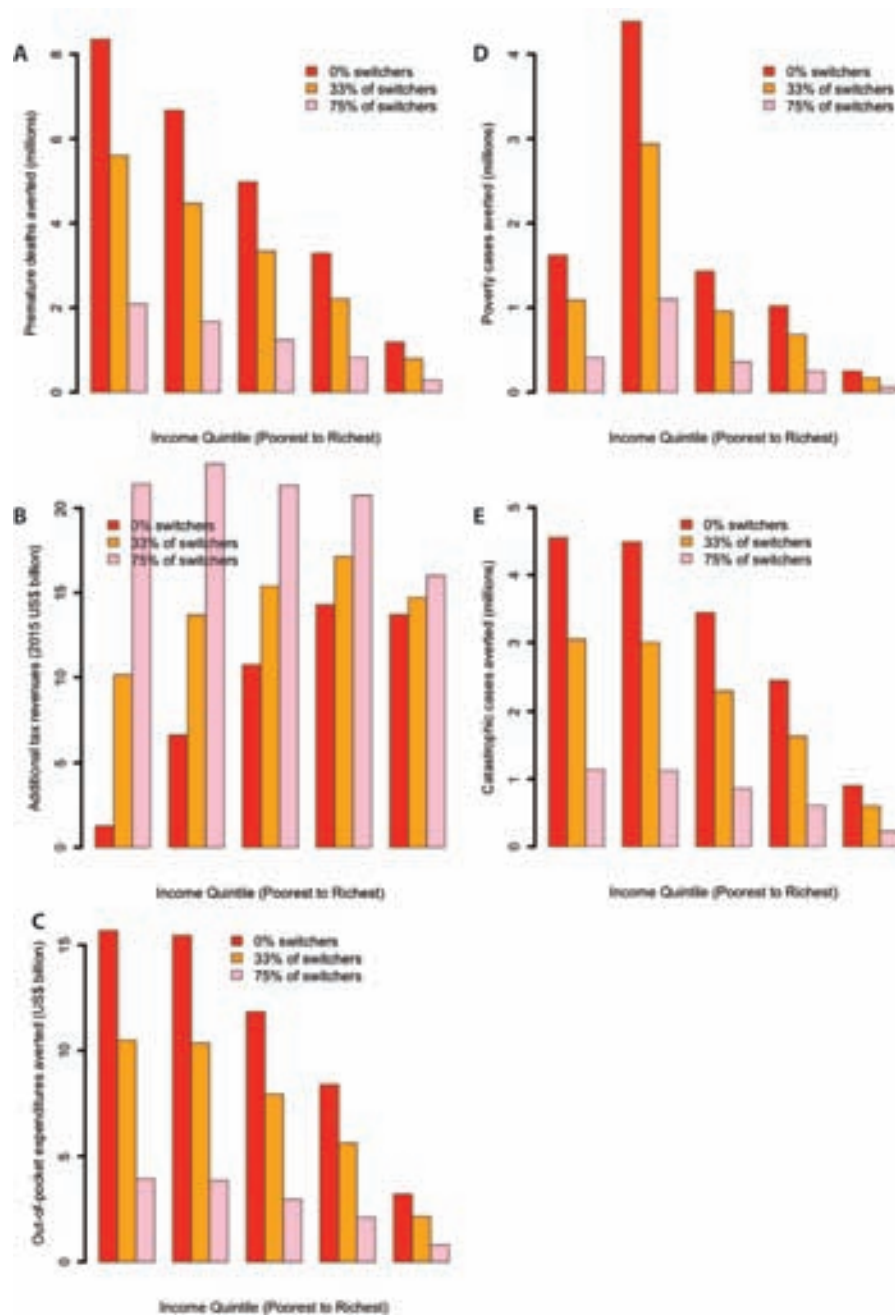


**Figure 5.** Impact of tobacco control policies (75% increase in the retail price of cigarettes through excise tax; workplace total smoking bans) in China, per income quintile, on the number of tobacco-related cases of catastrophic expenditures averted due to the prevention of out-of-pocket tobacco-related disease treatment costs.

By implementing a ban, premature deaths were estimated to decrease by 12 million and to be evenly distributed across quintiles, due to a more or less flat distribution of smoking prevalence and assumed equal responsiveness to smoking bans by income quintile. Four million cases of poverty would be prevented, primarily among the bottom two quintiles (52%), explained by the fact that poorer individuals had lower income. Compared with the excise tax increase, a reduction in smoking with workplace bans would decrease tax revenues by US\$ 7 billion (Figure 2) and will not be compensated by tax hike on tobacco products. Larger decreases (47% of the total decrease) would be observed among the bottom two income quintiles as a result of a larger number of cigarettes consumed among these groups in the first place.

Sensitivity analyses showed that policy impact was affected by a number of parameters and scenarios. First, the distributional analysis of the excise tax was largely influenced by the differential responsiveness to price changes per income quintile. Predictably, the excise tax hike progressiveness disappeared when all quintiles were given the same price elasticity (Figure S1 in **Online Supplementary Document**). Assuming a flat price elasticity would equalize the number of premature deaths averted, lead to a larger share of additional taxes borne by the poor, while some pro-poor aspect of the impoverishment averted would be maintained as the poorer income quintiles would evidently still have a lower income. Second, cigarette brand switching could significantly alter the findings (Figure 6). Introducing brand switching produced large reductions in averted deaths (compared with the base case) equivalent to the proportion of individuals switching brands (eg, assuming 75% of smokers switch, deaths averted would decrease by 75%). Additional tax revenues through excise tax hike could increase substantially and were more evenly distributed among income quintiles, minimizing policy progressiveness. In summary, less progressiveness followed greater switching. Third, using an alternative effect size for workplace bans (absolute reduction in smoking prevalence and cigarette consumption) would alter the conclusions (Figure S2 in **Online Supplementary Document**): both premature deaths and poverty cases averted would decrease; and the additional tax revenues would decrease further when smoking prevalence reduction is accompanied by consumption reduction. Likewise, a smaller prevalence reduction of 4.5% among the bottom quintile would decrease

substantially (by 50%) the premature deaths averted, increase the OOP expenditures and the poverty and catastrophic cases among the poor; it would however decrease the revenue losses among the bottom income quintile (Figure S2 in **Online Supplementary Document**). Fourth, we found that the poverty cases headcounts could be substantially affected when using distinct poverty thresholds (Figure S3 in **Online Supplementary Document**). Expectedly, progressiveness was enhanced when the poverty threshold was reduced (eg, from US\$ 3.00 to US\$ 1.00 per day).



**Figure 6.** Impact of a 75% increase in the retail price of cigarettes through excise tax (proportion of smokers switching to cheaper cigarette brands, ie, “switchers”, was set at either 0%, 33%, or 75%) in China, per income quintile, on: the number of tobacco-related premature deaths averted (a); the net change in annual tax revenues collected on cigarette sales among current smokers (15 years of age and above) (b); the amount of out-of-pocket tobacco-related disease treatment costs averted (c); the number of tobacco-related poverty cases averted due to the prevention of out-of-pocket tobacco-related disease treatment costs (d); and the number of tobacco-related cases of catastrophic expenditures averted due to the prevention of out-of-pocket tobacco-related disease treatment costs (e).

## DISCUSSION

We studied the distributional impact of expanding two tobacco control policies, aggressive increase in the excise tax on tobacco products and enforcement of workplace smoking bans, in China. On the one hand, excise tax hike passed onto the consumer in the form of a 75% retail price increase would prevent 24 million premature deaths (about 2% of China's population) and 9 million cases of poverty, and yield an annual US\$ 47 billion more in revenues. China's poorest would experience the greatest benefits in averted deaths and impoverishments while bearing a smaller burden of the tax hike. On the other hand, instituting workplace smoking bans would have a more moderate impact on mortality and impoverishment averted. Assuming a relative decrease in smoking prevalence of 9%, 12 million premature deaths (or 1% of the Chinese population) and 4 million poverty cases could be averted, while revenues would decrease.

This analysis has shown that expanded tobacco control could promote equity at the national level in China. Yet, a subnational examination of smoking-related inequalities is required to fine tune policy. Significant geographical variations in income and health exist in China, with the West and Southwest less economically advanced, uneven access to health care between urban and rural populations, and considerable intra-urban inequality. Even in the major cities like Beijing, large health and economic gains can accrue for the poor.

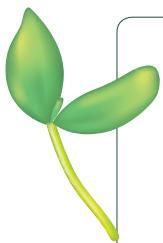
Our estimates of averted premature mortality and revenue gains are consistent with previous work [13,14]. Few models on impact of workplace smoking bans exist against which to compare our results. We chose to focus on workplace bans given availability of data and also because the working population of ages 25–54 has the highest smoking prevalence and thus is of greater relevance [28]. Unfortunately, the lack of clear evidence [18] for differential responses to smoke-free policies across socio-economic groups precluded us from examining whether such policies could redress inequities observed in outcomes.

Nevertheless, our analysis presents a number of limitations. First, our estimates are sensitive to assumptions about the price elasticity of demand for tobacco. We have used an elasticity of  $-0.38$  following norms from developed countries [11]. China has a few studies estimating price elasticity of demand for tobacco with ranges from  $-0.84$  to  $-0.01$  [51]. Nonetheless, we elected to use a value closer to high-income countries to get a more conservative estimate as these studies report a wide range of price elasticities. Second, our effect size for the relative reduction in prevalence due to smoking bans (eg, 9%) is limited by basing it on simulation inputs [30] that may not be generalizable. We certainly underestimated the potential impact of bans by our focus on mortality and men alone. Women could be the main beneficiaries of reduced SHS exposure [55,56], which could also lead to significant decreases in hospital admissions and associated OOP medical payments [57]. In addition, the distributional impact of smoke-free policies is unclear. In four countries, there was no correlation between socio-economic status and the introduction of smoking bans in workplaces [58]; and Dinno and Glantz [59] found that the decrease in smoking prevalence due to clean air laws in the US did not vary by socio-economic status. Third, in the past, tax increases have not generated significant behavior changes among Chinese smokers [51], which may make our model appear optimistic. For example, cigarette prices have not increased in China at the same rate as disposable incomes making them more affordable [51]. However, past excise tax increases have been low, and we emphasize here the importance of large excise tax hikes. Furthermore, cigarettes in China also have wide variation in prices, allowing consumers to switch to lower-priced brands when taxes increase [60–63]. Therefore, we have modeled different levels of brand-switching in sensitivity analyses to explore its potential impact on our estimates (Figure 4). This is equivalent to using a lower average price elasticity, similar to those seen in other models [13,30]. With the largest brand-switching modeled (equivalent to an average elasticity of  $-0.10$ ), five million premature deaths would still be averted, but policy progressivity would be diminished. A more moderate switching parameter (corresponding to an average elasticity of  $-0.25$ ) would project an averted 14 million premature deaths and preserve some progressivity. However, more research is needed for quantifying the extent of brand-switching and which smokers are more likely to switch to enact an optimum level of taxation. Fourth, since we examined the consumer perspective, one major limitation is not taking into account the role of the Chinese State Tobacco Monopoly Administration (STMA) and China's tobacco tax structure. STMA determines cigarette prices, including the use of central government and local government taxes, and thus, tax increases do not get necessarily passed onto the retail price of tobacco products [63,64]. Excise taxes will only have an effect when increases are passed onto the consumers through higher retail prices [60,63], which is what our analysis assumed. If the tax increase were not fully (but partially) passed onto the consumers, we would still observe reductions (though diminished) in premature mortality and tobacco-related OOP spending and

impoverishment. Fifth, as in all models, we had to balance interpretability and strength of existing evidence with realism. For example, we assumed that changes in smoking behaviors due to policy would occur among individuals who would otherwise not have quit on their own. Thus, we did not attempt to capture background quitting or consumption reductions. This simplifying assumption would have resulted in an overestimate of impact in presence of downward smoking trends. We also assumed that all changes in smoking behaviors were instantaneous and persisted over the life of individuals; and did not account for the fact that increased taxes may themselves be a source of impoverishment and enhance poverty, notably for those among the bottom income quintile who do not quit. Finally, for simplicity, our analysis studied one policy at a time, and thus did not model any synergies and interactions from the effects of both policies. We expect that as tobacco control in China grows and individual policies become integrated in coordinated national frameworks, evidence may be collected and research be conducted to examine complementarity of measures and their results.

Our analysis focuses on China but its findings are relevant to many other low- and middle-income countries. Other settings have already successfully implemented excise tax hikes and smoking ban policies validating our approach. Large increases in specific excise taxes can have a substantial impact on cigarette consumption [11,12,25]. For example, youth smoking is very responsive to cigarette prices as shown by data from 17 low- and middle-income countries [65]; and over 15 years, South Africa tripled cigarette prices and halved tobacco consumption with large tax hikes [66,67]. Likewise, a review [18] showed a consistent positive impact of national smoking bans on improving cardiovascular health outcomes and reducing mortality and morbidity from tobacco-related diseases, based on data from 21 countries including the middle-income countries of Argentina [68,69], Uruguay [70], Panama [71], and Turkey [72]. And a meta-analysis demonstrated that workplace bans in Australia, Canada, Germany, and the U.S led to a 4% absolute reduction in smoking prevalence [17].

Our results highlight the need to consider not only the overall impact of policies to decrease smoking but also how impact is distributed across sub-populations. More importantly, the distributional impact of tobacco control efforts provides governments with relevant evidence to make the biggest difference for the populations that most need it. We show here that increasing cigarette taxes and instituting smoke-free workplaces can potentially prevent millions from being impoverished as a consequence of smoking-related medical expenditures. Because of the structure of the existing tobacco policies and state-owned tobacco industry, we believe that a priority should be placed on full implementation of all MPOWER measures [2] including comprehensive smoking bans and large excise tax hikes. Following India's differentially taxed cigarettes based on the length of cigarettes may help mitigate brand-switching and other compensating behaviors [73]. Higher taxes and workplace smoking bans can work well together and are complimentary: by having a mutually strengthening effect (eg, smoking bans impose social norms and enhance the price effect of taxes) they can reinforce each other to both lower consumption and bring large health and economic benefits to households. In addition, higher taxes can lead to revenue increases partially offsetting revenue losses from smoking bans. Such benefits would arise from large excise tax hikes that explicitly narrow the price differentials from top to bottom cost cigarettes, combined with total (not partial) smoking bans. In summary, the simultaneous implementation of both policies would present great synergies: mutual strengthening and exponentiation of the health and financial protection benefits (eg, poverty cases averted); compensation of lost revenues by smoke-free places by increases in excise taxes; and providing a resulting combined pro-poor policy where the potentially flat distribution of benefits of smoke-free places are compensated by increases in excise taxes, and increases in taxes to the poor are compensated by reductions in cigarette spending due to bans.



**Acknowledgments:** We thank Angela Pratt and Andrea Pastorelli for valuable comments on an earlier version of the manuscript, as well as two reviewers for helpful and constructive suggestions.

**Disclaimer:** The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the World Health Organization, and the United Nations Development Programme.

**Funding:** This work was originally funded by the World Health Organization and the United Nations Development Programme. GT was initially supported in part by the National Institute of Environmental Health Sciences of the National Institutes of Health under award number T32ES015459.

**Authorship contributions:** SV and GT initiated and conceptualized the study. SV coordinated the research and GT did the analysis with CLG, SM, LL, SQ, CY, and KZ. SV, GT, and CLG wrote the first draft of the report. CLG, SM, KZ, and PJ reviewed the report and provided advice and suggestions.

**Competing interests:** The authors have completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author) and declare no conflicts of interests.

## REFERENCES

- 1 World Health Organization. WHO Report on the Global Tobacco Epidemic 2015. Available: [http://www.who.int/tobacco/global\\_report/2015/en/](http://www.who.int/tobacco/global_report/2015/en/). Accessed: 27 November 2016.
- 2 World Health Organization. MPOWER: a policy package to reverse the tobacco epidemic. Geneva: World Health Organization; 2011.
- 3 Yang G, Wang Y, Wu Y, Yang J, Wan X. The road to effective tobacco control in China. *Lancet*. 2015;385:1019-28. Medline:25784349 doi:10.1016/S0140-6736(15)60174-X
- 4 Zheng R. 2015 Tobacco tax reform in China: results and challenges. May 2016, World Bank, Washington, DC. Available: <http://pubdocs.worldbank.org/en/931281464893347787/6-VI-Rose-Zheng-ppt-China-2015-tobacco-tax-reform-Rose-ZHENG.pdf>. Accessed: 26 November 2016.
- 5 Liu Y, Rao K, Hu T-w, Sun Q, Mao Z. Cigarette smoking and poverty in China. *Soc Sci Med*. 2006;63:2784-90. Medline:16959391 doi:10.1016/j.socscimed.2006.06.019
- 6 Xin Y, Qian J, Xu L, Tang S, Gao J, Critchley JA. The impact of smoking and quitting on household expenditure patterns and medical care costs in China. *Tob Control*. 2009;18:150-5. Medline:19158112 doi:10.1136/tc.2008.026955
- 7 Hu TW, Mao Z, Liu Y, de Beyer J, Ong M. Smoking, standard of living, and poverty in China. *Tob Control*. 2005;14:247-50. Medline:16046687 doi:10.1136/tc.2004.010777
- 8 Wang H, Sindelar JL, Busch SH. The impact of tobacco expenditure on household consumption patterns in rural China. *Soc Sci Med*. 2006;62:1414-26. Medline:16137812 doi:10.1016/j.socscimed.2005.07.032
- 9 Li Q, Hyland A, O'Connor R, Zhao G, Du L, Li X, et al. Support for smoke-free policies among smokers and non-smokers in six cities in China: ITC China Survey. *Tob Control*. 2010;19 Suppl 2:i40-6. Medline:19679889 doi:10.1136/tc.2009.029850
- 10 New York Times. Proposed Regulations Aim to Clamp Down on Smoking in China. 25 November 2014. Available: [http://sinosphere.blogs.nytimes.com/2014/11/25/antismoking-regulations-seek-to-change-face-of-china/?\\_r=0](http://sinosphere.blogs.nytimes.com/2014/11/25/antismoking-regulations-seek-to-change-face-of-china/?_r=0). Accessed: 26 November 2016.
- 11 International Agency for Research on Cancer (IARC). IARC Handbooks of Cancer Prevention in Tobacco Control, vol. 14: Effectiveness of tax and price policies for tobacco control. Lyon: World Health Organization; 2011.
- 12 Jha P, MacLennan M, Chaloupka FJ, Yurekli A, Ramasundharahettige C, Palipudi K, et al. Global hazards of tobacco and the benefits of smoking cessation and tobacco taxes. In: Gelband H, Jha P, Sankaranaryanan R, Horton S. *Disease Control Priorities, Third Edition. Volume 3: Cancer*. Washington, DC: World Bank; 2015.
- 13 Jha P, Joseph R, Li D, Gauvreau C, Anderson I, Moser P, et al. Tobacco taxes: a win-win measure for fiscal space and health. Mandaluyong City: Asian Development Bank; 2012.
- 14 Verguet S, Gauvreau CL, Mishra S, MacLennan M, Muprhy SM, Brouwer ED, et al. The consequences of tobacco tax on household health and finances in rich and poor smokers in China: an extended cost effectiveness analysis. *Lancet Glob Health*. 2015;3:e206-16. Medline:25772692 doi:10.1016/S2214-109X(15)70095-1
- 15 Salti N, Brouwer E, Verguet S. The health, financial and distributional consequences of increases in the tobacco excise tax among smokers in Lebanon. *Soc Sci Med*. 2016;170:161-9. Medline:27792922 doi:10.1016/j.socscimed.2016.10.020
- 16 US Department of Health and Human Services. *The Health Consequences of Involuntary Exposure to Tobacco Smoke: a Report of the Surgeon General*. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, Coordinating Center for Health Promotion, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2006.
- 17 Fichtenberg CM, Glantz SA. Effect of smoke-free workplaces on smoking behaviour: systematic review. *BMJ*. 2002;325:188. Medline:12142305 doi:10.1136/bmj.325.7357.188
- 18 Frazer K, Callinan JE, McHugh J, van Baarsel S, Clarke A, Doherty K, et al. Legislative smoking bans for reducing secondhand smoke exposure, smoking prevalence and tobacco consumption [Review]. *Cochrane Database Syst Rev*. 2016;2:CD005992. Medline:26842828
- 19 Wilson LM, Avila Tang E, Chander G, Hutton HE, Odelola OA, Elf JL, et al. Impact of tobacco control interventions on smoking initiation, cessation, and prevalence: a systematic review. *J Environ Public Health*. 2012;2012:961724. Medline:22719777 doi:10.1155/2012/961724

- 20 World Health Organization Western Pacific Region and University of Waterloo. ITC Project. Smoke-free policies in China: evidence of effectiveness and implications for action. Manila: World Health Organization Regional Office for the Western Pacific; 2015.
- 21 Wan X, Stillman F, Liu H, Spires M, Dai Z, Tamplin S, et al. Development of policy performance indicators to assess the implementation of protection from exposure to secondhand smoke in China. *Tob Control*. 2013;22 Suppl 2:ii9-15. Medline:23760608 doi:10.1136/tobaccocontrol-2012-050890
- 22 Schwartländer B, Pratt A. Tobacco in China: taming the smoking dragon. *Lancet*. 2015;385:2123-4. Medline:26068250 doi:10.1016/S0140-6736(15)60996-5
- 23 Verguet S, Laxminarayan R, Jamison DT. Universal public finance of tuberculosis treatment in India: an extended cost-effectiveness analysis. *Health Econ*. 2015;24:318-32. Medline:24497185 doi:10.1002/hec.3019
- 24 Verguet S, Kim JJ, Jamison DT. Extended cost-effectiveness analysis for health policy assessment: a tutorial. *Pharmacoeconomics*. 2016;34:913-23. Medline:27374172 doi:10.1007/s40273-016-0414-z
- 25 World Health Organization. WHO technical manual on tobacco tax administration. Geneva: World Health Organization; 2011.
- 26 Fong GT, Sansone G, Yan M, Math M, Craig L, Quah ACK, et al. Evaluation of smoke-free policies in seven cities in China: longitudinal findings from the ITC project (2007-2012). *Tob Control*. 2015;24:iv14-20. Medline:26407720 doi:10.1136/tobaccocontrol-2015-052508
- 27 United Nations Population Division Department of Economic and Social Affairs, World Population Prospects: The 2012 Revision, June 2013. Available: [https://esa.un.org/unpd/wpp/publications/Files/WPP2012\\_HIGHLIGHTS.pdf](https://esa.un.org/unpd/wpp/publications/Files/WPP2012_HIGHLIGHTS.pdf). Accessed: 1 August 2017.
- 28 Global Adult Tobacco Survey. Global Adult Tobacco Survey China 2010 Country Report. Geneva: World Health Organization; 2010.
- 29 Lam TH, Xu L, Schooling CM, Chan WM, Lee SY, Leung GM. Smoking and mortality in a prospective cohort study of elderly Chinese in Hong Kong. *Addiction*. 2015;110:502-10. Medline:25331629 doi:10.1111/add.12776
- 30 Levy D, Rodriguez-Buno RL, Hu T-w, Moran AE. The potential effects of tobacco control in China: projections from the China SimSmoke simulation model. *BMJ*. 2014;348:g1134. Medline:24550245 doi:10.1136/bmj.g1134
- 31 Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ*. 2004;328:1519. Medline:15213107 doi:10.1136/bmj.38142.554479.AE
- 32 Global Burden of Disease Study. 2010. Global Burden of Disease Study 2010. Results by risk factor 1990-2010. Seattle: Institute for Health Metrics and Evaluation; 2012.
- 33 Le C, Zhankun S, Jun D, Keying Z. The economic burden of hypertension in rural south-west China. *Trop Med Int Health*. 2012;17:1544-51. Medline:22973901 doi:10.1111/j.1365-3156.2012.03087.x
- 34 Lee VW, Chan WK, Lam NL, Lee KK. Cost of acute myocardial infarction in Hong Kong. *Dis Manag Health Outcomes*. 2005;13:281-5. doi:10.2165/00115677-200513040-00006
- 35 Wei JW, Heeley EL, Jan S, Huang Y, Huang Q, Wang J-G, et al. Variations and determinants of hospital costs for acute stroke in China. *PLoS One*. 2010;5:e13041. Medline:20927384 doi:10.1371/journal.pone.0013041
- 36 Ma Y, Liu Y, Fu HM, Wang XM, Wu BH, Wang SX, et al. Evaluation of admission characteristics, hospital length of stay and costs for cerebral infarction in a medium-sized city in China. *Eur J Neurol*. 2010;17:1270-6. Medline:20402751 doi:10.1111/j.1468-1331.2010.03007.x
- 37 Heeley E, Anderson CS, Huang Y, Jan S, Li Y, Liu M, et al. Role of health insurance in averting economic hardship in families after acute stroke in China. *Stroke*. 2009;40:2149-56. Medline:19359646 doi:10.1161/STROKEAHA.108.540054
- 38 He QY, Zhou X, Xie CM, Liang ZA, Chen P, Wu CG, et al. Impact of chronic obstructive pulmonary disease on quality of life and economic burden in Chinese urban areas. *Zhonghua Jie He He Hu Xi Za Zhi*. 2009;32:253-7. Medline:19576036
- 39 Zeng X, Karnon J, Wang S, Wu B, Wan X, Peng L. The cost of treating advanced non-small cell lung cancer: estimates from the Chinese experience. *PLoS One*. 2012;7:e48323. Medline:23118985 doi:10.1371/journal.pone.0048323
- 40 She J, Yang P, Hong Q, Bai C. Lung cancer in China: challenges and interventions. *Chest*. 2013;143:1117-26. Medline:23546484 doi:10.1378/chest.11-2948
- 41 Zhao D, Liu J, Wang W, Zeng Z, Cheng J, Liu J, et al. Epidemiological transition of stroke in China twenty-one-year observational study from the Sino-MONICA Beijing project. *Stroke*. 2008;39:1668-74. Medline:18309149 doi:10.1161/STROKEAHA.107.502807
- 42 Chai Y, Xu H, Wang W, Liu B, Yang D, Fan H, et al. A survey of factors associated with the utilization of community health centers for managing hypertensive patients in Chengdu, China. *PLoS One*. 2011;6:e21718. Medline:21750725 doi:10.1371/journal.pone.0021718
- 43 Zhong N, Wang C, Yao W, Chen P, Kang J, Huang S, et al. Prevalence of chronic obstructive pulmonary disease in China: a large, population-based survey. *Am J Respir Crit Care Med*. 2007;176:753-60. Medline:17575095 doi:10.1164/rccm.200612-1749OC
- 44 Shanghai Municipal Center for Disease Control & Prevention (SCDC). Study on global AGEing and adult health (SAGE), Wave 1. China National Report. Study Report October, 2012. Available: <http://apps.who.int/healthinfo/systems/survey-data/index.php/catalog/13/download/1874>. Accessed: 23 April 2014.
- 45 Yip WC, Hsiao WC, Chen W, Hu S, Ma J, Maynard A. Early appraisal of China's huge and complex health-care reforms. *Lancet*. 2012;379:833-42. Medline:22386036 doi:10.1016/S0140-6736(11)61880-1
- 46 National Bureau of Statistics of China. China Statistical Yearbook 2014, Table 6-1 per capital income and consumption expenditure nationwide. Available: <http://www.stats.gov.cn/tjsj/ndsj/2014/zk/html/Z0601e.htm>. Accessed: 26 November 2016.
- 47 World Bank. World development indicators. Available: <http://data.worldbank.org/data-catalog/world-development-indicators>. Accessed: 26 November 2016.

- 48 World Health Organization. World Bank. Tracking universal health coverage: first global monitoring report. Geneva: World Health Organization; 2015.
- 49 World Health Organization Western Pacific Region and United Nations Development Programme. The bill China cannot afford to pay. Manila: World Health Organization Regional Office for the Western Pacific; 2017.
- 50 Chen Z, Peto R, Zhou M, Iona A, Smith M, Yang L, et al. Contrasting male and female trends in tobacco-attributed mortality in China: evidence from successive nationwide prospective studies. *Lancet*. 2015;386:1447-56. Medline:26466050 doi:10.1016/S0140-6736(15)00340-2
- 51 Hu T-w, Mao Z, Shi J, Chen W. Tobacco taxation and its potential impact in China. Paris: International Union Against Tuberculosis and Lung Disease, 2008.
- 52 Kemp-Benedict E. Income distribution and poverty – methods for using available data in global analysis. 2001. Available: [http://gdrs.sourceforge.net/docs/PoleStar\\_TechNote\\_4.pdf](http://gdrs.sourceforge.net/docs/PoleStar_TechNote_4.pdf). Accessed: 26 November 2016.
- 53 Salem ABZ, Mount TD. A convenient descriptive model of income distribution: the gamma density. *Econometrica*. 1974;42:1115-27. doi:0012-9682(197411)42:6<1115:ACDMOI>2.0.CO;2-U
- 54 China Labour Bulletin. 2014. Employment in China, 22 June, 2013. Available: <http://www.clb.org.hk/en/content/employment-china>. Accessed: 15 January 2014.
- 55 Gan Q, Smith KR, Hammond SK, Hu T-w. Disease burden of adult lung cancer and ischaemic heart disease from passive tobacco smoking in China. *Tob Control*. 2007;16:417-22. Medline:18048620 doi:10.1136/tc.2007.021477
- 56 Oberg M, Jaakkola MS, Woodward A, Peruga A, Pruss-Ustun A. Worldwide burden of disease from exposure to second-hand smoke: a retrospective analysis of data from 192 countries. *Lancet*. 2011;377:139-46. Medline:21112082 doi:10.1016/S0140-6736(10)61388-8
- 57 Tan CE, Glantz SA. Association between smoke-free legislation and hospitalizations for cardiac, cerebrovascular, and respiratory diseases: a meta-analysis. *Circulation*. 2012;126:2177-83. Medline:23109514 doi:10.1161/CIRCULATIONAHA.112.121301
- 58 King BA, Hyland AJ, Borland R, McNeill A, Cummings KM. Socioeconomic variation in the prevalence, introduction, retention, and removal of smoke-free policies among smokers: findings from the International Tobacco Control (ITC) Four Country Survey. *Int J Environ Res Public Health*. 2011;8:411-34. Medline:21556194 doi:10.3390/ijerph8020411
- 59 Dinno A, Glantz S. Tobacco control policies are egalitarian: a vulnerabilities perspective on clean indoor air laws, cigarette prices, and tobacco use disparities. *Soc Sci Med*. 2009;68:1439-47. Medline:19282078 doi:10.1016/j.socscimed.2009.02.003
- 60 White JS, Li J, Hu T-w, Fong GT, Jiang Y. The effect of cigarette prices on brand-switching in China: a longitudinal analysis of data from the ITC China Survey. *Tob Control*. 2014;23 Suppl 1:i54-60. Medline:23697645 doi:10.1136/tobaccocontrol-2012-050922
- 61 Li J, White JS, Hu T-w, Fong GT, Yuan J. The heterogeneous effects of cigarette prices on brand choice in China: implications for tobacco control policy. *Tob Control*. 2015;24 Suppl 3:iii25-32. Medline:25855642 doi:10.1136/tobaccocontrol-2014-051887
- 62 Huang J, Zheng R, Chaloupka FJ, Fong GT, Li Q, Jiang Y. Chinese smokers' cigarette purchase behaviors, cigarette prices and consumption: findings from the ITC China Survey. *Tob Control*. 2014;23 Suppl 1:i67-72. Medline:24048580 doi:10.1136/tobaccocontrol-2013-051057
- 63 Gao S, Zheng R, Hu T-w. Can increases in the cigarette tax rate be linked to cigarette retail prices? Solving mysteries related to the cigarette pricing mechanism in China? *Tob Control*. 2012;21:560-2. Medline:23076787 doi:10.1136/tobaccocontrol-2011-050027
- 64 Hu TW, Mao Z, Shi J, Chen W. The role of taxation in tobacco control and its potential economic impact in China. *Tob Control*. 2010;19:58-64. Medline:20008158 doi:10.1136/tc.2009.031799
- 65 Kostova D, Ross H, Blecher E, Markowitz S. Is youth smoking responsive to cigarette prices? Evidence from low- and middle-income countries. *Tob Control*. 2011;20:419-24. Medline:21737858 doi:10.1136/tc.2010.038786
- 66 Jha P, Peto R. Global effects of smoking, of quitting, and of taxing tobacco. *N Engl J Med*. 2014;370:60-8. Medline:24382066 doi:10.1056/NEJMr1308383
- 67 Van Walbeek C. Industry responses to the tobacco excise tax increases in South Africa. *S Afr J Econ*. 2006;74:110-22. doi:10.1111/j.1813-6982.2006.00051.x
- 68 Ferrante D, Linetzky B, Virgolini M, Schoj V, Apelberg B. Reduction in hospital admissions for acute coronary syndrome after the successful implementation of 100% smoke-free legislation in Argentina: a comparison with partial smoking restrictions. *Tob Control*. 2012;21:402-6. Medline:21602536 doi:10.1136/tc.2010.042325
- 69 Konfino J, Ferrante D, Mejia R, Coxson P, Moran A, Goldman L, et al. Impact on cardiovascular disease events of the implementation of Argentina's national tobacco control law. *Tob Control*. 2014;23:e6. Medline:23092886 doi:10.1136/tobaccocontrol-2012-050599
- 70 Sebríe EM, Sandoya E, Bianco E, Hyland A, Cummings MK, Glantz SA. Hospital admissions for acute myocardial infarction before and after implementation of a comprehensive smoke-free policy in Uruguay. *Tob Control*. 2013;22:e16-20. Medline:22337557 doi:10.1136/tobaccocontrol-2011-050134
- 71 Jan C, Lee M, Roa R, Herrera V, Politis M, Motta J. The association of tobacco control policies and the risk of acute myocardial infarction using hospital admissions data. *PLoS One*. 2014;9:e88784. Medline:24520421 doi:10.1371/journal.pone.0088784
- 72 Yildiz F, Baris SA, Basyigit I, Boyaci H, Aydinlik H, Sonmez PO. Role of smoke-free legislation on emergency department admissions for smoking-related diseases in Kocaeli, Turkey. *East Mediterr Health J*. 2015;20:774-80. Medline:25664515
- 73 Jha P, Guindon E, Joseph RA, Nandi A, John RJ, Rao K, et al. A rational taxation system of bidis and cigarettes to reduce smoking deaths in India. *Econ Polit Wkly*. 2011;42:44-51.



# Effect of physician characteristics and knowledge on the quality of dyslipidemia management and LDL-C target goal achievement in China: Subgroup analysis of the Dyslipidemia International Study

Rongjing Ding<sup>1</sup>, Ping Ye<sup>2</sup>, Shuiping Zhao<sup>3</sup>, Dong Zhao<sup>4</sup>, Xiaowei Yan<sup>5</sup>, Yugang Dong<sup>6</sup>, Jihu Li<sup>7</sup>, Yuqin Ran<sup>8</sup>, Dayi Hu<sup>1</sup>; on behalf of the DYSIS-China Study Investigators

<sup>1</sup> Department of Cardiology, Peking University People's Hospital, Beijing, China

<sup>2</sup> Department of Gerontology, Chinese PLA General Hospital, Medical School of Chinese PLA, Beijing, China

<sup>3</sup> Department of Cardiology, The Second Xiangya Hospital of Central South University, Changsha, China

<sup>4</sup> Institute of Epidemiology, Beijing Anzhen Hospital, Capital Medical University, Beijing, China

<sup>5</sup> Department of Cardiology, Peking Union Medical College Hospital, Beijing, China

<sup>6</sup> Department of Cardiology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China

<sup>7</sup> Outcome Research, Merck Sharp & Dohme (China) Holding Ltd., Shanghai, China

<sup>8</sup> Medical Affairs, Merck Sharp & Dohme (China) Holding Ltd., Shanghai, China

**Objective** This study aimed to investigate the effect of physicians' characteristics and knowledge of LDL-C target goals on the quality of lipid management in China.

**Methods** A total of 25 317 dyslipidemia patients who had taken lipid-lowering medication for >3 months were enrolled in our study. Patients' demographic data, medical history, lipid profile, their physician's specialty and professional title and their hospital level as well as their LDL-C goal opinions were recorded.

**Results** Questionnaires were completed by 926 physicians with 6 different specialties and 4 professional statuses, in 3 different-level hospitals. Most (74.5%) of the physicians recognized the importance of considering LDL-C serum concentration for treating dyslipidemia, and set target LDL-C goals according to the 2007 Chinese guidelines for 83.4% of their patients. The LDL-C goal achievement rate was significantly higher for patients whose physicians' knowledge of LDL-C target goals was consistent with guideline recommendations, compared with those whose physicians' knowledge was inconsistent with the guidelines (60.4% vs 31.1%,  $P < 0.0001$ ). Physicians working in tier 1 (odds ratio (OR) = 2.95; 95% CI 2.37–3.67), (OR = 1.56; 95% CI 1.34–1.81) and tier 2 (OR = 2.53; 95% CI 2.22–2.88), (OR = 1.16; 95% CI 1.06–1.27) hospitals, specialized in neurology (OR = 1.13; 95% CI 0.93–1.36), (OR = 1.57; 95% CI 1.40–1.77), internal medicine (OR = 1.07; 95% CI 0.90–1.27), (OR = 1.58; 95% CI 1.39–1.80), endocrinology (OR = 1.02; 95% CI 0.87–1.21), (OR = 1.63; 95% CI 1.47–1.82) and being a resident vs attending physician (OR = 1.05; 95% CI 0.92–1.20), (OR = 1.00; 95% CI 1.00–1.19) were independent risk factors for low knowledge of LDL-C target goals and low LDL-C goal achievement.

**Conclusion** Chinese physicians' characteristics and knowledge of LDL-C target goals were associated with patients' LDL-C goal achievement.

## Correspondence to:

Dayi Hu  
Department of Cardiology  
Peking University People's Hospital  
No. 11 Xizhimen South Street  
Xicheng District, Beijing, 100044  
China  
hudyai080615@163.com

An elevated level of serum cholesterol has been suggested to be the most important risk factor for ischemic cardiovascular disease (CVD) [1–7]. The Cholesterol Treatment Trialists' meta-analysis of statin clinical trials found that a 1 mmol/L reduction in LDL-C reduced major adverse CVD

events by approximately 20% [8,9]. Hence, prescription statin and LDL-C target goal achievement for patients with high- and very high-risk CVD is considered as an important strategy in CVD prevention and the epidemiology of LDL-C target goal achievement has been widely investigated. In 2014, the Dyslipidemia International Study in China (DYSIS-China) revealed that, after almost 10 years of cholesterol education for the primary and secondary prevention of CVD, 45% of high-risk and 60% of very high-risk dyslipidemia patients in China had not attained their LDL-C target goals [10]. Although these data reflect some improvement compared to previous reports that 69% of high-risk and 78% of very high-risk patients in China failed to attain their LDL-C goal between 2004 and 2006 [11], large gaps still exist, when compared with developed countries, regarding LDL-C goal achievement [12]. High-quality care ultimately comes from high-quality health professionals [13]. Previous studies have found that physicians play an important role in lipid management [14–17]. While physicians' attitude and behavior may affect the efficiency of lipid management, inadequate knowledge of LDL-C targets by physicians is also an important factor in dyslipidemia patients' failure to attain their LDL-C goals [6,18]. A previous study has reported that physicians' knowledge of lipid management was related to their professional status [19]. In China, the main characteristics of physicians include their professional status, as well as their specialty, and the quality and region of the hospital where they work. However, it is unclear how Chinese physicians' knowledge of guideline-recommended LDL-C targets and other characteristics affect their patients' LDL-C goal achievement. Therefore, we performed a subgroup analysis of DYSIS-China patients to evaluate the association between Chinese physicians' fundamental knowledge of lipid management and their patients' LDL-C goal achievement rate, and how it relates to the physicians' specialty and professional status, as well as their hospital's quality and location.

## METHODS

### Patients and study design

The DYSIS-China trial was an observational, cross-sectional, multicenter international study. Participants were recruited at 122 hospitals, including 58 tier 3 (teaching hospital), 31 tier 2 (territory hospital), and 33 tier 1 (community hospital) institutions in 6 representative regions of China between April 2012 and October 2012 (Table S1 in **Online Supplementary Document**). Patients who were 45 years of age or older and had been treated with a lipid-lowering drug for at least 3 months were included in our study. Data for each patient were collected from the baseline clinical examination, medical records, and a single outpatient follow-up visit. Lipid-lowering medications included statins, cholesterol absorption inhibitors, fibrates, nicotinic acid, and Xuezhikang. The 10-year risk of CVD (10YRCVD) for each patient was classified as follows, based on the 2007 Chinese Guidelines on Prevention and Treatment of Dyslipidemia in Adults: (a) very high-risk (CVD with diabetes or acute coronary syndrome); (b) high-risk (coronary heart disease with a 10YRCVD of 10% to 15%); (c) moderate risk (10YRCVD of 5% to 10%); and (d) low-risk (10YRCVD of <5%).

All of the participating physicians were asked to complete a questionnaire that asked whether they regarded guideline recommendations for LDL-C as an important clinical reference for lipid management. If the physician answered yes, he/she was asked to select an LDL-C target goal based on the patient's relevant risk category, which was defined in the 2007 Chinese lipid management guidelines as follows: (a) <4.14 mmol/L (<160 mg/dL) for low-risk patients; (b) <3.37 mmol/L (<130mg/dL) for moderate-risk patients; (c) <2.59 mmol/L (<100 mg/dL) for high-risk patients; (d) <2.07 mmol/L (<80 mg/dL) for very high-risk patients; (e) other; and (f) not sure. Consistency between each physician-suggested LDL-C target goal and the LDL level recommended by the 2007 guidelines was defined as follows: (a) Yes (consistent), if the physician-suggested LDL-C target goal was lower than or equal to the LDL-C level recommended by the guidelines. (b) No (not consistent), if the physician's suggested LDL-C target goal was higher than that recommended by the guidelines. Data for each physician category including professional status (professor, associate professor, attending physician, or resident physician), specialty (including internal medicine, general medicine, cardiology, neurology, endocrinology, and geriatrics), as well as the level of their hospital (tier 1, tier 2, or tier 3), were collected. Each patient provided written informed consent before participation, and our study protocol was approved by the ethics committee of each participating hospital. All of the hospitals and physicians consented to the publication of information related to hospital status and the treating physicians' specialties and professional titles.

## Statistical analysis

All of the statistical analyses were performed using the SAS, version 9.1, software (SAS Institute, Cary, NC, USA). A hypothesis test could not be used for the principal analysis, due to the primarily descriptive nature of our study design. Categorical variables are presented as absolute numbers and percentages. Continuous variables are reported as the mean  $\pm$  standard deviation (SD). Intergroup differences in categorical variables were evaluated using chi-square analysis or Fisher exact test, depending on the number of patients in each group. Intergroup differences in continuous variables were evaluated using an analysis of variance. Logistic regression models were used to calculate the odds ratio and 95% confidence interval for the probability of consistency between the physician-suggested and guideline-recommended LDL-C target levels based on hospital status, type of physician or department, and adjusting for the 10YRCVD category, comorbidities, or other risk factors, including hypertension, body mass index (BMI)  $\geq 28$  kg/m<sup>2</sup>, male  $>45$  years or female  $>55$  years, smoking status, family history of premature CVD, and hospital level, as well as the physician's specialty and professional title. All of the evaluations were 2-tailed, and results with  $P < 0.05$  were considered to be statistically significant.

## RESULTS

### Baseline patient characteristics

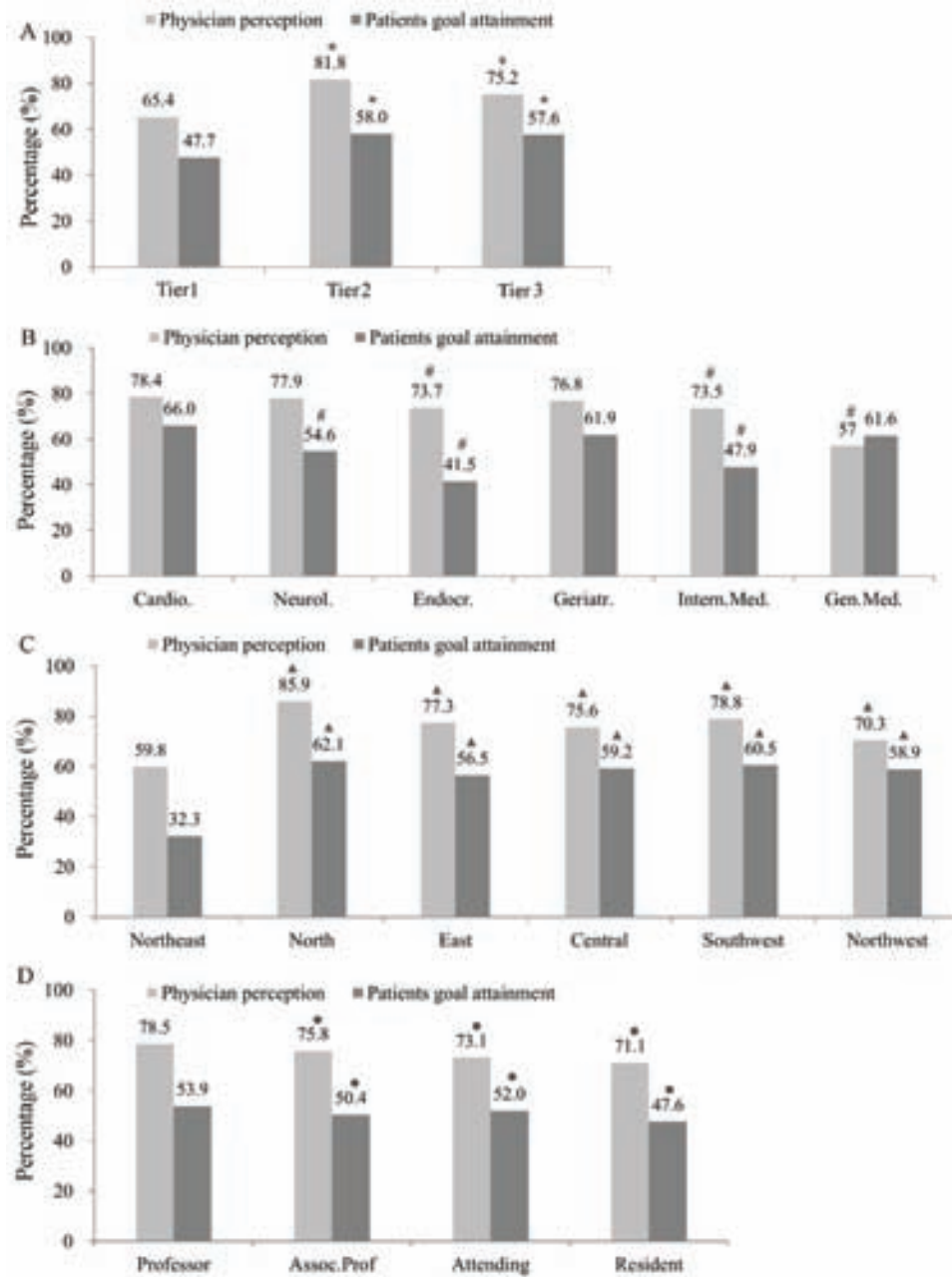
A total of 25 317 dyslipidemia patients (51.3% male) with a mean age of 65.4 years were included in our study. Most of the patients (97.5%) were Han Chinese. The most prevalent comorbidities were hypertension (65.8%), diabetes mellitus (34.8%), and coronary heart disease (37.2%). Behavioral risk factors included sedentary lifestyle (19.7%) and smoking (12.4%). The distribution of patients based on 10YRCVD was 12.2% for the very high-risk, 58.9% for high-risk, 11.0% for moderate-risk, and 17.9% for the low-risk groups. The normal dosage of statin is equivalent to simvastatin 20–40 mg/d (71%) (Table 1). The LDL-C goal achievement rates for each category are listed in Table 1.

### Physician characteristics

A total of 926 physicians participated in our study, 168 (18.1%) of whom were from tier 1 hospitals, 199 (21.5%) from tier 2 hospitals, and 559 (60.4%) from tier 3 hospitals. The physicians' specialties were as follows: 228 (24.6%) cardiology, 156 (16.3%) neurology, 185 (19.3%) endocrinology, 113 (12.2%) geriatrics, 210 (22.7%) internal medicine, and 34 (3.7%) general medicine. At tier 1 hospitals, all of the physicians were from the internal medicine department, 73.8% of whom were attending physicians and residents. Most physicians at tier 2 hospitals were associate professors or attending physicians from the cardiology or internal medicine departments, followed by the endocrinology, neurology, and geriatric departments. In tier 3 hospitals, more physicians were professors in all departments, compared with tier 2 and tier 1 hospitals (24.7% vs 11.1%  $P < 0.0001$  and 24.7 vs 6.5%  $P < 0.0001$ , respectively). Also the distribution of physicians' status differed between some hospital departments particularly within the Tier 2 and tier 3 groups (Table 2).

### Physician characteristics and knowledge of LDL-C targets as a clinical reference for LDL-C goal achievement

The proportion of physicians who considered the guideline-recommended LDL-C goal as an important clinical reference was different depending on physician characteristics, and was higher at tier 2 and tier 3 hospitals. It was also somewhat higher among physicians who specialized in cardiology, neurology, and geriatrics. Physicians' low knowledge of LDL-C goal as a clinical reference predicted lower LDL-C goal achievement. At tier 1, tier 2 and tier 3 hospitals, 65.4%, 81.8% and 75.2% of the physicians recognized that LDL-C goal is important in clinical practice, the overall goal achievement rates were only 47.7%, 58.0% and 57.6%, respectively (Figure 1A, Table S2 in Online Supplementary Document), and the lowest rate of LDL target goal attainment was in tier 1 hospitals. Regarding different specialties, 73.5% to 78.4% of the physicians recognized that LDL-C goal achievement is important in clinical practice, and the goal achievement rate ranged from 41.5% to 66.0% across the various specialties. The greatest difference between patients' goal attainment and physician perception was found to be among endocrinologists (Figure 1B, Table S2 in Online Supplementary Document). The lowest LDL-C goal achievement rate and the least knowledge by physicians regarding LDL-C as an important clinical reference existed in the Northeast (Figure 1C, Table S2 in Online Supplementary Document), and among resident physicians (Figure 1D, Table S2 in Online Supplementary Document).



**Figure 1.** Comparison of questionnaire results regarding whether physicians recognized that the guideline–recommended low–density lipoprotein cholesterol (LDL–C) goal is an important clinical reference for dyslipidemia management and the LDL–C goal achievement rates based on (A) hospital level, (B) physician specialty, (C) professional status, and (D) physician perception. \* $P < 0.05$ , compared to Tier1, # $P < 0.05$ , compared to Cardiology department, ▲ $P < 0.05$ , compared to Northeast, ● $P < 0.05$ , compared to Professor title.

### Effects of physician characteristics and knowledge of guideline–recommended LDL–C targets on LDL–C goal achievement

Our study showed that less than 75% of physicians in China are familiar with guideline–recommended LDL–C targets, regardless of physician characteristics. There was an association between the rate of LDL–C goal achievement by patients and their physicians’ knowledge of LDL–C targets. The concordance between physicians’ knowledge of LDL–C targets and guideline recommendations was the lowest in tier 1 hospitals (Figure 2A, Table S3 in Online Supplementary Document), endocrinology departments. (Figure 2B, Table S3 in Online Supplementary Document), Northeast China (Figure 2C, Table S3 in Online Supplementary Document) and among resident physicians (Figure 2D, Table S3 in Online Supplementary Document).

**Table 1.** Patients' basic characteristics and LDL-C goal achievement rate

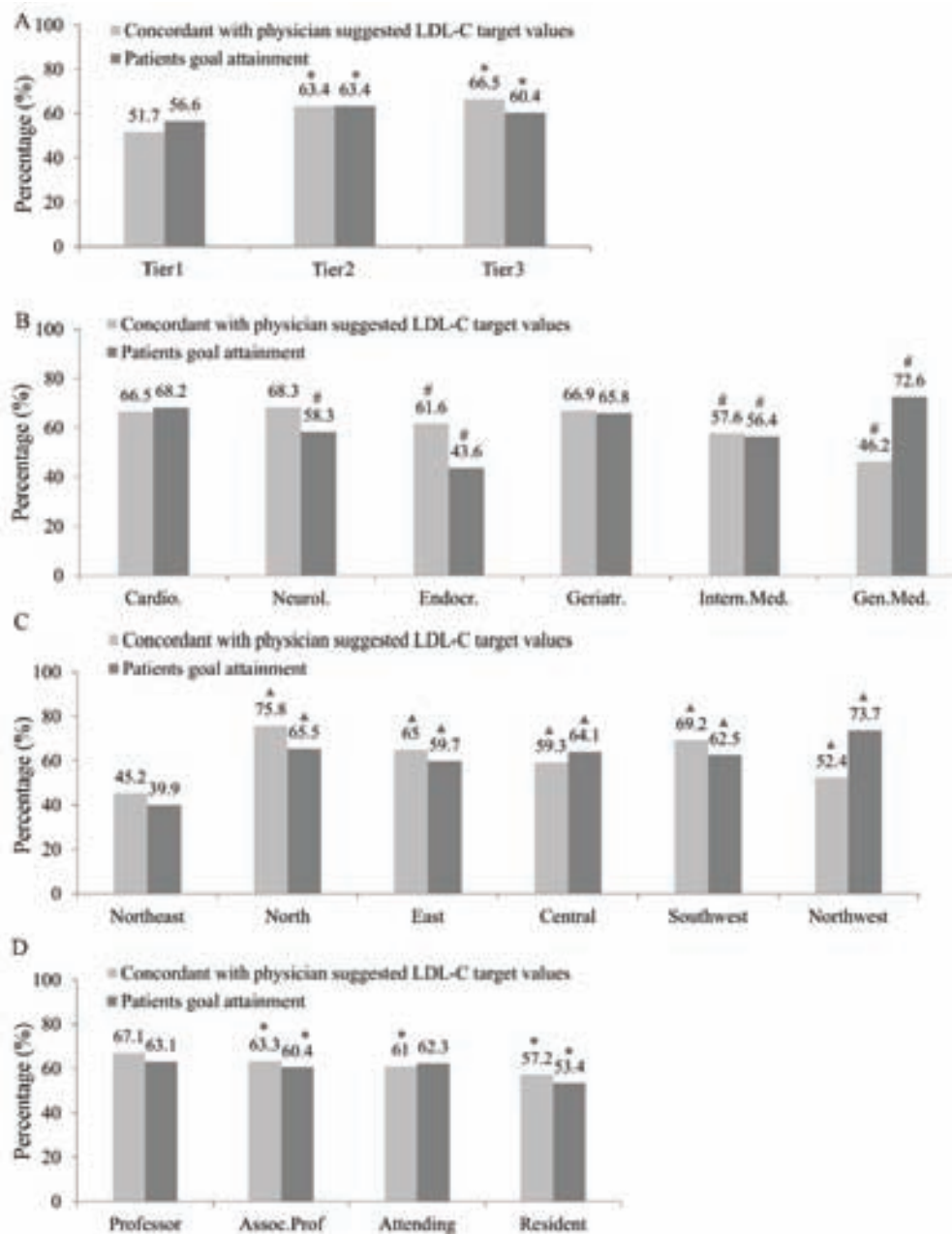
	NUMBER (%) N = 25 317	LDL-C GOAL ACHIEVEMENT N (%)
<b>Age category:</b>		
≥65 years	13 100 (51.7%)	7803 (59.6%)
<b>Gender:</b>		
Male	12 975 (51.3%)	8431 (65.0%)
Female	12 342 (48.7%)	7140 (57.9%)
<b>Coronary heart disease</b>	9420 (37.2%)	5247 (55.7%)
<b>Cerebrovascular disease</b>	4281 (16.9%)	2304 (53.8%)
<b>Peripheral arterial disease</b>	263 (1.0%)	153 (58.2%)
<b>Chronic kidney disease</b>	1846 (7.3%)	1095 (59.3%)
<b>Diabetes mellitus</b>	8798 (34.8%)	3978 (45.3%)
<b>Hypertension</b>	16650 (65.8%)	10075 (60.5%)
<b>Other cardiovascular risk factors:</b>		
<b>Smoking:</b>		
Current	3143 (12.4%)	1902 (60.5%)
Quit	4445 (17.6%)	2881 (64.8%)
Never	17 729 (70.0%)	10788 (60.9%)
HDL-C <1.04mmol/L	6682 (26.4%)	4468 (68.9%)
Family history of CHD	2294 (9.1%)	1381 (60.2%)
BMI >28kg/m <sup>2</sup>	3457 (13.7%)	1947 (56.3%)
Family history of early onset of ischemic cardiovascular disease	2294 (9.1%)	1381 (60.2%)
Sedentary lifestyle	4997 (19.7%)	2962 (59.3%)
<b>10YRCVD category:</b>		
Very high	3092 (12.2%)	1226 (40.0%)
High	14 916 (58.9%)	8174 (54.8%)
Moderate	2782 (11.0%)	2041 (73.4%)
Low	4527 (17.9%)	4130 (91.2%)
<b>Statin dosage potency:*</b>		
Potency 1	299 (1.2%)	192/299 (64.2%)
Potency 2	2605 (10.3%)	1512/2605 (58.0%)
Potency 3	9958 (39.3%)	6014/9958 (60.4%)
Potency 4	7179 (28.4%)	4547/7179 (63.3%)
Potency 5	1725 (2.9%)	1083/1725 (62.8%)
Potency 6	96	57/96 (59.4%)

HDL-C – high density lipoprotein cholesterol, CHD – chronic disease, 10YRCVD – 10-y risk of cardiovascular disease

\*Statin dosage level: Potency1 is equivalent to simvastatin 5 mg/d; Potency 2 is equivalent to simvastatin 10 mg/d; Potency 3 is equivalent to simvastatin 20 mg/ day; Potency 4 is equivalent to simvastatin 40 mg/d; Potency 5 is equivalent to simvastatin 80 mg/d; Potency 6 is equivalent to Atorvastatin 80 mg/d.

### Effects of physicians' knowledge of LDL-C as a clinical reference and guideline-recommended LDL-C targets on the rate of LDL-C goal achievement

In total, 74.5% of physicians recognized that LDL-C is an important clinical reference. The LDL-C goal achievement rate was significantly higher among those patients whose physicians' perceptions of the LDL-C target goal were consistent with guideline recommendations (60.4% vs 31.1%,  $P < 0.0001$ , **Table 3**). The LDL-C goal achievement rates in patients with very high, high, and moderate risk whose physicians' knowledge of the LDL-C target goal was consistent with guideline recommendations were 39.9%, 51.5%, and 70%, respectively. The corresponding rates among patients whose physicians' knowledge of the LDL-C target goal was inconsistent with guideline recommendations were 29.7%, 32.1%, and 30.1%, respectively ( $P < 0.0001$  for all). These results suggest that the LDL-C goal achievement rate significantly correlated with physicians' knowledge of the guideline-recommended LDL-C target goal.



**Figure 2.** LDL-C goal achievement rate and consistency between physicians' knowledge of LDL-C target goal and guideline recommendations based on (A) hospital level, (B) physician specialty, (C) China's geographic regions, and (D) professional status. \* $P < 0.05$ , compared to Tier 1, # $P < 0.05$ , compared to Cardiology department, ▲ $P < 0.05$ , compared to Northeast, ● $P < 0.05$ , compared to Professor title.

### Predictors of inconsistent physicians' knowledge with guideline-recommended LDL-C target goals using multivariate logistic regression analysis

Multivariate logistic regression analysis revealed that predictors of inconsistency between physicians' perceptions of the LDL-C target goal and 2007 Chinese guideline recommendations included patients diagnosed with diabetes, coronary artery disease, cerebrovascular disease, and peripheral arterial disease. Other predictors of inconsistency between physicians' perceptions of the LDL-C target goal and 2007 Chinese guideline recommendations included the following physician characteristics: working in a tier 1 or tier 2 hospital, specializing in neurology, endocrinology, geriatrics, or general medicine, and being a resident physician. In addition, there were differences regarding the regions (Table 4).

**Table 2.** Characteristics of physicians based on hospital status and specialty

HOSPITAL STATUS	NUMBER OF PHYSICIANS (%)						
	Cardiology	Neurology	Endocrinology	Geriatrics	Internal Medicine	General Medicine	Total*
Tier 1	Professor			0	10 (7.5%)	1 (2.9%)	11 (6.5%) <sup>a</sup>
	Assoc.Prof.			0	26 (19.4%)	7 (20.6%)	33 (19.6%) <sup>a</sup>
	Attending			0	50 (37.3%)	17 (50.0%)	67 (39.9%) <sup>a</sup>
	Resident			0	48 (35.8%)	9 (26.5%)	57 (33.9%) <sup>a</sup>
	Total			0	134 (100.0%)	34 (100.0%)	168 (100.0%)
Tier 2	Professor	7 (9.3%)	4 (36.4%)	4 (11.4%)	0	7 (9.2%)	22 (11.1%) <sup>a</sup>
	Assoc.Prof.	19 (25.3%)	3 (27.3%)	11 (31.4%)	2 (100.0%)	22 (28.9%)	57 (28.6%) <sup>b</sup>
	Attending	30 (40.0%)	1 (9.1%)	10 (28.6%)	0	30 (39.5%)	71 (35.7%) <sup>a</sup>
	Resident	19 (25.3%)	3 (27.3%)	10 (28.6%)	0	17 (22.4%)	49 (24.6%) <sup>b</sup>
	Total	75 (100.0%)	11 (100.0%)	35 (100.0%)	2 (100.0%)	76 (100.0%)	199 (100.0%)
Tier 3	Professor	41 (26.8%)	34 (23.4%)	28 (18.7%)	35 (31.5%)		138 (24.7%) <sup>b</sup>
	Assoc.Prof.	47 (30.7%)	47 (32.4%)	42 (28.0%)	24 (21.6%)		160 (28.6%) <sup>b</sup>
	Attending	46 (30.1%)	45 (31.0%)	50 (33.3%)	41 (36.9%)		182 (32.6%) <sup>a</sup>
	Resident	19 (12.4%)	19 (13.1%)	30 (20.0%)	11 (9.9%)		79 (14.1%) <sup>c</sup>
	Total	153 (100.0%)	145 (100.0%)	150 (100.0%)	111 (100.0%)		559 (100.0%)
Total	Professor	48 (21.1%)	38 (24.4%)	32 (17.3%)	35 (31.0%)	17 (8.1%)	171 (18.5%)
	Assoc.Prof.	66 (28.9%)	50 (32.1%)	53 (28.6%)	26 (23.0%)	48 (22.9%)	250 (27.0%)
	Attending	76 (33.3%)	46 (29.5%)	60 (32.4%)	41 (36.3%)	80 (38.1%)	320 (34.6%)
	Resident	38 (16.7%)	22 (14.1%)	40 (21.6%)	11 (9.7%)	65 (31.0%)	185 (20.0%)
	Total	228 (100.0%)	156 (100.0%)	185 (100.0%)	113 (100.0%)	210 (100.0%)	926 (100.0%)

Assoc. Prof – Associate Professor

\*Indicates containing sub-groups which are marked with a, b and c. Different letter marks indicate significant differences of the indicated professional title group within departments of the hospitals, same letters indicate no significant differences ( $P > 0.05$ ).

**Table 3.** LDL-C goal achievement rates according to physicians' knowledge of LDL-C as a clinical reference and consistency between physicians' acceptance of LDL-C target goal and guideline recommendations based on 10YRCVD\*

	LDL-C GOAL ACHIEVEMENT (YES/TOTAL) FOR 10YRCVD CATEGORIES					
		Very high	High	Moderate	Low	Total
Total		1226/3092 (39.7%)	8174/14916 (54.8%)	2041/2782 (73.4%)	4130/4527 (91.2%)	15571/25317 (61.5%)
LDL-C as a clinical reference	Yes (74.5%)	836/2363 (35.4%)	5411/11222 (48.2%)	1393/2054 (67.8%)	2852/3231 (88.3%)	10492/18870 (55.6%)
	No	164 (29.8%)	780 (27.4%)	138 (27.3%)	305 (34.0%)	1387 (28.9%)
	P-value	NA	NA	NA	NA	0.0014
Consistency with guideline-recommended LDL-C target goal <sup>†</sup>	Yes (83.4%)	515/1290 (39.9%)	4777/9270 (51.5%)	1358/1940 (70%)	2852/3231 (100%)	9502/15731 (60.4%)
	No <sup>‡</sup>	318/1069 (29.7%)	624/1941 (32.1%)	34/113 (30.1%)	NA	976/3123 (31.3%)
	P-value	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$	NA	$P < 0.0001$

HDL-C – high density lipoprotein cholesterol, CHD – chronic disease, 10YRCVD – 10-year risk of cardiovascular disease

\*10YRCVD categories were based on serum LDL-C levels of  $< 2.07$  mmol/L ( $< 80$  mg/dL) for very high risk,  $< 2.59$  mmol/L ( $< 100$  mg/dL) for high risk,  $< 3.37$  mmol/L ( $< 130$  mg/dL) for moderate risk, and  $< 4.14$  mmol/L ( $< 160$  mg/dL) for low risk.

<sup>†</sup>Included patients in whom the risk classification was lower than guideline recommendations.

<sup>‡</sup>2007 Chinese lipid management guidelines as the reference.

## Predictors of failure for not achieving guideline-recommended LDL-C target goals using multivariate logistic regression analysis

The result showed that predictors of failure to achieve the LDL-C goal included patients with diabetes, coronary artery disease, cerebrovascular disease, BMI  $> 28$  kg/m<sup>2</sup> and male  $\geq 45$  years or female  $\geq 55$  years, physician working in a tier 1 or tier 2 hospital, specializing in neurology, endocrinology or general medicine, and being a resident physician (Table 5).

**Table 4.** Multivariate logistic regression analysis of consistence between physician suggested–LDL–C target goal and recommendations of 2007 Chinese guidelines

	PARAMETER ESTIMATION VALUE	OR	95% CI	P-VALUE
<b>Comorbidity (yes vs no):</b>				
Diabetes mellitus	1.24	3.47	3.15–3.83	<0.0001
Coronary heart disease	1.33	3.78	3.42–4.18	<0.0001
Cerebrovascular disease	0.41	1.51	1.33–1.70	<0.0001
Peripheral arterial disease	1.08	2.94	2.13–4.06	<0.0001
Chronic kidney disease	–0.07	0.93	0.80–1.09	0.3755
<b>Risk factors:</b>				
Hypertension (yes vs no)	0.04	1.05	0.94–1.16	0.4118
BMI >28kg/m <sup>2</sup>	–0.01	0.99	0.87–1.13	0.9256
Male ≥45 years or Female ≥55 years (yes vs no)	0.22	1.24	1.03–1.50	0.0227
Smoking (yes vs no)	–0.11	0.89	0.77–1.03	0.1196
Family history of early onset of ischemic cardiovascular disease (yes vs no)	–0.13	0.88	0.75–1.03	0.1198
<b>Hospital level:</b>				
Tier 1 vs Tier 3	1.08	2.95	2.37–3.67	<0.0001
Tier 2 vs Tier 3	0.93	2.53	2.22–2.88	<0.0001
<b>Specialty:</b>				
Neurology vs Cardiology	0.12	1.13	0.93–1.36	0.2180
Endocrinology vs Cardiology	0.02	1.02	0.87–1.21	0.7733
Geriatrics vs Cardiology	0.24	1.27	1.06–1.53	0.0112
Internal medicine vs Cardiology	0.06	1.07	0.90–1.27	0.4649
General medicine vs Cardiology	–0.03	0.97	0.73–1.30	0.8352
<b>Professional title:</b>				
Professor vs attending physician	–0.06	0.95	0.83–1.08	0.4092
Associate professor vs attending physician	–0.11	0.90	0.80–1.01	0.0755
Resident vs attending physician	–0.05	1.05	0.92–1.20	0.4454
<b>Region:</b>				
North vs Northeast	–1.14	0.32	0.27–0.38	<0.0001
East vs Northeast	–0.71	0.49	0.42–0.58	<0.0001
Central vs Northeast	–0.16	0.85	0.73–0.99	0.0346
Southwest vs Northeast	–1.01	0.37	0.31–0.43	<0.0001
Northwest vs Northeast	–0.60	0.55	0.46–0.65	<0.0001

LDL–C – low density lipoprotein cholesterol, OR – odds ratio; CI – confidence interval, BMI – body mass index

## DISCUSSION

The results of our study are consistent with previously published reports, which found that more than 60% of very high–risk patients and over 45% of high–risk patients did not attain their LDL–C goals, according to the 2007 Chinese lipid management guidelines, after at least 3 months of lipid–lowering treatment [10]. We found that physicians' knowledge of guideline–recommended LDL–C target goals significantly correlated with patients' LDL–C target goal achievement in China. Physician's knowledge varied depending on the hospital level, medical specialty, professional status, and geographic region. Working in a tier 1 or tier 2 hospital, being a resident physician, and specializing in neurology, endocrinology, or internal medicine were independent risk factors for lower LDL–C target goal achievement. Our results indicate that physician knowledge and characteristics affect the quality of lipid management in China, and more attention should be paid to this fact.

Previous studies have suggested that both patient– and physician–dependent factors contributed to LDL–C goal achievement failure [20–22]. Since the start of physician–initiated LDL–C treatment, more attention has been paid to studying the role of physicians in LDL–C goal achievement. Behavior change theory [23] suggests that knowledge should be the most important factor for behavior change. According to a global survey of physicians' perceptions of serum cholesterol management in 10 countries, 80.0% of

**Table 5.** Multivariate logistic regression analysis of LDL-C goal achievement rates

	PARAMETER ESTIMATION VALUE	OR	95% CI	P-VALUE
<b>Comorbidity (yes vs no):</b>				
Diabetes	1.08	2.94	2.75–3.14	1.0783
Coronary artery disease	0.61	1.84	1.72–1.97	0.6109
Cerebrovascular disease	0.29	1.33	1.22–1.44	0.2850
Peripheral arterial disease	–0.09	0.92	0.70–1.21	–0.0857
<b>Risk factors:</b>				
Hypertension (yes vs no)	0.01	1.01	0.95–1.08	0.7462
BMI >28kg/m <sup>2</sup>	0.19	1.21	1.11–1.31	<0.0001
Male ≥45 years or Female ≥55 years (yes vs no)	0.14	1.15	1.03–1.29	0.0104
Smoking (yes vs no)	0.01	1.01	0.93–1.11	0.8020
Family history of early onset of ischemic cardiovascular disease (yes vs no)	0.04	1.04	0.94–1.15	0.4729
<b>Hospital level:</b>				
Tier 1 vs Tier3	0.44	1.56	1.34–1.81	0.4427
Tier 2 vs Tier3	0.15	1.16	1.06–1.27	0.1500
<b>Specialty:</b>				
Neurology vs Cardiology	0.45	1.57	1.40–1.77	<0.0001
Endocrinology vs Cardiology	0.49	1.63	1.47–1.82	<0.0001
Geriatrics vs Cardiology	0.30	1.35	1.20–1.51	<0.0001
Internal Medicine vs Cardiology	0.46	1.58	1.39–1.80	<0.0001
General Medicine vs Cardiology	–0.09	0.91	0.75–1.11	0.3513
<b>Professional title:</b>				
Professor vs attending physician	–0.09	0.92	0.84–1.00	0.0420
Associate professor vs attending physician	–0.01	0.92	0.92–1.08	0.9055
Resident vs attending physician	0.09	1.00	1.00–1.19	0.0329
<b>Region:</b>				
North vs Northeast	–0.86	1.10	0.38–0.47	<0.0001
East vs Northeast	–0.85	0.42	0.39–0.48	<0.0001
Central vs Northeast	–0.69	0.43	0.45–0.56	<0.0001
Southwest vs Northeast	–0.89	0.50	0.37–0.46	<0.0001
Northwest vs Northeast	–0.87	0.41	0.38–0.47	<0.0001

LDL-C – low density lipoprotein cholesterol, OR – odds ratio; CI – confidence interval, BMI – body mass index

physicians were concordant with dyslipidemia-management guidelines with regard to setting LDL-C goals for their patients, and 61.0% of them believed that a sufficient number of their patients attained their serum cholesterol goals. However, only 47.0% of their patients reached and maintained their serum cholesterol goals, and physicians expressed their frustration regarding their inability to effectively treat some CVD patients [24]. Another recent study in Croatia reported that 80.6% of physicians believed that they provided effective treatment to their dyslipidemia patients, but only 53.3% could accurately state the target LDL-cholesterol values for high-risk patients without consulting the most recent guidelines [25]. These reports suggest that physicians' knowledge of LDL-C target goals is an important factor for patients' LDL-C goal achievement.

In the current study, by evaluating correlations between physicians' knowledge of guideline-recommended LDL-C levels and LDL-C goal achievement, we aimed to investigate the physician's role in failing to attain recommended LDL-C goals in Chinese patients who were treated with lipid-lowering medications. We found that 74.5% of Chinese physicians recognized the importance of considering LDL-C serum concentration for treating dyslipidemia, and set target LDL-C goals that were concordant with the 2007 Chinese guidelines for 83.4% of their patients. Further analysis showed that 55.6% of these patients attained their serum LDL-C goal, indicating a success rate of 66.7%. We also found that the LDL-C goal achievement rate was significantly higher for patients whose physicians' knowledge of the LDL-C target goals was consistent with guideline recommendations, compared with those whose physicians' knowledge was inconsistent with the guidelines (60.4% vs 31.1%,  $P < 0.0001$ ). This finding supports the idea that enhancing Chinese physicians' awareness of LDL-C target goals may improve the quality of lipid control in China. We also found that hospital level, as well as medical specialty and professional status were significantly related to physicians' knowledge and their patients' LDL-C goal achievement. Physicians from tier 1 and tier 2 hospitals, who were specializing in endocrinology, internal medicine, or general medicine,

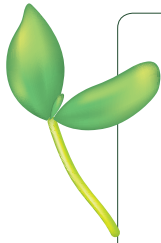
and resident physicians demonstrated lower awareness of guideline–recommended LDL–C target goals than their colleagues. Multivariate logistic regression analysis revealed a relatively strong correlation between physician characteristics and physician–suggested LDL–C target goals and guideline–recommended LDL–C goals. These findings suggest that certain physician characteristics and inadequate knowledge of guideline–recommended LDL–C goals may be important reasons for patients’ failure to attain their LDL–C target goals in China. Guideline–recommended LDL–C goals should be considered among the most important pieces of information delivered to physicians. These results also suggest that continuing medical education strategy in China should be based on physician characteristic.

This is the first study to investigate factors underlying suboptimal LDL–C goal achievement rates in China from the physicians’ perspective in an innovative attempt to improve the quality of cardiovascular disease prevention in China. There are, however, certain limitations to our study. Our results point to inadequate awareness among Chinese physicians regarding lipid control. However, the questionnaire completed by physicians addressed their knowledge of LDL–C only. Thus, our results do not provide a comprehensive view of all aspects of the physicians’ skills regarding lipid management. Furthermore, due to our focus on physicians’ characteristics and lipid knowledge in our study, we did not collect data on all potential confounding factors, such as the physician’ age, gender, and behavior regarding lipid management, as well as patients’ lifestyle habits that might have influenced LDL–C goal achievement. However, there are some advantages to our study. First, this is a large–scale cohort study, which included patients’ and physicians’ medical information together, related to lipid medication; therefore, we can directly analyze the relationship between physicians’ knowledge and patients’ LDL–C goal achievement. Second, the physicians’ information in our cohort was collected with the randomized cohort stratified sample method, so it is representative of physician characteristics throughout China. Physician characteristics included their hospital level, specialty, professional status, and service regions, so we can clearly tell the effect of physician characteristics on knowledge of guideline–recommended LDL–C target goals and on LDL–C goal achievement. Thus, we were able to uncover shortcomings in China regarding physician characteristics and knowledge. Although we chose only 2 questions, which focused on LDL–C, those 2 questions are fundamental to lipid control relating to physician behavior. In this study, we used 2007 Chinese guideline–recommended LDL–C target goals and did not use the updated 2013 ESC guidelines. The reason for this was that LDL–C target goals in the Chinese guidelines were the most widely disseminated in China. This way we could tell whether or not the effect of knowledge of LDL–C target goals on LDL–C target achievement related to physician characteristics.

Our study has identified some shortcomings in health services in Chinese community medical centers (tier 1 hospitals) regarding physician characteristics and inadequate knowledge. The lower percentage of specialists and professors in community medical centers result in less knowledge and lower LDL–C target achievement. This may be just the tip of the iceberg for medical service shortcomings in China. Since community hospitals have been considered as the first bastion of chronic disease prevention in China, efficient work by community physicians is of great importance for the health of the population [19]. It is imperative to improve the level of health services in community medical centers. Arduous educational efforts must be ongoing in Chinese community medical centers, and more attention should be paid to physicians’ knowledge of LDL–C target goals.

## CONCLUSIONS

The results of our study suggest that LDL–C goal achievement in dyslipidemia patients in China is associated with physicians’ characteristics and their knowledge of LDL–C guidelines. Deficiencies in physicians’ knowledge of LDL–C targets reveal serious shortcomings in health services in China. Strategies aimed at enhancing familiarity with and acceptance of LDL–C guidelines by Chinese physicians based on the region, professional status, medical specialty, and hospital level may be an efficient way to improve both guideline adherence and LDL–C goal achievement.



**Acknowledgments:** We sincerely thank Dr Philippe Brudi for his great contribution to the global DYSIS study protocol design.

**Ethics approval:** Our study protocol was approved by the ethics committee of each participating hospital

**Funding:** This study was funded by a research grant from Merck & Co., Ltd (IISP#39298).

**Authorship contributions:** Dr Hu Dayi conceived and initiated the study, supervised its conduct and data analysis and had primary responsibility for writing the report. Dr Ding Rongjing wrote the report of this study. Dr Ding Rongjing, Dr Ye Ping, Zhao Shuiping, Dr Zhao Dong, and Dr Yan Xiaowei contributed to the conception, design, and assembly of all data. All of the authors reviewed the report and provided critical input for its revision.

**Competing interests:** The authors have completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author) and declare no conflicts of interests.

## REFERENCES

- 1 Expert Panel on Detection E. Treatment of High Blood Cholesterol in A. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *JAMA*. 2001;285:2486-97. Medline:11368702 doi:10.1001/jama.285.19.2486
- 2 Graham I, Atar D, Borch-Johnsen K, Boysen G, Burell G, Cifkova R, et al. European guidelines on cardiovascular disease prevention in clinical practice: executive summary. Fourth Joint Task Force of the European Society of Cardiology and other societies on cardiovascular disease prevention in clinical practice (constituted by representatives of nine societies and by invited experts). *Eur J Cardiovasc Prev Rehabil*. 2007;14 Suppl 2:E1-40. Medline:17726406 doi:10.1097/01.hjr.0000277984.31558.c4
- 3 Grundy SM, Cleeman JI, Merz CN, Brewer HB Jr, Clark LT, Hunninghake DB, et al. Implications of recent clinical trials for the National Cholesterol Education Program Adult Treatment Panel III guidelines. *Circulation*. 2004;110:227-39. Medline:15249516 doi:10.1161/01.CIR.0000133317.49796.0E
- 4 National Cholesterol Education Program Expert Panel on Detection E. Treatment of High Blood Cholesterol in A. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. *Circulation*. 2002;106:3143-421. Medline:12485966
- 5 Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97:1837-47. Medline:9603539 doi:10.1161/01.CIR.97.18.1837
- 6 Sager HB, Linsel-Nitschke P, Mayer B, Lieb W, Franzel B, Elsasser U, et al. Physicians' perception of guideline-recommended low-density lipoprotein target values: characteristics of misclassified patients. *Eur Heart J*. 2010;31:1266-73. Medline:20219745 doi:10.1093/eurheartj/ehq026
- 7 Wagner AM, Perez A, Calvo F, Bonet R, Castellvi A, Ordonez J. Apolipoprotein(B) identifies dyslipidemic phenotypes associated with cardiovascular risk in normocholesterolemic type 2 diabetic patients. *Diabetes Care*. 1999;22:812-7. Medline:10332687 doi:10.2337/diacare.22.5.812
- 8 Baigent C, Keech A, Kearney PM, Blackwell L, Buck G, Pollicino C, et al. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet*. 2005;366:1267-78. Medline:16214597 doi:10.1016/S0140-6736(05)67394-1
- 9 Cholesterol Treatment Trialists C, Kearney PM, Blackwell L, Collins R, Keech A, Simes J, et al. Efficacy of cholesterol-lowering therapy in 18,686 people with diabetes in 14 randomised trials of statins: a meta-analysis. *Lancet*. 2008;371:117-25. Medline:18191683 doi:10.1016/S0140-6736(08)60104-X
- 10 Zhao S, Wang Y, Mu Y, Yu B, Ye P, Yan X, et al. Prevalence of dyslipidaemia in patients treated with lipid-lowering agents in China: results of the DYSlipidemia International Study (DYSIS). *Atherosclerosis*. 2014;235:463-9. Medline:24950001 doi:10.1016/j.atherosclerosis.2014.05.916
- 11 Wu YF The second multi-center survey of dyslipidemia management in China: goal attainment rate and related factors. *Zhonghua Xin Xue Guan Bing Za Zhi*. 2007;35:420-7. Medline:17711683
- 12 Gitt AK, Lautsch D, Ferrieres J, Kastelein J, Drexel H, Horack M, et al. Low-density lipoprotein cholesterol in a global cohort of 57,885 statin-treated patients. *Atherosclerosis*. 2016;255:200-9. Medline:27667299 doi:10.1016/j.atherosclerosis.2016.09.004
- 13 How to attain the ambitious goals for health reform in China. *Lancet*. 2015;386:1419. Medline:26466023 doi:10.1016/S0140-6736(15)00452-3
- 14 Yarzebski J, Bujor CF, Goldberg RJ, Spencer F, Lessard D, Gore JM. A community-wide survey of physician practices and attitudes toward cholesterol management in patients with recent acute myocardial infarction. *Arch Intern Med*. 2002;162:797-804. Medline:11926854 doi:10.1001/archinte.162.7.797
- 15 Cacoub P, Tocque-Le Gousse E, Fabry C, Hermant S, Petzold L. Application in general practice of treatment guidelines for patients with dyslipidaemia: the RESPECT study. *Arch Cardiovasc Dis*. 2008;101:715-21. Medline:19059566 doi:10.1016/j.acvd.2008.09.011

- 16 Davidson MH, Maki KC, Pearson TA, Pasternak RC, Deedwania PC, McKenney JM, et al. Results of the National Cholesterol Education (NCEP) Program Evaluation Project Utilizing Novel E-Technology (NEPTUNE) II survey and implications for treatment under the recent NCEP Writing Group recommendations. *Am J Cardiol.* 2005;96:556-63. Medline:16098311 doi:10.1016/j.amjcard.2005.04.019
- 17 Frolkis JP, Pearce GL, Nambi V, Minor S, Sprecher DL. Statins do not meet expectations for lowering low-density lipoprotein cholesterol levels when used in clinical practice. *Am J Med.* 2002;113:625-9. Medline:12505111 doi:10.1016/S0002-9343(02)01303-7
- 18 Hwang JY, Jung CH, Lee WJ, Park CY, Kim SR, Yoon KH, et al. Low Density Lipoprotein Cholesterol Target Goal Attainment Rate and Physician Perceptions about Target Goal Achievement in Korean Patients with Diabetes. *Diabetes Metab J.* 2011;35:628-36. Medline:22247906 doi:10.4093/dmj.2011.35.6.628
- 19 Guan F, Xie J, Wang GL, Wang JH, Wang JS, Yu JM, et al. Community-wide survey of physicians' knowledge of cholesterol management. *Chin Med J (Engl).* 2010;123:884-9. Medline:20497682
- 20 Casula M, Tragni E, Catapano AL. Adherence to lipid-lowering treatment: the patient perspective. *Patient Prefer Adherence.* 2012;6:805-14. Medline:23152673
- 21 Plana N, Ibarretxe D, Cabre A, Ruiz E, Masana L. Prevalence of atherogenic dyslipidemia in primary care patients at moderate-very high risk of cardiovascular disease. Cardiovascular risk perception. *Clin Investig Arterioscler.* 2014;26:274-84. Medline:24931442 doi:10.1016/j.arteri.2014.04.002
- 22 McCrate F, Godwin M, Murphy L. Attainment of Canadian Diabetes Association recommended targets in patients with type 2 diabetes: a study of primary care practices in St John's, Nfld. *Can Fam Physician.* 2010;56:e13-9. Medline:20090056
- 23 Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA.* 1999;282:1458-65. Medline:10535437 doi:10.1001/jama.282.15.1458
- 24 Erhardt LR, Hobbs FD. A global survey of physicians' perceptions on cholesterol management: the From The Heart study. *Int J Clin Pract.* 2007;61:1078-85. Medline:17577295 doi:10.1111/j.1742-1241.2007.01420.x
- 25 Reiner Z, Sonicki Z, Tedeschi-Reiner E. Physicians' perception, knowledge and awareness of cardiovascular risk factors and adherence to prevention guidelines: the PERCRO-DOC survey. *Atherosclerosis.* 2010;213:598-603. Medline:20947087 doi:10.1016/j.atherosclerosis.2010.09.014



# The national and subnational prevalence and burden of age-related macular degeneration in China

Peige Song, Yuhang Du,  
Kit Yee Chan, Evropi  
Theodoratou, Igor Rudan;  
on behalf of the Global  
Health Epidemiology  
Research Group (GHERG)

Centre for Global Health Research,  
Usher Institute of Population  
Health Sciences and Informatics,  
University of Edinburgh, Edinburgh,  
Scotland, United Kingdom

## Correspondence to:

Peige Song  
Centre for Global Health Research  
Usher Institute of Population Health  
Sciences and Informatics  
University of Edinburgh  
Edinburgh EH8 9AG  
United Kingdom  
p.song@sms.ed.ac.uk

**Background** Age-related macular degeneration (AMD) is the third most common cause of blindness, and the fourth leading cause of visual impairment worldwide, but little is known about the burden of this disease in the most populous country—China. This study provides the first comprehensive estimates of the prevalence and burden of AMD in China from 1990 to 2015, with projections till 2050.

**Methods** In this study, a systematic review and meta-analysis was conducted to estimate the prevalence of AMD in China. China National Knowledge Infrastructure (CNKI), Wanfang, Chinese Biomedicine Literature Database (CBM-SinoMed), PubMed, Embase and Medline were searched before September 2016. Multilevel mixed-effect meta-regression was performed to define the prevalence rates of AMD and its subtypes. UN population data were used to estimate and project the number of people affected from 1990 to 2050. Based on different demographic and geographic features, the national burden of AMD in 2000 and 2010 was distributed to different regions in China.

**Results** Our search returned 2016 citations, of which 25 met the inclusion criteria. The prevalence of any AMD ranged from 2.44% (95% CI=1.85–3.22) in people aged 45–49 years to 18.98% (95% CI=15.05–23.66) in people aged 85–89 years. Prevalence of early AMD ranged from 1.79% (95% CI=1.05–3.02) to 10.05% (95% CI=6.17–15.97), and, in the case of late AMD, from 0.38% (95% CI=0.16–0.97) to 3.88% (95% CI=1.68–9.13). In late AMD, the prevalence of geographic atrophy (GA) was 0.15% (95% CI=0.05–0.47) in people aged 45–49 years and 1.09% (95% CI=0.35–3.36) in those aged 85–89 years, and the prevalence of neovascular AMD (NVAMD) ranged between 0.24% (95% CI=0.11–0.50) and 2.79% (95% CI=1.33–5.77). The number of people with any AMD was 12.01 million (95% CI=9.29–15.46) in 1990 and 26.65 million (95% CI=20.62–34.27) in 2015. Within the same period, the number of people with early AMD increased from 9.44 million (95% CI=7.74–11.15) to 20.91 million (95% CI=17.16–24.68), and those with late AMD rose from 2.58 million (95% CI=1.56–4.30) to 5.74 million (95% CI=3.46–9.59). In late AMD, the number of people living with GA ranged from 0.87 million (95% CI=0.40–1.83) in 1990 to 1.93 million (95% CI=0.89–4.08) in 2015, and NVAMD from 1.71 million (95% CI=1.16–2.47) to 3.81 million (95% CI=2.57–5.51). The projected number of people with any AMD in 2020 is 31.23 million (95% CI=24.18–40.14), increasing to 55.19 million (95% CI=43.04–70.30) in 2050. Between different regions, the South Central owed the most AMD cases (5.50 million in 2000 and 7.52 million in 2010), whereas the North–West China the least (0.66 million in 2000 and 0.95 million in 2010).

**Conclusions** The estimates in this study suggest a substantial burden of AMD in China, with the ageing process in Chinese society, this burden will be increasing in the foreseen future. Primary and secondary prevention and treatment and effective government response are urgently needed. Improved epidemiological studies are also required to better develop eye-care strategies and health services.

Age-related macular degeneration (AMD), a degenerative disease of the macula, is a leading cause of severe and irreversible loss of vision globally, and most notably in developed countries [1–3]. In 2010, it was estimated that AMD was the third most common cause of blindness, and the fourth leading cause of visual impairment worldwide [4]. Although AMD is not a life threatening disorder, up to one-third of the affected individuals will experience various degrees of disability and depression during the course of the disease, even when only one eye is affected [5,6]. Moreover, AMD is notably associated with falls and other injuries, resulting in increased economic and social burden for the individual, caregiver and community to bear [7–10]. Ageing is consistently documented as the most important risk factor for AMD [1,3,10]. In addition, other factors, such as cigarette smoking, female gender, ethnicity, and genetic predisposition may also play a role [3,11,12]. The combined effect of continuous exposure to different risk factors and different demographic ageing speed resulted in the global epidemic of AMD showing substantial variation across different ethnic groups and geographic regions [3,13–15].

The clinical course of AMD can be broadly divided into two stages: early and late (advanced) [1,16,17]. Early AMD is characterised by soft drusen and/or pigmentary changes, but many early cases do not progress to the advanced form [16,17]. Late AMD includes two types: geographic atrophy (GA) and neovascular (exudative) AMD (NVAMD). Compared with early AMD, late AMD is far less frequent but most damaging to the sight [18]. According to the latest global estimate of AMD prevalence, both early and late AMD were most frequent in populations of European ancestry (11.20% and 0.50%). Early AMD is least common in Asians (6.81%) while late AMD is least common in populations of African ancestry (0.28%) [3]. With Asia having the largest share of the world's population, and understanding that AMD is an age-driven disorder, it was estimated that Asia had the greatest number of people with AMD in 2014 (59 million). Furthermore, this number is expected to increase at the fastest pace in Asia in comparison to other regions – to 113 million by 2040. China, the most populous country in the world, is experiencing the most rapid ageing trend among all developing countries. It has been estimated that more than one-third of Chinese people living in China will be aged 60 years and over by 2050 [19]. It is, therefore, important to have an up-to-date summary of the magnitude and distribution of AMD in the general population to inform stakeholders and guide eye-related health policy-making and health services allocation in China.

In the last two decades, an increasing number of epidemiological studies of AMD have been conducted in China. The estimates were, however, contingent upon the characteristics of individual studies: the age structure of the study sample, case definition and classification of AMD [20–22]. Another important feature of AMD is that its prevalence is likely to be associated with geographic factors. In the most recent global geo-epidemiology analysis of AMD, both latitude and longitude were inversely correlated with AMD prevalence, providing a new clue to study the geographic distribution of AMD [15].

Until recently, there were no systematic estimates of AMD prevalence in China. With that said, the sheer volume of data available on the prevalence of AMD in Chinese bibliographical databases makes it possible to summarise the prevalence and burden of AMD from a modelling perspective [23,24]. Moreover, the large territory area with great variation of latitude and longitude in China provides a good opportunity to explore the influence of geographic factors within the same country. In this study, we undertook a comprehensive systematic review, in both Chinese and English databases, to retrieve population-based studies of AMD prevalence in China from 1990 onwards. Based on the existing evidence, we estimated and projected the prevalence and burden of AMD and its sub-types. The aims of this study were to 1) ascertain the AMD prevalence in China by using epidemiological modelling; 2) estimate and project the overall prevalence and number of people living with AMD at the national level from 1990 to 2050; 3) estimate the regional prevalence and number of people with AMD from 2000 to 2010.

## METHODS

### Systematic review

For developing epidemiological models to estimate the prevalence of AMD and its subtypes in the general population, a systematic review was conducted by two independent reviewers (PS and YD) in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines and the Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER) statement [25,26]. To ensure that all possible informative studies are included, a comprehensive literature search (title, abstract and keywords) was conducted in order to identify relevant studies. First, three Chinese bibliographic databases and three English bibliographic databases were searched from inception to 17 September

2016. These were the China National Knowledge Infrastructure (CNKI), Wanfang, Chinese Biomedicine Literature Database (CBM-SinoMed), PubMed, Embase and Medline. The source of studies in the three Chinese databases included journal articles, abstracts, dissertations and conference proceedings, whereas those in the three English databases included journal articles only. A combination of search terms for prevalence (prevalence, incidence, mortality, morbidity, epidemiology), AMD (age-related macular degeneration, age-related maculopathy, retina\* macula\* age related degeneration, retina\* macular degeneration, macular degeneration) and China (China, Chinese, Hong Kong, Macau, Taiwan) was adopted for the comprehensive search. The final search strategy is presented in Table S1 in **Online Supplementary Document**. Note that the search strategy for the different bibliographic databases was slightly different based on the database's specific search features. Snowball searching of reference lists of publications retrieved in the first step was then conducted to further identify studies of interest. Only studies published since 1990 were retrieved and no language restrictions were imposed.

Only population-based studies that quantified the prevalence of AMD were included in this study. This is because studies conducted at institutional sites tend to have poor representativeness of the surrounding general population, especially for affected people living in poor and rural areas where access to health is not universal. Studies that relied on self-reported diagnosis were also excluded, due to recall bias. Studies that only reported the number of eyes affected by AMD, rather than the number of affected individuals, were also excluded because no prevalence of AMD could be derived from such studies. Duplicate publications of the same study were compared and the study providing more details was retained. Some additional criteria were also applied to ensure the quality of included studies. The detailed selection criteria are shown in **Table 1**.

Before reviewing the retrieved records, duplicates were removed manually. Records were screened for relevance in two stages: screening of titles and abstracts followed by the retrieval and check of full-text articles. All non-English or non-Chinese language documents were reviewed after translation into English by Google Translate. For studies that fulfilled the criteria, three main categories of data were extracted: characteristics of the study, characteristics of the investigated population, and prevalence estimates of AMD and its subtypes. The data extraction tables were pilot tested on ten randomly selected included studies and refined accordingly before the final extraction.

The final data extraction table included:

- 1) Characteristics of the study: authors, publication year, study setting, year of survey, sampling method, study design (cross-sectional or cohort), AMD assessment method, and AMD grading system;
- 2) Characteristics of the investigated population: number of the sample, population type (urban, rural or mixed), gender (male, female or mixed), and age (age range, mean or median age, or midpoint of the age range);
- 3) Prevalence data: number of people with AMD and the number of participants who had been tested, by age group, gender, setting and AMD subtype where available.

The geographic indicators of interest (latitude, longitude and average annual insolation) were assigned to each study accordingly. The latitude and longitude data were obtained using Google Maps GPS coordinates (<http://www.gps-coordinates.net/>). The average annual insolation data (ie, the amount of solar radiation incident on the surface of the earth) on the horizontal surface, expressed in kWh/m<sup>2</sup>/d, was obtained from the National Aeronautics and Space Administration (NASA) Atmospheric Science Data Centre (<http://eosweb.larc.nasa.gov/sse/>). When study settings were defined as larger regions, such as at province, or regional levels, the mean centre point of the setting was calculated and the corresponding

**Table 1.** Selection criteria of studies in the systematic review

Inclusion criteria
1) Community-based study of AMD in China (including Hong Kong, Macao and Taiwan)
2) Studies conducted to examine the epidemiology of AMD
3) Studies reported numerical prevalence measure of AMD
Exclusion criteria
1) Multiple publications of the same study
2) Studies with no professional assessment methods or relied on self-reported diagnoses
3) Studies that were conducted in a population with characteristics that were clearly unrepresentative, eg, visual impaired population, diabetes population
4) Studies with inconsistencies between reported methods and presented results

geographic data of the centre point was used. Studies that reported raw prevalence data in more than one geographic area (eg, a single study presented prevalence of AMD for three different cities) were recorded separately for each geographic area. For studies that reported aggregated AMD prevalence data for different geographic areas, the average geographic data of the different areas were calculated and recorded. For studies with missing data of survey year, three years were subtracted from the published year to impute the survey year, which was based on the average time from survey to publication in studies with available data. In studies where censoring age groups were reported, eg, older than 80 years, the missing age band was taken as the same width as other age groups in the same study.

The classification systems used to define AMD and its subtypes include the Wisconsin age-related maculopathy system (WARMGS) [27], the International Classification and Grading system (IC) [28], the Clinical Age-Related Maculopathy Grading System (CARMS) [29], and the “Age-related Macular Degeneration Clinical Diagnosis Standard” proposed by the China Medical Association in 1986 (CMA1986) [30]. For studies adopting different classification systems, the prevalence of any AMD, early AMD, late AMD, which included GA and NVAMD, was extracted or calculated (if necessary) separately according to the definitions below:

- 1) Early AMD: any soft drusen (distinct or indistinct) and pigmentary abnormalities, or large soft drusen 125  $\mu\text{m}$  or more in diameter with a large drusen area (>500  $\mu\text{m}$  diameter circle) or large soft indistinct drusen in the absence of signs of late-stage disease;
- 2) Late AMD: the presence of geographic atrophy or pigment epithelial detachment, subretinal haemorrhage or visible subretinal new vessel, or subretinal fibrous scar or laser treatment scar.

### Statistical analysis

Due to high heterogeneity between studies that reported prevalence rates for any AMD, early AMD, late AMD, GA and NVAMD (Table S2 in **Online Supplementary Document**), random-effect models were adopted throughout the analysis. In the data extraction process, data were stratified by age, gender and setting. Some studies provided more than one data point. To take this hierarchical data structure into account, a multilevel mixed-effect meta-regression was conducted [31,32]. Given that:

$$\text{Prevalence} = p = \frac{\text{Number of cases}}{\text{Sample size}}$$

Then, the binomial distribution of prevalence rates was transferred to the normal distribution by using logit link:

$$\text{logit}(p) = \log_e \left( \frac{p}{1-p} \right) = \log_e(\text{odds}) = \alpha + \beta_1 \times x_1 + \beta_2 \times x_2 + \dots \beta_n \times x_n$$

Estimates were back transformed and expressed as conventional prevalence:

$$p = \frac{e^{\alpha + \beta_1 \times x_1 + \beta_2 \times x_2 + \dots \beta_n \times x_n}}{1 + e^{\alpha + \beta_1 \times x_1 + \beta_2 \times x_2 + \dots \beta_n \times x_n}}$$

To develop the overall “envelope” of AMD cases in China from 1990 to 2015, five models were first developed to establish the prevalence of any AMD, early AMD, late AMD, GA and NVAMD as a function of age:

$$\text{logit}(p) = \alpha + \beta_1 \times (\text{age})$$

Thus, the prevalence of AMD is:

$$p = \frac{e^{\alpha + \beta_1 \times (\text{age})}}{1 + e^{\alpha + \beta_1 \times (\text{age})}}$$

The total number of AMD cases (“envelope”) in China was calculated by multiplying the age-specific prevalence of AMD for each 5-year age group estimated in the above models with the corresponding 5-year population subgroups in China, available from the United Nations Population Division (UNPD) [19]. This was performed for any AMD, early AMD, late AMD, GA and NVAMD separately in the years 1990, 2000, 2010 and 2015.

To investigate whether study-level demographic and geographic factors might affect the prevalence of AMD, variables of interest were added into the multilevel mixed-effect meta-regression to test the significance [33]. As a rule, at least seven data points should be available for each variable [34]. These variables included gender, setting, latitude, longitude and average annual insolation. Investigation year was

also tested so as to assess if there were any significant time trends. All variables that individually associated AMD prevalence in univariable analyses were included in the subsequent multivariable regression model, where variables that were not statistically significant were removed, starting from the one with the highest p value.

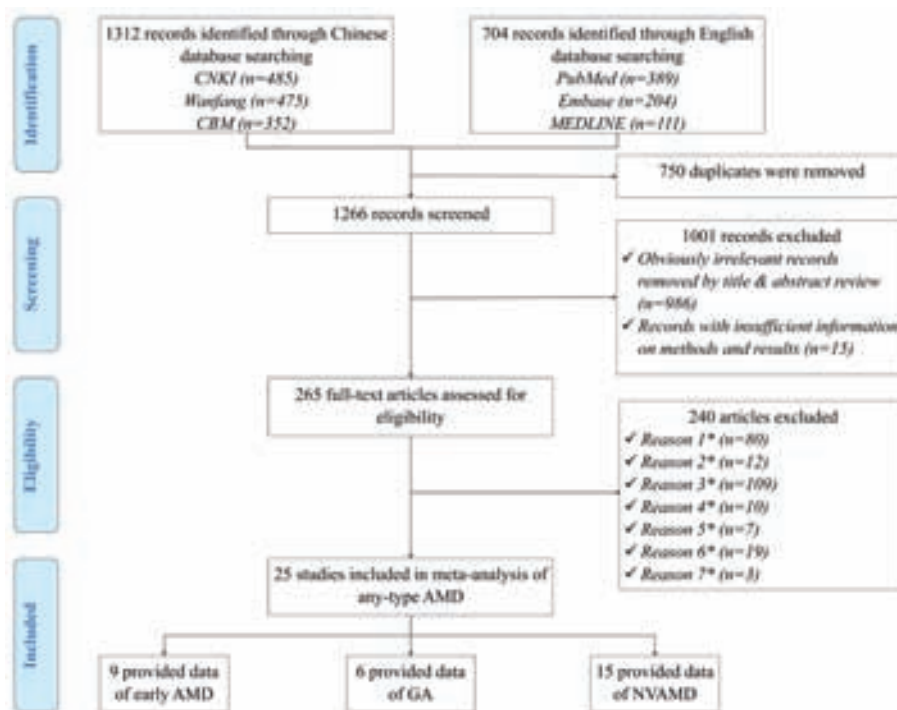
For our projection to the year 2050, age-specific prevalence rates of AMD were assumed to be constant over the next 33 years, the number of individuals with AMD from 2020 to 2050 was calculated by multiplying the age-specific prevalence rates to the UNDP Prospects data [19].

Based on the final multivariable regression models that take the effects of demographic and geographic factors into consideration, the estimated national population with AMD was distributed into six geographical regions, namely, East China, North China, Northeast China, Northwest China, South Central China, Southwest China (Table 2) [35–37]. This method was initially proposed by the Child Health Epidemiology Reference Group (CHERG), and has, since, been adopted widely in disease burden research [38–40]. First, AMD prevalence in each geographic region was calculated, based on the final regression equation. Second, the regional population with AMD was estimated by multiplying the regional AMD prevalence and corresponding population for the years 2000 and 2010, where regional population data were available from the fifth and sixth census [36,37]. Finally, the regional population of AMD was adjusted to fit the national AMD “envelope”.

Non-dichotomous variables were analysed as continuous. A two-sided p value less than 0.05 was regarded as statistically significant for all analyses. All statistical analyses were performed in R Studio (ver-

**Table 2.** The six geographical regions in China

REGION	INCLUDED PROVINCES
North China	Beijing Municipality, Hebei province, Inner Mongolia Autonomous Region, Shanxi province, Tianjin Municipality
Northeast China	Heilongjiang province, Jilin province, Liaoning province;
East China	Anhui province, Fujian province, Jiangsu province, Jiangxi province, Shandong province, Shanghai Municipality, Zhejiang province
South Central China	Guangdong province, Guangxi Zhuang Autonomous Region, Hainan province, Henan province, Hubei province, Hunan province
Southwest China	Chongqing Municipality, Guizhou province, Sichuan province, Tibet Autonomous Region, Yunnan province
Northwest China	Gansu province, Ningxia Hui Autonomous Region, Qinghai province, Shaanxi province, Xinjiang Uyghur Autonomous Region



**Figure 1.** Systematic review flow diagram. Note: \*Reason 1 – Studies that were not population-based; \*Reason 2 – Studies that were not based in China; \*Reason 3 – Papers with no numerical prevalence measure of AMD; \*Reason 4 – Studies that had no professional assessment methods or relied on self-reported diagnoses; \*Reason 5 – Studies that were conducted in a population with unrepresentative characteristics; \*Reason 6 – Multiple publications of the same study; \*Reason 7 – Papers with inconsistency between reported methods and presented results.

**Table 3.** Main characteristics of the included prevalence studies (n=25)

CHARACTERISTICS OF STUDY	NUMBER OF STUDIES (%)
<b>Year published:</b>	
1990–1999	7 (28.0)
2000–2009	7 (28.0)
2010–2016	11 (44.0)
<b>Setting:</b>	
Urban	7 (28.0)
Rural	9 (36.0)
Mixed	9 (36.0)
<b>Sample size:</b>	
600–1000	5 (20.0)
1001–2000	5 (20.0)
2001–3000	7 (28.0)
3001–5000	4 (16.0)
5001–8000	4 (16.0)
<b>Grading system:</b>	
CMA1986*	12 (48.0)
WARMGS†	6 (24.0)
CARMS‡	5 (20.0)
IC§	1 (4.0)
Other	1 (4.0)

\*CAM 1986 – the “Age-related Macular Degeneration Clinical Diagnosis Standard” proposed by the China Medical Association in 1986.

†WARMGS – the Wisconsin age-related maculopathy system.

‡CARMS, the Clinical Age-Related Maculopathy Grading System.

§IC, the International Classification and Grading system.

||Other, definition in the “Ophthalmology” (7th version).

sion 1.0.136) built on R (version 3.3.0). All included studies in the analysis were mapped by ArcGIS software (Version 10.1). The China base map was obtained as a shapefile from the Global Administrative Areas (GADM) database (GADM, 2015, version 2.0; www.gadm.org).

## RESULTS

### Summary of systematic review

**Figure 1** Shows the process of systematic review for studies included in the final meta-analysis. In brief, the initial search identified 2016 citations. After removing 750 duplications, 986 apparently irrelevant citations by title and abstract review, and 15 citations with no sufficient information on methods and results, 265 papers were reviewed at the full-text level to assess their eligibility. Ultimately, 25 AMD prevalence studies were included in the final analysis.

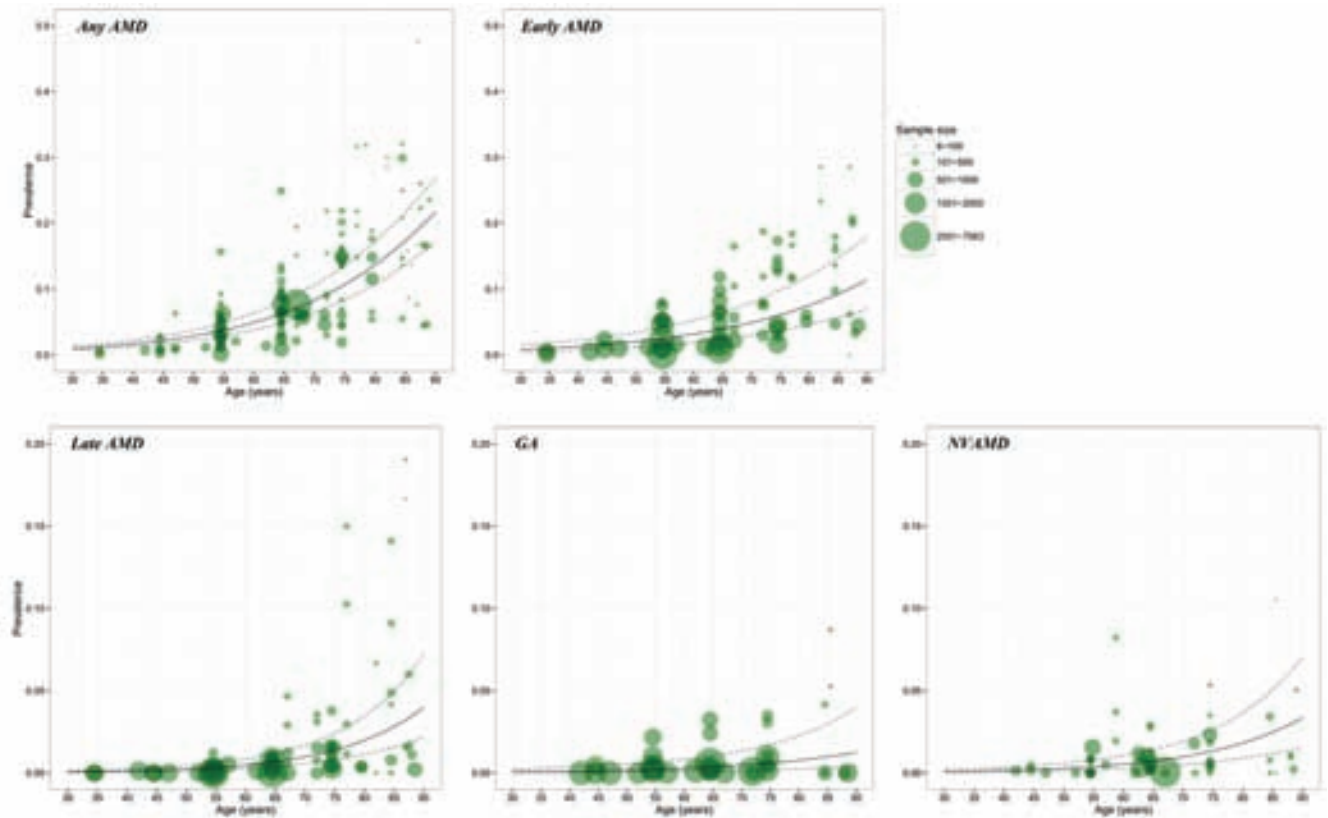
A full list of included studies is shown in Table S3 in **Online Supplementary Document**, the included data involved 3016 AMD cases in a total of 43 420 examined individuals. **Table 3** shows the main characteristics of the studies, and the detailed characteristics of every study can be found in Table S4 in **Online Supplementary Document**. All included studies were cross-sectional studies that assessed AMD by using fundus imaging. Almost half of the retained studies were published in the past six years (44.0%), with CMA1986 the most widely adopted grading system (48.0%), followed by WARMGS (24.0%) and CARMS (20.0%). The geographic distribution of the 25 included studies is demonstrated in **Figure 2**.

### Age-specific prevalence of AMD

In each model (**Figure 3**), a substantial number of data points were available for constructing the relationship between AMD prevalence and age. The age spectrum ranged from around 35 years to less than 90 years. However, for GA and NVAMD, few data points were available at younger ages (30–40 years). In this study, to ensure that the estimated prevalence was comparable, the lower bound of age range was set as 45 years and the upper bound as 89 years where data were available for model construction at all AMD subtype groups.



**Figure 2.** Geographic distribution of the included studies in China.

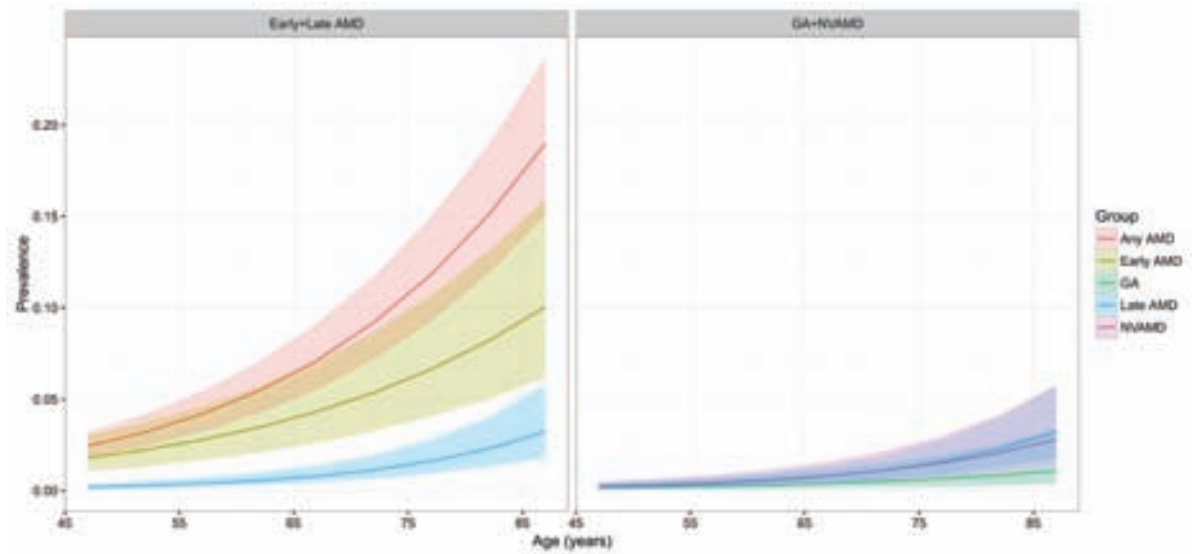


**Figure 3.** Prevalence of age-related macular degeneration (AMD) and its subtypes by age in retained studies. Note: The size of each bubble is proportional to the sample size. There were 124 data points for constructing the relation between prevalence and age for any AMD, 67 for early AMD, 67 for late AMD, 35 for geographic atrophy (GA) and 54 for neovascular AMD (NVAMD).

The estimated age-specific prevalence of any AMD, early AMD, late AMD, GA and NVAMD is shown in **Figure 4** and **Table 4**. The prevalence of any AMD ranged from 2.44% (95% CI=1.85–3.22) in people aged 45–49 years to 18.98% (95% CI=15.05–23.66) in people aged 85–89 years. Prevalence of early AMD ranged from 1.79% (95% CI=1.05–3.02) to 10.05% (95% CI=6.17–15.97), and, in the case of late AMD, from 0.38% (95% CI=0.16–0.97) to 3.88% (95% CI=1.68–9.13). In late AMD, the prevalence of GA was 0.15% (95% CI=0.05–0.47) in people aged 45–49 years and 1.09% (95% CI=0.35–3.36) in those aged 85–89 years, and the prevalence of NVAMD ranged between 0.24% (95% CI=0.11–0.50) and 2.79% (95% CI=1.33–5.77).

### National population affected with AMD from 1990 to 2015

By applying the age-specific prevalence of AMD to the national population in 1990, 2000, 2010 and 2015, the number of people living with AMD in China was estimated (**Table S5** in **Online Supplementary Document**). During this period, the national prevalence of any AMD slightly decreased by 0.41%, from 5.26% (95% CI=4.07–6.76) in 1990 to 5.24% (95% CI=4.05–6.73) in 2015. This declining trend was also witnessed in early AMD and late AMD, with decreasing rates of 0.50% and 0.07% respectively. In late AMD, GA also showed a decreasing trend within this time frame, whereas the prevalence of NVAMD increased slightly (**Table 5**). Despite this decreasing prevalence trend during 1990–2015, the overall number of people with any AMD or its subtypes all increased dramatically due to the rapidly ageing population. The national number of people with any AMD increased by 121.80%, from 12.01 million (95% CI=9.29–15.46) in 1990 to 26.65 million (95% CI=20.62–34.27) in 2015. Within the same period, the number of people with early AMD increased from 9.44 million (95% CI=7.74–11.15) to 20.91 million (95% CI=17.16–24.68), and those with late AMD rose from 2.58 million (95% CI=1.56–4.30) to 5.74 million (95% CI=3.46–9.59), which yielded increasing rates of 121.60% and 122.55% respectively. In late AMD, increase in the number of people living with GA was similar to those with NVAMD (121.99%



**Figure 4.** Estimated age-specific prevalence of age-related macular degeneration (AMD) and its subtypes in China, with 95% confidence intervals. GA – geographic atrophy, NVAMD – neovascular AMD.

**Table 4.** Estimated age-specific prevalence (% and 95% confidence interval) of age-related macular degeneration (AMD) and its subtypes in China

AGE (YEARS)	ANY AMD	EARLY AMD	LATE AMD	GA	NVAMD
45–49	2.44 (1.85–3.22)	1.79 (1.05–3.02)	0.38 (0.16–0.97)	0.15 (0.05–0.47)	0.24 (0.11–0.50)
50–54	3.21 (2.45–4.19)	2.23 (1.32–3.74)	0.51 (0.22–1.24)	0.19 (0.06–0.58)	0.32 (0.16–0.67)
55–59	4.20 (3.22–5.45)	2.78 (1.67–4.62)	0.68 (0.30–1.60)	0.24 (0.08–0.72)	0.44 (0.22–0.89)
60–64	5.47 (4.23–7.06)	3.47 (2.09–5.71)	0.91 (0.41–2.09)	0.31 (0.11–0.90)	0.60 (0.30–1.19)
65–69	7.11 (5.52–9.12)	4.31 (2.61–7.05)	1.22 (0.55–2.76)	0.40 (0.14–1.15)	0.82 (0.41–1.60)
70–74	9.20 (7.17–11.72)	5.36 (3.26–8.68)	1.63 (0.74–3.67)	0.52 (0.18–1.48)	1.11 (0.56–2.19)
75–79	11.81 (9.26–14.96)	6.63 (4.05–10.68)	2.18 (0.98–4.94)	0.66 (0.23–1.93)	1.52 (0.76–3.01)
80–84	15.05 (11.85–18.92)	8.18 (5.01–13.09)	2.91 (1.29–6.70)	0.85 (0.28–2.54)	2.06 (1.01–4.16)
85–89	18.98 (15.05–23.66)	10.05 (6.17–15.97)	3.88 (1.68–9.13)	1.09 (0.35–3.36)	2.79 (1.33–5.77)

GA – geographic atrophy, NVAMD – neovascular AMD

**Table 5.** Estimated prevalence and number of people living with age-related macular degeneration (AMD) in China from 1990 to 2015, by AMD type

AMD TYPE	PREVALENCE OF AMD (% , 95% CI)		NUMBER OF PEOPLE WITH AMD (MILLION, 95% CI)		RATE OF CHANGE (% , 1990–2015)	
	1990	2015	1990	2015	prevalence	AMD cases
Any AMD	5.26 (4.07–6.76)	5.24 (4.05–6.73)	12.01 (9.29–15.46)	26.65 (20.62–34.27)	–0.41	121.80
Early AMD	4.13 (3.39–4.88)	4.11 (3.37–4.85)	9.44 (7.74–11.15)	20.91 (17.16–24.68)	–0.50	121.60
Late AMD	1.13 (0.68–1.88)	1.13 (0.68–1.88)	2.58 (1.56–4.30)	5.74 (3.46–9.59)	–0.07	122.55
GA	0.38 (0.17–0.80)	0.38 (0.17–0.80)	0.87 (0.40–1.83)	1.93 (0.89–4.08)	–0.33	121.99
NVAMD	0.75 (0.51–1.08)	0.75 (0.51–1.08)	1.71 (1.16–2.47)	3.81 (2.57–5.51)	0.05	122.84

GA – geographic atrophy, NVAMD – neovascular AMD

vs 122.84%), which ranged from 0.87 million (95% CI=0.40–1.83) to 1.93 million (95% CI=0.89–4.08), and 1.71 million (95% CI=1.16–2.47) to 3.81 million (95% CI=2.57–5.51) throughout this time frame respectively (Table 5). In 2015, the age group that contributed the most cases of any AMD, early AMD, late AMD, GA and NVAMD was 60–64 years (Figure 5).

### Effects of demographic and geographic factors on the prevalence of AMD

Findings from the univariable meta-regression analyses (Table S6 in Online Supplementary Document) showed that age, setting and latitude were significantly associated with the prevalence of any AMD. For early AMD, age, gender, setting and latitude also had a significant influence on the prevalence. For late AMD, age and latitude were found to be significantly associated with the prevalence. However, in late AMD, only age was found to be significantly associated with the prevalence of GA, and age, gender, latitude and insolation were significantly associated with the prevalence of NVAMD. For any AMD and all subtypes, the investigation year was identified to have no influence on prevalence rates and increased age was the only constantly significant risk factor.

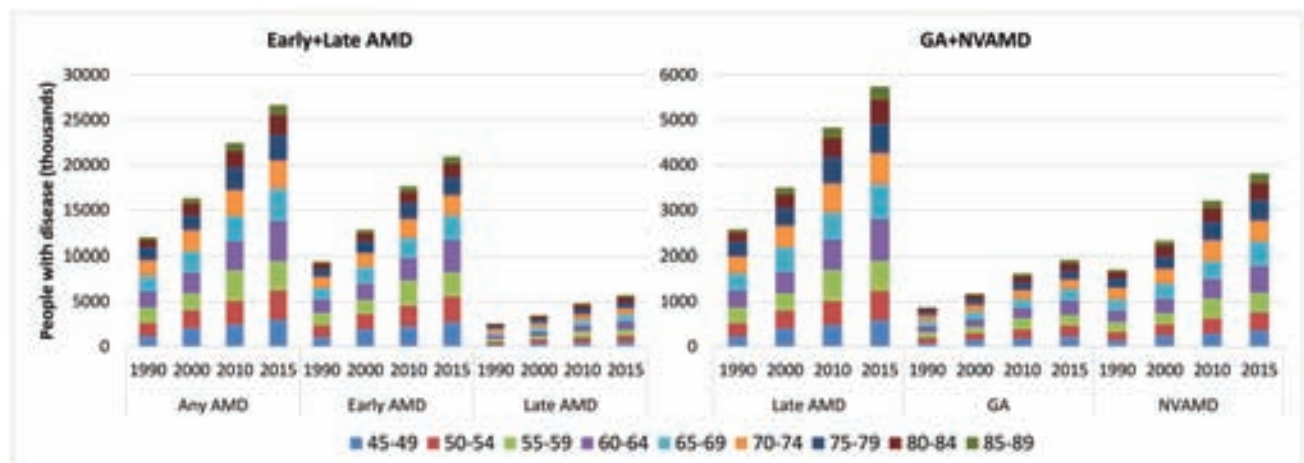
Although most studies provided multiple data points of prevalence rates, these data were mainly stratified by age groups. For AMD subtype groups (early AMD, late AMD, GA and NVAMD), and after controlling the difference of age structures, insufficient data were available for conducting multivariable meta-regression that simultaneously included all statistically significant factors identified in the univariable analyses. Thus, here the multivariable regression model was only conducted and reported for any AMD. The formula generated from the multivariable regression is shown below:

$$\text{logit}(p) = -4.230 + 0.056 \times \text{age} + (-0.6013) \times \text{setting}_{\text{rural}} + 0.053 \times \text{setting}_{\text{urban}} + (-0.060) \times \text{latitude}$$

Where  $p$  indicates the prevalence of any AMD;  $\text{setting}_{\text{rural}} = 1$  for rural setting and =0 otherwise;  $\text{setting}_{\text{urban}} = 1$  for urban setting and =0 otherwise;  $\text{latitude}$  refers to the absolute value of latitude.

### Projection of national population affected with AMD from 2020 to 2050

No secular trend of the prevalence of any AMD, early AMD, late AMD, GA and NVAMD was observed in the included studies, thus age-specific prevalence was assumed as constant for the projection analysis. By applying the age-specific prevalence of AMD to the national population in 2020, 2030, 2040 and 2050, the number of people with AMD was estimated (Tables S5 in Online Supplementary Document). Unlike the slightly fluctuating trend of AMD prevalence during 1990 to 2015, the prevalence rates of all subtypes of AMD will increase notably during 2020 and 2050. In 2020, the prevalence of any AMD will be 5.39% (95% CI=4.18–6.93) and is expected to increase by 41.66%, reaching to 7.64% (95% CI=5.96–9.73) in 2050. Among all subtypes of AMD, NVAMD will show the greatest increasing rate of 57.48%, from 0.78% (95% CI=0.52–1.12) in 2020 to 1.22% (95% CI=0.83–1.75) in 2050, whereas the increasing rate of early AMD will be the smallest (38.45%), from 4.23% (95% CI=3.47–4.99) to 5.21% (95% CI=4.31–6.08) during this period (Table 6).



**Figure 5.** Estimate of the national number of people with age-related macular degeneration (AMD) and contributing age groups in China from 1990 to 2015, by AMD type. GA – geographic atrophy, NVAMD – neovascular AMD.

From 2020 to 2050, the number of cases of any AMD in China will rise by 76.72%, from 31.23 million (95% CI=24.18–40.14) to 55.19 million (95% CI=43.04–70.30). The increasing rate of late AMD cases will be greater than early AMD cases (91.12% vs 72.70%), with the number of people affected by early AMD increasing from 24.47 million (95% CI=20.10–28.87) in 2020 to 42.26 million (95% CI=35.15–49.05) in 2050, and those affected by late AMD from 6.76 million (95% CI=4.08–11.28) to 12.92 million (95% CI=7.89–21.26). In late AMD, the number of people with GA will increase by 80.78%, from 2.26 million (95% CI=1.04–4.78) in 2020, to 4.09 million (95% CI=1.89–8.59) in 2050. Furthermore, the number of those with NVAMD will grow even further (96.45%), from 4.50 million (95% CI=3.04–6.50) to 8.84 million (95% CI=6.00–12.66) (Table 6). From 2020 to 2050, the age groups to contribute the most cases will shift from 65–69 years to 80–84 years for any AMD, late AMD, GA and NVAMD, and from 65–69 years to 75–79 years for early AMD (Figure 6).

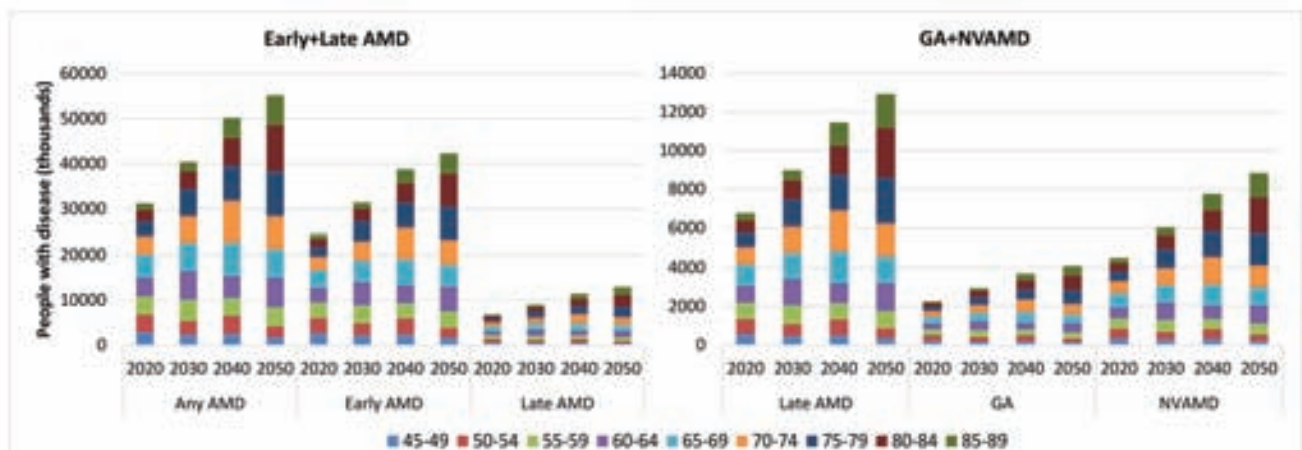
### Regional population affected with AMD from 2000 to 2010

The total number of AMD cases in China in 2000 and 2010 was distributed across the six geographical regions according to the final multivariable model that took into account three main factors: age, setting and latitude. In 2000, the national prevalence of AMD in China was 5.16% (95% CI=3.99–6.64), with the regional prevalence estimates ranging from 2.69% (95% CI=1.67–4.29) in North–East China to 6.64% (95% CI=5.12–8.52) in South Central China. In 2010, the prevalence was still the highest in South Central China (6.74% [95% CI=5.20–8.65]) and the lowest in North–East China (2.65% [95% CI=1.66–4.20]), with the overall prevalence in Chinese population increasing to the level of 5.24% (95% CI=4.05–6.73). During 2000 to 2010, the overall prevalence of AMD increased by 1.44%, and the most marked

**Table 6.** Projected prevalence and number of people living with age-related macular degeneration (AMD) in China from 2020 to 2050, by AMD type

AMD TYPE	PREVALENCE OF AMD (% , 95% CI)		NUMBER OF PEOPLE WITH AMD (MILLION, 95% CI)		RATE OF CHANGE (% , 2020–2050)	
	2020	2050	2020	2050	prevalence	AMD cases
Any AMD	5.39 (4.18–6.93)	7.64 (5.96–9.73)	31.23 (24.18–40.14)	55.19 (43.04–70.30)	41.66	76.72
Early AMD	4.23 (3.47–4.99)	5.85 (4.87–6.79)	24.47 (20.10–28.87)	42.26 (35.15–49.05)	38.45	72.71
Late AMD	1.17 (0.70–1.95)	1.79 (1.09–2.94)	6.76 (4.08–11.28)	12.92 (7.89–21.26)	53.28	91.21
GA	0.39 (0.18–0.83)	0.57 (0.26–1.19)	2.26 (1.04–4.78)	4.09 (1.89–8.59)	44.92	80.78
NVAMD	0.78 (0.52–1.12)	1.22 (0.83–1.75)	4.50 (3.04–6.50)	8.84 (6.00–12.66)	57.48	96.45

GA – geographic atrophy, NVAMD – neovascular AMD



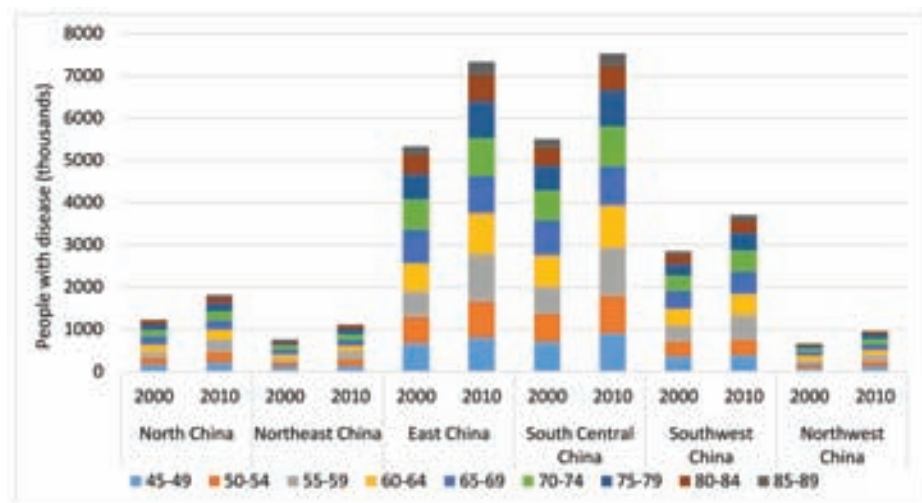
**Figure 6.** Projection of the national number of people with age-related macular degeneration (AMD) and contributing age groups in China from 2020 to 2050, by AMD type. GA – geographic atrophy, NVAMD – neovascular AMD.

increase was in Southwest China (6.51%) while the prevalence rate of AMD declined by 1.37% in North-East China (Table 7).

Estimates of the number of people living with AMD in different regions are shown in Table 7 and Figure 7. With the ageing trend of the Chinese population, the total number of people living with AMD in China increased by 37.50%, from 16.31million (95% CI=12.62–20.99) in 2000 to 22.43 million (95% CI=17.36–28.85) in 2010. In 2000, more than one-third (33.72%) of Chinese AMD cases were found living in South Central China (5.50 million, 95% CI=4.24–7.05) and only 4.05% were in North-West China (0.66 million, 95% CI=0.49–0.89). In 2010, this distribution of AMD cases remained the same across the six geographical regions, with most (33.53%) of the AMD cases in South Central China (7.52 million, 95% CI=4.24–7.05) and the least (4.24%) in North-West China (0.95 million, 95% CI=0.71–1.28). From 2000 to 2010, the most striking increases in the number of AMD cases were in North China (48.64%) and North-East China (47.06%), and the least in South-West China (29.97%). Throughout this decade, the age groups that contributed the most AMD cases shifted from 55–59 years to 60–64 years in all of the six regions.

**Table 7.** Estimated prevalence and number of people living with any age-related macular degeneration (AMD) in China in the years 2000 and 2010, by geographical region

REGION	PREVALENCE OF AMD (% , 95% CI)		NUMBER OF PEOPLE WITH AMD (MILLION, 95% CI)		RATE OF CHANGE (% , 2000–2010)	
	2000	2010	2000	2010	prevalence	AMD cases
North China	3.28 (2.35–4.55)	3.36 (2.42–4.64)	1.23 (0.88–1.70)	1.82 (1.31–2.51)	2.40	48.18
North-East China	2.69 (1.67–4.29)	2.65 (1.66–4.20)	0.76 (0.47–1.21)	1.11 (0.69–1.76)	–1.37	46.70
East China	5.46 (4.39–6.74)	5.57 (4.49–6.86)	5.32 (4.29–6.57)	7.33 (5.91–9.04)	1.99	37.67
South Central China	6.64 (5.12–8.52)	6.74 (5.20–8.65)	5.50 (4.24–7.05)	7.52 (5.80–9.64)	1.54	36.74
South-West China	5.68 (4.50–7.11)	6.05 (4.81–7.56)	2.85 (2.26–3.56)	3.70 (2.94–4.62)	6.51	29.97
North-West China	3.29 (2.44–4.43)	3.40 (2.53–4.55)	0.66 (0.49–0.89)	0.95 (0.71–1.28)	3.23	44.65
China	5.16 (3.99–6.64)	5.24 (4.05–6.73)	16.31 (12.62–20.99)	22.43 (17.36–28.85)	1.44	37.50



**Figure 7.** Estimate of the regional number of people with age-related macular degeneration (AMD) and contributing age groups in China in the years 2000 and 2010.

## DISCUSSION

In this systematic review and meta-analysis, data-driven estimates and projections of AMD prevalence and burden in China were presented, both at the national level and at the regional levels. The results from this synthesised population-based data show that the burden of AMD in China is substantial. From 1990 to 2015, the prevalence of AMD fluctuated at around 5.2%, which translates to a total of 26.65 million affected individuals in 2015. By 2050, prevalence of AMD is expected to increase to 7.64%, with the corresponding number of affected individuals being 55.19 million. Substantial regional variation was found across the country, with AMD prevalence being the highest in South Central China and the lowest in the North-East. In terms of the total number of AMD cases, the greatest burden was in the South Central area, and the smallest in North-West China.

To the best of our knowledge, this study is the first attempt to estimate the prevalence and the burden of AMD in China and to make future projections. The comprehensive search strategies and strict inclusion and exclusion criteria ensured a well-designed analysis. Furthermore, the current study provided estimates of the prevalence and the number of affected AMD cases by AMD subtype, with this additional information being of particular clinical and public health relevance. Indeed, such information offers valuable, detailed insights into the burden of AMD in China. Critically, this study used the best available data to portray a complete picture of the public health burden of AMD in different regions. Therefore, it can serve as the basis for health policy making and resource allocation for AMD prevention and treatment initiatives. From a global perspective, this study complements the most recent Global AMD study, where insufficient data were available for national estimates and projections [3].

However, this study is not free from limitations. First, significant heterogeneity existed between all of the included studies, despite the strict inclusion and exclusion criteria applied. Like any meta-analysis, the findings of this study are only as good as the included primary investigations. Included studies did not come from across the country, thus the ability to generate provincial estimates of AMD prevalence and cases may be limited. Second, one NVAMD-like disease, polypoidal choroidal vasculopathy, is markedly more common in Asians [3,41,42], and taking into consideration that most population-based studies may have limited ability to distinguish between these two diseases (as suggested previously [3]), the prevalence and burden of NVAMD in the current study may be overestimated. In addition, as suggested by a previous meta-analysis, studies using fundus imaging with classifications, rather than the internationally recognised grading systems, are more likely to diagnose late AMD [13]. In this study, almost half of the included studies adopted the grading system proposed by the Chinese Medical Association, which may also have contributed to the peculiarly elevated prevalence rate of late AMD. A further point to raise is that only a limited set of variables were included and explored in the meta-regression analysis. This means that there could have also been further explanatory variables that may influence the presence of AMD. Moreover, the included demographic and geographic variables were mainly aggregate level data, and although efforts were made to extract data stratified by age, gender and location, the variation at the individual level may still be hidden. This may include smoking exposure, the habit of wearing sunglasses, and others. A further limitation of the study is that the estimates of regional prevalence and burden of AMD were based on the assumption that the pooled prevalence estimate for a specific region was homogeneous across all included provinces within this region, but this is quite unrealistic. Additionally, for regions that contributed only a few, or no actual AMD prevalence data points to the model, the model-based estimates may diverge quite considerably from the true prevalence. Finally, as reported in both previous reviews and substantiated in the current study, the prevalence of AMD and all its subtypes was stable over time [3,13]. Based on this assumption, the projections of the national prevalence and burden of AMD were actually based on the model-based age-specific prevalence and demographic changes during the next three decades. Thereby, the uncertainty of these projections may be largely dependent on the accuracy of age-specific prevalence model and the UNPD population projection. Bearing these limitations in mind, estimates presented in the current study should be interpreted judiciously.

In this study, the overall prevalence of AMD among the Chinese population was lower than the estimates in the Global AMD study, which reported an overall AMD prevalence of 6.86% in people living in Asia [3]. In the Global AMD study, the prevalence estimates for Asia were based on eleven studies conducted across Asia, among which, six came from south Asia (India, Singapore, and Thailand), three from China and two from Japan. Given the fact that AMD prevalence increases with decreasing latitude, as detected in both this study and a previous global geo-epidemiology study of AMD [15], it is not surprising that the overall prevalence of AMD in the current study is lower than that in Asia – as estimated by the Glob-

al AMD study. Moreover, the eleven studies in the Global AMD study were each published in the 21st century, whereas those included in this study distributed from 1990 to 2014. Although no secular trend of AMD was detected in either the Global AMD study or this study, the difference of the estimated AMD prevalence in these two studies can still be partly explained by the difference of ageing demographic structure [3,13,15].

In line with previous population-based investigations and synthesised analysis [3,14,43], this study confirms two common notions of AMD with strong evidence. First, AMD is a degenerative and progressive disease, with the prevalence of AMD dramatically increasing with age, and with age also found to be the only constant risk factor in the presence of any AMD and all its subtypes. Second, the prevalence of early AMD was found to be much higher than that of late AMD. This finding, however, should not be misinterpreted as late AMD contributing a smaller burden. Rather, most individuals with early AMD may not go on to develop the late-stage disease, and late AMD is a much more severe disease than early AMD [9,13]. In this study, the prevalence of late AMD in Chinese people in 2015 was found to be even higher than that of people living in Europe (1.13% vs 0.75), the continentally highest prevalence of both early and late AMD as revealed by the Global AMD study [3]. In view of the large population size in China, this striking finding highlights an urgent need for action on the prevention and treatment of late AMD, given its clinical significance. Compared to GA, the group of NVAMD represents a larger burden in Chinese population because the prevalence and number of people with NVAMD were estimated as around twice higher than those of GA. This phenomenon has been reported in some individual investigations [44,45], although it has not been universally acknowledged [13]. This finding is still of particular importance for the secondary prevention, especially for NVAMD, whose progress to sight loss could be slowed considerably by current treatment approaches – such as the use of anti-vascular endothelial growth factor agents [46,47].

In this study, AMD was found to be more prevalent in urban populations than in rural populations, with possible explanations for this disparity being the difference in environments (eg, UV exposure), as well as lifestyles (eg, education, profession and level of physical activity). While it is not possible to say precisely what the determinants are, this study clearly shows that people living in rural areas with a self-sustained economy are less likely to be affected by AMD [48,49].

A gradient of decreasing prevalence of AMD was noted in increasing latitude, which suggests that the special climate and environmental factors in geographical areas approximating the equator may accelerate the development of AMD. One common hypothesis is that AMD is associated with the amount of insolation [15]. However, the indicator of average annual insolation was only found to be significantly associated with late AMD in this present study. There are two possible reasons for this. First, annual insolation data were averaged over a 22-year period (July 1983 – June 2005), which may represent a considerable time-lag [50,51]. Second, the relation between insolation and the prevalence of AMD may not be a monotone function, the global geo-epidemiology study of AMD revealed higher prevalence rates of AMD in locations with insolation  $\leq 3$  kWh/m<sup>2</sup>/d compared with those with insolation  $> 3$  kWh/m<sup>2</sup>/d [15]. Although this interesting relation was not studied further because of limited data availability, the negative relation between latitude and AMD prevalence is an interesting hypothesis to explore in future Chinese AMD epidemiological studies.

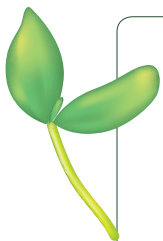
Male gender was indicated as a risk factor for early AMD and NVAMD in the univariable regression analysis of this study. This is in contrast to previous reviews and individual investigations of populations of European ancestry, where females were reported to have a higher risk of developing NVAMD [13,45,52]. However, this study is underpowered to further confirm the observed gender difference in multivariable regression analysis. In the multivariable analysis of the prevalence of any AMD, no evidence of gender difference was found after adjusting for *a priori* demographic and geographic variables. This finding is consistent with the previous Global AMD study [3].

Variation in AMD prevalence and burden was noted in different geographic locations in China. The variation was mainly driven by the different demographic structures and the intrinsic environmental characteristics of these regions. According to the estimates for the six regions, AMD epidemics continue to be concentrated in the most populous South Central China. Taken together, these findings are of a particular public health interest in national health service allocation. Based on this study, more epidemiological investigations are required in order to make the regional estimates of subtypes of AMD in the future.

This work has important implications both in academic and public health areas. Future epidemiological studies of AMD in China would benefit from greater standardisation and improved design, ideally adopting internationally recognised grading systems and presenting results for different subtypes. In addition,

as AMD is a priority eye disease that may lead to severe visual impairment or even blindness, its potential burden on individuals and health systems is particularly large in resource-limited settings [1,2,17]. Thus, localised epidemiological surveys should be conducted in socio-economically disadvantaged provinces, such as Tibet. From the national perspective, the public health impact of AMD is not only limited to the number of people affected, but also brings about multiple diagnostic and treatment challenges arising from this condition [53,54]. It is prudent to address the importance of primary prevention, such as smoking cessation [11,55], lifestyle modification, antioxidant therapy [56], and the use of hats and sunglasses [57]. In the meanwhile, the treatment of NVAMD is already available (although rather expensive) [47,58,59]. Given the remarkable potential economic burden on the society, government efforts must be taken to ensure the availability of health services to address AMD from the points of diagnosis and treatment, and even prevention when available.

To conclude, this systematic review and meta-analysis provides the first comprehensive and up-to-date estimate of AMD prevalence and burden in China. The results from this study indicate that the burden of AMD is substantial in China, with great variance among different subtypes and geographic regions. In the next decade and beyond, the ageing demographic will make this burden even larger. Improved epidemiological studies are still needed to inform optimal implementation of eye care programmes in China.



**Acknowledgements:** The authors would like to thank the China Scholarship Council for the scholarship to Peige Song.

**Funding:** Peige Song is supported by the China Scholarship Council.

**Authorship contribution:** PS and IR conceptualised and designed the study, PS and YD conducted the systematic review. PS conducted the analysis and prepared the first draft. IR, KC and ET critically reviewed the manuscript. All authors have read and approved the final manuscript as submitted.

**Competing interests:** The author completed the Unified Competing Interest form at [www.icmje.org/doi\\_disclosure.pdf](http://www.icmje.org/doi_disclosure.pdf) (available on request from the corresponding author). The authors declare no competing interests. IR is the 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 organisations.

## REFERENCES

- 1 Lim LS, Mitchell P, Seddon JM, Holz FG, Wong TY. Age-related macular degeneration. *Lancet*. 2012;379:1728-38. Medline:22559899 doi:10.1016/S0140-6736(12)60282-7
- 2 Jager RD, Mieler WF, Miller JW. Age-related macular degeneration. *N Engl J Med*. 2008;358:2606-17. Medline:18550876 doi:10.1056/NEJMra0801537
- 3 Wong WL, Su X, Li X, Cheung CMG, Klein R, Cheng C-Y, et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *Lancet Glob Health*. 2014;2:e106-16. Medline:25104651 doi:10.1016/S2214-109X(13)70145-1
- 4 Pascolini D, Mariotti SP. Global estimates of visual impairment: 2010. *Br J Ophthalmol*. 2012;96:614-8. Medline:22133988 doi:10.1136/bjophthalmol-2011-300539
- 5 Casten RJ, Rovner BW, Tasman W. Age-related macular degeneration and depression: a review of recent research. *Curr Opin Ophthalmol*. 2004;15:181-3. Medline:15118503 doi:10.1097/01.icu.0000120710.35941.3f
- 6 Mitchell J, Bradley C. Quality of life in age-related macular degeneration: a review of the literature. *Health Qual Life Outcomes*. 2006;4:97. Medline:17184527 doi:10.1186/1477-7525-4-97
- 7 Hassan SE, Lovie-Kitchin JE, Woods RL. Vision and mobility performance of subjects with age-related macular degeneration. *Optom Vis Sci*. 2002;79:697-707. Medline:12462538 doi:10.1097/00006324-200211000-00007
- 8 Brody BL, Gamst AC, Williams RA, Smith AR, Lau PW, Dolnak D, et al. Depression, visual acuity, comorbidity, and disability associated with age-related macular degeneration. *Ophthalmology*. 2001;108:1893-900. Medline:11581068 doi:10.1016/S0161-6420(01)00754-0
- 9 Lamoureux EL, Mitchell P, Rees G, Cheung G, Yeo I, Lee SY, et al. Impact of early and late age-related macular degeneration on vision-specific functioning. *Br J Ophthalmol*. 2011;95:666-70. Medline:20956281 doi:10.1136/bjo.2010.185207
- 10 Bonastre J, Le Pen C, Anderson P, Ganz A, Berto P, Berdeaux G. The epidemiology, economics and quality of life burden of age-related macular degeneration in France, Germany, Italy and the United Kingdom. *Eur J Health Econ*. 2002;3:94-102. Medline:15609135 doi:10.1007/s10198-002-0104-y
- 11 Thornton J, Edwards R, Mitchell P, Harrison R, Buchan I, Kelly S. Smoking and age-related macular degeneration: a review of association. *Eye (Lond)*. 2005;19:935-44. Medline:16151432 doi:10.1038/sj.eye.6701978

- 12 Swaroop A, Branham KE, Chen W, Abecasis G. Genetic susceptibility to age-related macular degeneration: a paradigm for dissecting complex disease traits. *Hum Mol Genet.* 2007;16 Spec No. 2:R174-82. Medline:17911160 doi:10.1093/hmg/ddm212
- 13 Rudnicka AR, Jarrar Z, Wormald R, Cook DG, Fletcher A, Owen CG. Age and gender variations in age-related macular degeneration prevalence in populations of European ancestry: a meta-analysis. *Ophthalmology.* 2012;119:571-80. Medline:22176800 doi:10.1016/j.ophtha.2011.09.027
- 14 Kawasaki R, Yasuda M, Song SJ, Chen S-J, Jonas JB, Wang JJ, et al. The prevalence of age-related macular degeneration in Asians: a systematic review and meta-analysis. *Ophthalmology.* 2010;117:921-7. Medline:20110127 doi:10.1016/j.ophtha.2009.10.007
- 15 Reibaldi M, Longo A, Pulvirenti A, Avitabile T, Russo A, Cillino S, et al. Geo-epidemiology of age-related macular degeneration: new clues into the pathogenesis. *Am J Ophthalmol.* 2016;161:78-93.e1. Medline:26432929 doi:10.1016/j.ajo.2015.09.031
- 16 de Jong PT. Age-related macular degeneration. *N Engl J Med.* 2006;355:1474-85. Medline:17021323 doi:10.1056/NEJMra062326
- 17 Gottlieb JL. Age-related macular degeneration. *JAMA.* 2002;288:2233-6. Medline:12425683 doi:10.1001/jama.288.18.2233
- 18 Bowens Rickman C, Farsi S, Toth CA, Klingeborn M. Dry age-related macular degeneration: mechanisms, therapeutic targets, and imaging. *Invest Ophthalmol Vis Sci.* 2013;54:ORSF68-80. Medline:24335072 doi:10.1167/iovs.13-12757
- 19 United Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects, the 2015 Revision. 2015. Available: <https://esa.un.org/unpd/wpp/>. Accessed: 8 January 2017.
- 20 Li T, He T, Tan X, Yang S, Li J, Peng Z, et al. Prevalence of age-related cataract in high-selenium areas of China. *Biol Trace Elem Res.* 2009;128:1-7. Medline:18972073 doi:10.1007/s12011-008-8248-y
- 21 Zhou Q, Friedman DS, Lu H, Duan X, Liang Y, Yang X, et al. The epidemiology of age-related eye diseases in Mainland China. *Ophthalmic Epidemiol.* 2007;14:399-407. Medline:18161614 doi:10.1080/09286580701331974
- 22 Yabing Z, Shanlian H, Jinghua C. Burden of wet age-related macular degeneration in China (湿性老年性黄斑变性的疾病负担研究). *Chinese Health Economics.* 2015;1:63-5. *中国卫生经济.*
- 23 Fung IC. Chinese journals: a guide for epidemiologists. *Emerg Themes Epidemiol.* 2008;5:20. Medline:18826604 doi:10.1186/1742-7622-5-20
- 24 Xia J, Wright J, Adams CE. Five large Chinese biomedical bibliographic databases: accessibility and coverage. *Health Info Libr J.* 2008;25:55-61. Medline:18251914 doi:10.1111/j.1471-1842.2007.00734.x
- 25 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg.* 2010;8:336-41. Medline:20171303 doi:10.1016/j.ijsu.2010.02.007
- 26 Stevens GA, Alkema L, Black RE, Boerma JT, Collins GS, Ezzati M, et al. Guidelines for accurate and transparent health estimates reporting: the GATHER statement. *PLoS Med.* 2016;13:e1002056. Medline:27351744 doi:10.1371/journal.pmed.1002056
- 27 Klein R, Davis MD, Magli YL, Segal P, Klein BE, Hubbard L. The Wisconsin age-related maculopathy grading system. *Ophthalmology.* 1991;98:1128-34. Medline:1843453 doi:10.1016/S0161-6420(91)32186-9
- 28 Bird AC, Bressler N, Bressler S, Chisholm I, Coscas G, Davis M, et al. An international classification and grading system for age-related maculopathy and age-related macular degeneration. *Surv Ophthalmol.* 1995;39:367-74. Medline:7604360 doi:10.1016/S0039-6257(05)80092-X
- 29 Seddon JM, Sharma S, Adelman RA. Evaluation of the clinical age-related maculopathy staging system. *Ophthalmology.* 2006;113:260-6. Medline:16458093 doi:10.1016/j.ophtha.2005.11.001
- 30 China Medical Association. Age-related Macular Degeneration Clinical Diagnosis Standard (老年性黄斑变性临床诊断标准). *Zhonghua Yan Ke Za Zhi.* 1987;23:F02. *中华眼科杂志.*
- 31 Hox JJ, Moerbeek M, van de Schoot R. Multilevel analysis: techniques and applications. London: Routledge; 2010.
- 32 Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw.* 2010;36:1-48. doi:10.18637/jss.v036.i03
- 33 Higgins JP, Green S. *Cochrane handbook for systematic reviews of interventions.* Vol 5. Oxford: Wiley Online Library; 2008.
- 34 Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am J Epidemiol.* 2007;165:710-8. Medline:17182981 doi:10.1093/aje/kwk052
- 35 Zhang L, Chow EP, Jing J, Zhuang X, Li X, He M, et al. HIV prevalence in China: integration of surveillance data and a systematic review. *Lancet Infect Dis.* 2013;13:955-63. Medline:24107261 doi:10.1016/S1473-3099(13)70245-7
- 36 National Bureau of Statistics. Tabulation on the 2010 population census of the People's Republic of China. Beijing: China Statistics Press; 2012.
- 37 National Bureau of Statistics. Tabulation on the 2000 population census of the People's Republic of China. Beijing: China Statistics Press; 2002.
- 38 Rudan I, Tomaskovic L, Boschi-Pinto C, Campbell H. Global estimate of the incidence of clinical pneumonia among children under five years of age. *Bull World Health Organ.* 2004;82:895-903. Medline:15654403
- 39 Fowkes FGR, Rudan D, Rudan I, Aboyans V, Denenberg JO, McDermott MM, et al. Comparison of global estimates of prevalence and risk factors for peripheral artery disease in 2000 and 2010: a systematic review and analysis. *Lancet.* 2013;382:1329-40. Medline:23915883 doi:10.1016/S0140-6736(13)61249-0
- 40 Adeloje D, Chua S, Lee C, Basquill C, Papana A, Theodoratou E, et al. Global and regional estimates of COPD prevalence: systematic review and meta-analysis. *J Glob Health.* 2015;5:020415. Medline:26755942 doi:10.7189/jogh.05.020415

- 41 Liu Y, Wen F, Huang S, Luo G, Yan H, Sun Z, et al. Subtype lesions of neovascular age-related macular degeneration in Chinese patients. *Graefes Arch Clin Exp Ophthalmol*. 2007;245:1441-5. Medline:17406882 doi:10.1007/s00417-007-0575-8
- 42 Lee GK, Wong AL, Luk FO, Lai TY. Visual outcome of retinal angiomatous proliferation in Chinese patients following photodynamic therapy or direct laser photocoagulation. *Hong Kong Journal of Ophthalmology*. 2009;13:5-8.
- 43 Klein R, Klein BE, Knudtson MD, Meuer SM, Swift M, Gangnon RE. Fifteen-year cumulative incidence of age-related macular degeneration: the Beaver Dam Eye Study. *Ophthalmology*. 2007;114:253-62. Medline:17270675 doi:10.1016/j.ophtha.2006.10.040
- 44 Owen CG, Fletcher A, Donoghue M, Rudnicka A. How big is the burden of visual loss caused by age related macular degeneration in the United Kingdom? *Br J Ophthalmol*. 2003;87:312-7. Medline:12598445 doi:10.1136/bjo.87.3.312
- 45 Friedman DS, O'Colmain BJ, Munoz B, Tomany SC, McCarty C, De Jong P, et al. Prevalence of age-related macular degeneration in the United States. *Arch Ophthalmol*. 2004;122:564-72. Medline:15078675 doi:10.1001/archophth.122.4.564
- 46 Cook HL, Patel P, Tufail A. Age-related macular degeneration: diagnosis and management. *Br Med Bull*. 2008;85:127-49. Medline:18334518 doi:10.1093/bmb/ldn012
- 47 Brechner RJ, Rosenfeld PJ, Babish JD, Caplan S. Pharmacotherapy for neovascular age-related macular degeneration: an analysis of the 100% 2008 medicare fee-for-service part B claims file. *Am J Ophthalmol*. 2011;151:895.e1. Medline:21310390 doi:10.1016/j.ajo.2010.11.017
- 48 Gemmy Cheung CM, Li X, Cheng C-Y, Zheng Y, Mitchell P, Wang JJ, et al. Prevalence and risk factors for age-related macular degeneration in Indians: a comparative study in Singapore and India. *Am J Ophthalmol*. 2013;155:764-73. Medline:23246273 doi:10.1016/j.ajo.2012.10.013
- 49 Ye H, Zhang Q, Liu X, Cai X, Yu W, Yu S, et al. Prevalence of age-related macular degeneration in an elderly urban Chinese population in China: The Jiangning Eye Study. *Invest Ophthalmol Vis Sci*. 2014;55:6374-80. Medline:25190650 doi:10.1167/iovs.14-14899
- 50 Wild M, Gilgen H, Roesch A, Ohmura A, Long CN, Dutton EG, et al. From dimming to brightening: Decadal changes in solar radiation at Earth's surface. *Science*. 2005;308:847-50. Medline:15879214 doi:10.1126/science.1103215
- 51 Shi G-Y, Hayasaka T, Ohmura A, Chen Z-H, Wang B, Zhao J-Q, et al. Data quality assessment and the long-term trend of ground solar radiation in China. *J Appl Meteorol Climatol*. 2008;47:1006-16. doi:10.1175/2007JAMC1493.1
- 52 Klein R, Klein BE, Jensen SC, Meuer SM. The five-year incidence and progression of age-related maculopathy: the Beaver Dam Eye Study. *Ophthalmology*. 1997;104:7-21. Medline:9022098 doi:10.1016/S0161-6420(97)30368-6
- 53 Brown MM, Brown GC, Stein JD, Roth Z, Campanella J, Beauchamp GR. Age-related macular degeneration: economic burden and value-based medicine analysis. *Can J Ophthalmol*. 2005;40:277-87. Medline:15947797 doi:10.1016/S0008-4182(05)80070-5
- 54 Soubrane G, Cruess A, Lotery A, Pauleikhoff D, Moncs J, Xu X, et al. Burden and health care resource utilization in neovascular age-related macular degeneration: findings of a multicountry study. *Arch Ophthalmol*. 2007;125:1249-54. Medline:17846366 doi:10.1001/archophth.125.9.1249
- 55 Khan JC, Thurlby D, Shahid H, Clayton D, Yates J, Bradley M, et al. Smoking and age related macular degeneration: the number of pack years of cigarette smoking is a major determinant of risk for both geographic atrophy and choroidal neovascularisation. *Br J Ophthalmol*. 2006;90:75-80. Medline:16361672 doi:10.1136/bjo.2005.073643
- 56 Wong IYH, Koo SCY, Chan CWN. Prevention of age-related macular degeneration. *Int Ophthalmol*. 2011;31:73-82. Medline:20862519 doi:10.1007/s10792-010-9397-5
- 57 Tomany SC, Cruickshanks KJ, Klein R, Klein BE, Knudtson MD. Sunlight and the 10-year incidence of age-related maculopathy: the Beaver Dam Eye Study. *Arch Ophthalmol*. 2004;122:750-7. Medline:15136324 doi:10.1001/archophth.122.5.750
- 58 Damico FM, Gasparin F, Scolari MR, Pedral LS, Takahashi BS. New approaches and potential treatments for dry age-related macular degeneration. *Arq Bras Oftalmol*. 2012;75:71-6. Medline:22552424 doi:10.1590/S0004-27492012000100016
- 59 Takeda AL, Colquitt J, Clegg AJ, Jones J. Pegaptanib and ranibizumab for neovascular age-related macular degeneration: a systematic review. *Br J Ophthalmol*. 2007;91:1177-82. Medline:17475698 doi:10.1136/bjo.2007.118562



# Prevalence of chronic obstructive pulmonary disease (COPD) in China in 1990 and 2010

Kit Yee Chan<sup>1,2,\*</sup>, Xue Li<sup>1,\*</sup>, Wanjing Chen<sup>1,3,4,\*</sup>, Peige Song<sup>1,\*</sup>, Nuen Wing Katy Wong<sup>1,5</sup>, Adrienne N Poon<sup>1,6</sup>, Wenyan Jian<sup>7</sup>, Ireneous Y Soyiri<sup>8</sup>, Simon Cousens<sup>9</sup>, Davies Adeloye<sup>1,10</sup>, Aziz Sheikh<sup>8</sup>, Harry Campbell<sup>1</sup>, Igor Rudan<sup>1</sup>; on behalf of Global Health Epidemiology Research Group (GHERG)

<sup>1</sup> Centre for Global Health Research, Usher Institute for Population Health Sciences and Informatics, The University of Edinburgh, Scotland, UK

<sup>2</sup> Nossal Institute for Global Health, University of Melbourne, Melbourne, Australia

<sup>3</sup> Sichuan Academy of Medical Sciences, Chengdu, China

<sup>4</sup> Sichuan Provincial People's Hospital, Chengdu, China

<sup>5</sup> Hospital Authority, Hong Kong, China

<sup>6</sup> Department of Internal Medicine, School of Medicine & Health Sciences, The George Washington University, Washington, D.C., USA

<sup>7</sup> Department of Health Policy and Management, School of Public Health, Peking University Health Science Centre, Beijing, China

<sup>8</sup> Centre for Medical Informatics, Usher Institute for Population Health Sciences and Informatics, The University of Edinburgh, Scotland, UK

<sup>9</sup> Department of Infectious Disease Epidemiology London School of Hygiene and Tropical Medicine, Keppel Street, London, UK

<sup>10</sup> Demography and Social Statistics, and the e-Health Research Cluster, Covenant University, Ota, Ogun State, Nigeria

\*Joint-first authors

## Correspondence to:

Dr Kit Yee Chan  
Centre for Global Health Research  
Usher Institute for Population Health Sciences and Informatics  
The University of Edinburgh  
Teviot Place, Edinburgh EH8 9AG  
Scotland, UK  
k.chan@ed.ac.uk

**Background** Chronic obstructive pulmonary disease (COPD) is set to become the third most frequent cause of death and also the third largest cause of global morbidity by 2020. In China, where the population is aging rapidly, COPD has become one of the leading causes of disability and a large economic burden. An epidemiological assessment of the COPD in China is required, with a focus on the number of cases living with disease, main determinants of the disease and time trends.

**Methods.** We systematically searched large Chinese bibliographic databases and English databases to identify spirometry-based epidemiological studies of the prevalence of COPD in China diagnosed according to GOLD criteria. We estimated age- and gender-specific prevalence of COPD using a multilevel mixed-effect logistic regression. We also presented the time trends of COPD between 1990 and 2010 by age, gender and setting (urban vs rural).

**Findings** In 1990, the prevalence of COPD ranged from 0.49% (95% CI=0.29–0.85) in <20 years group to 20.95% (95% CI=14.04–27.04) in ≥80 years group, and the crude prevalence for China was 2.70% (95% CI=1.86–3.51). In 2010, the prevalence in <20 years was 0.55% (95% CI=0.37–1.04) and in ≥80 years was 22.89% (95% CI=18.13–28.96), with the crude prevalence for China of 3.84% (95% CI=3.30–4.77). The COPD prevalence in males was about two-fold higher than in females, and it increased with increasing age. Between 1990–2010, the total number of Chinese people living with COPD increased by 66.73%, from 30.90 million (95% CI=21.28–40.02) in 1990 to 51.52 million (95% CI=44.26–63.93) in 2010. This increase was most striking in middle age, and greater in females than in males from 30 years up to 64 years. Our estimates, which used an independent approach to acquiring data and development of analytical methods, and were based on a more complete data set, are remarkably similar to those produced recently by the GBD 2013 collaboration, differing by only about 5% in the estimated number of COPD cases in 1990 and by 1% in 2010.

**Conclusions** COPD is a highly prevalent disease in China and its importance is growing steadily. The number of people living with COPD has increased substantially between 1990 and 2010. COPD is more frequent in males and in rural areas. Optimised primary and secondary prevention and treatment is urgently needed to counter this growing trend. Improved epidemiological studies will be required to assist development of more effective strategies of prevention and treatment of COPD in China in the next decade and beyond.

Chronic obstructive pulmonary disease (COPD), characterized by progressive airflow limitation that is not fully reversible, is a major cause of chronic morbidity and mortality. It also carries a substantial financial burden on the society [1–4]. It is estimated that COPD was the sixth most frequent cause of death worldwide in 1990, and will become the third by 2020 [5,6]. Within the same time frame, COPD is also projected to become the third largest cause of global morbidity [5,7]. Cigarette smoking is the major risk factor for COPD, and others include occupational exposures, air pollution, airway hyper-responsiveness and asthma, while genetic predisposition may also play some role [8–11]. The startling spread of this global health epidemic is occurring as a result of the continuous exposure to COPD risk factors and the general aging of the global human population [5,12,13]. In China, where the population is aging very rapidly as a result of improved living conditions, urbanization, high economic growth and family planning [14,15], COPD has become the leading cause of disability [16], resulting in a very large economic burden [17–19]. Furthermore, the Chinese population is also at high risk of developing respiratory diseases because of indoor and outdoor pollution [10,20,21]. All the aforementioned factors indicate that COPD may present a large challenge for the Chinese health care system in the foreseeable future.

Despite growing evidence of epidemiological and economic impact, the availability of accurate disease burden estimates for COPD in China is challenged by the fact that COPD is often underdiagnosed or misdiagnosed [1,11]. Up-to-date information about the prevalence of COPD in general population is essential to inform stakeholders and guide health services allocation [22]. However, the evidence on the prevalence of COPD in the Chinese population is quite inconsistent, probably due to variation in case definition and diagnostic methods, or study population characteristics [20,23]. The latest Global Burden of Disease (GBD) 2013 study revealed a prevalence of COPD of 7.3% (95% CI=6.7%–7.9%) in Chinese population aged 40 years and above in the year 2013 [23]. The GBD 2013 study only included studies of high quality, but the limited number of included studies on COPD prevalence in China could have also limited the authors' ability to address the burden across the whole country. In addition, COPD becomes progressively more common with age, and it could also develop relatively early in life, while previous studies in China mainly focused on middle-aged populations [24,25]. It is, therefore, worth exploring the burden of COPD across the entire lifespan. Another feature of COPD is that its prevalence is likely associated with the greatest socioeconomic inequality in comparison to other common diseases, as COPD is generally believed to disproportionately affect the poor [17,22,26]. In China, difference between the urban and rural setting provides a chance to explore the influence of socioeconomic factors and we are not aware of a systematic analysis of urban–rural differences at the national level.

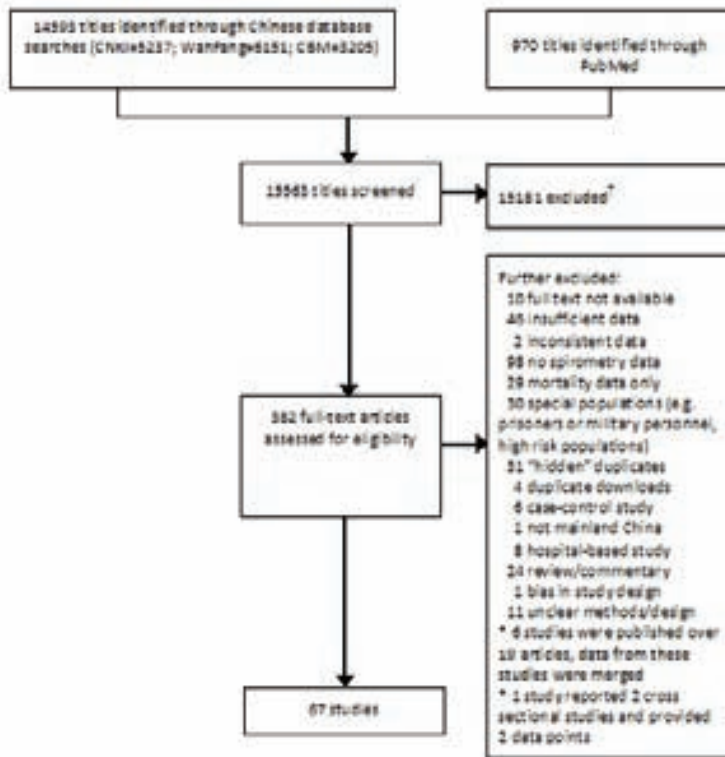
Large Chinese bibliographic databases provide a considerable amount of information that should allow studying the epidemiology of disease in China [27,28]. In addition, spirometry method – which is the gold standard for the diagnosis of COPD [1,29] – has been widely adopted as an accurate estimate of the true burden of COPD in more recent epidemiological studies in China. This provides an opportunity to systematically assess the burden of COPD in China based on a standard case definition and diagnostic method. In this study, we conducted a comprehensive systematic review, in both Chinese and English databases, to analyze published population-based studies of COPD prevalence in China from 1990 onwards, based on spirometry as a diagnostic tool. We also presented the time trends of COPD between 1990 and 2010 by age, gender and setting (urban vs rural).

## METHODS

### Systematic review and data extraction

We used CNKI, WanFang, CBM and PubMed academic databases to retrieve publications on COPD prevalence in China. Parallel searches of the databases were conducted by two independent trained reviewers (Chinese databases: XL and WC; English database: KYC and ANP). The search was limited to studies published between 1990 and 2014 (**Figure 1**). The initial screening of the four databases returned 5237 results from CNKI, 6151 from the WanFang, 3205 from CBM, and 970 from PubMed, respectively. 382 studies were retained after excluding duplicates within and between databases, studies with no numerical values, studies conducted outside mainland China, review articles, viewpoints, conference abstracts. The process is documented in detail in **Figure 1**.

Further steps included obtaining full-text versions of the retained 382 articles and checking whether they met the minimum quality criteria to be used for further analyses and modelling. Studies were subsequently excluded if they were not prospective in design, not population based, did not report spirometry data on COPD prevalence, did not follow the Global Initiative for Chronic Obstructive Lung Disease (GOLD)



**Figure 1.** PRISMA diagram of study selection. † –The large majority was excluded because they were duplicate returns of the same reference under different search terms or in different databases; the others were excluded because they were irrelevant to the topic of our study, provided no numerical estimates, or studied Chinese populations outside of mainland China. \* – These are duplicate publications of the same results in different journals.

standard of a fixed post-bronchodilator ratio of FEV1/FVC less than 0.7, did not have full text available, or contained insufficient or inconsistent data.

### Statistical analysis

In the data extraction procedure, individual studies were used to provide multiple data points which contributed to the overall data set. To take into account the sample size and the availability of different data points from the same study, a multilevel mixed-effect logistic regression was adopted with a restricted cubic spline [30,31]. This was then used to model the COPD prevalence as a function of age. Based on a total of 419 data points extracted from 67 studies with spirometry-based diagnosis of COPD, the gender- and age-specific prevalence of COPD in China was estimated. We further estimated the corresponding 95% confidence intervals (CIs) by applying the semiparametric bootstrapping method [32].

The number of gender-specific COPD cases in China in a population 20 years and older in a given year was calculated by multiplying the estimated gender-specific prevalence for each 5-year age group with the corresponding 5-year population in the same year, available from the United Nations Population Division (UNPD). This was performed for the years 1990, 2000 and 2010. Then, the gender-specific prevalence of COPD for the population aged <20 years was calculated by dividing the aggregated number of COPD cases observed in this age group (which was rather small) with the aggregated size of the corresponding population in this age group. This was done across the age groups 0–4, 5–9, 10–14 and 15–19 years old. The total number of people living with COPD was derived by adding together the numbers of males and females living with COPD, and the overall COPD prevalence was calculated by dividing the total number of COPD cases with the total size of the population in a given year. All the analyses were conducted in R v3.3.0 (R Development Core Team; <http://www.R-project.org>).

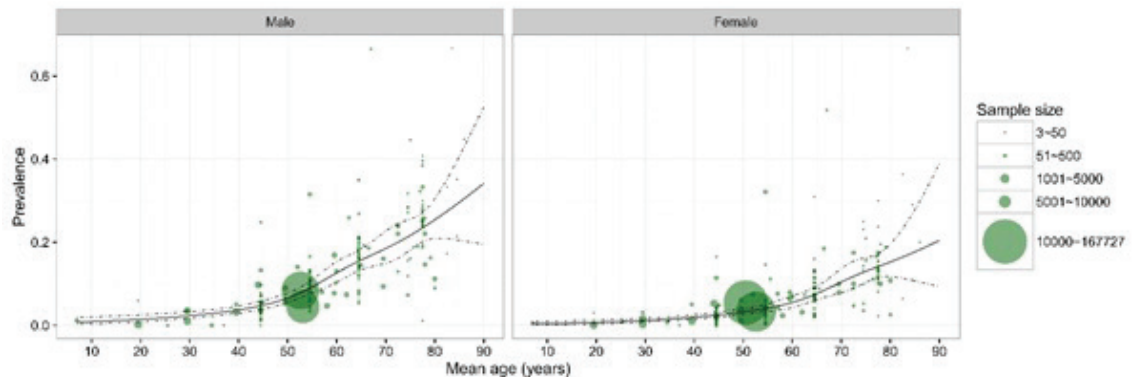
## RESULTS

The characteristics of the 67 retained cross-sectional studies are shown in Table S1 and Table S2 in **Online Supplementary Document**: the studies were typically large, published mainly in the past decade and led by multi-disciplinary teams of specialists. Geographically, the studies covered all 31 provinces/municipalities/autonomous regions of mainland China.

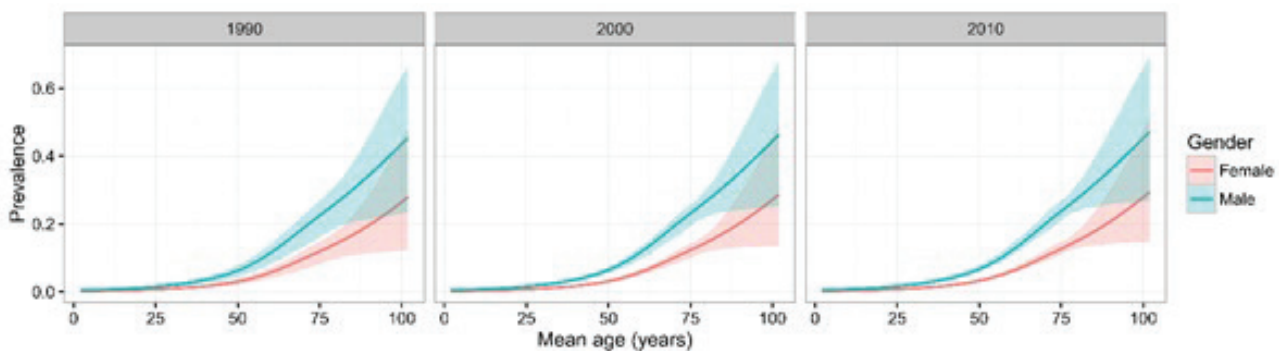
The 67 retained studies incorporated 759 461 people tested for COPD, 45 197 of whom were diagnosed with COPD using the spirometry method. Across most of the age spectrum in adulthood (20 years and

more), there were large sample sizes providing substantial numbers of data points, except for the very old age (>80 years). A total of 47 studies provided the gender-specific prevalence information. **Figure 2** shows the gender-specific relationship between the COPD prevalence and age, based on all the informative data points. Generally, the prevalence of COPD increased steadily with age in both males and females. In males, it reached a prevalence of about 25% by the age of 80 years, and in females it was about twice less prevalent. Based on the information from the 47 studies, the estimates of COPD prevalence for males and females in the years 1990, 2000 and 2010 are shown in **Figure 3**. Across all three years, the prevalence of COPD was consistently greater in males than in females, and the difference was most pronounced in older adults.

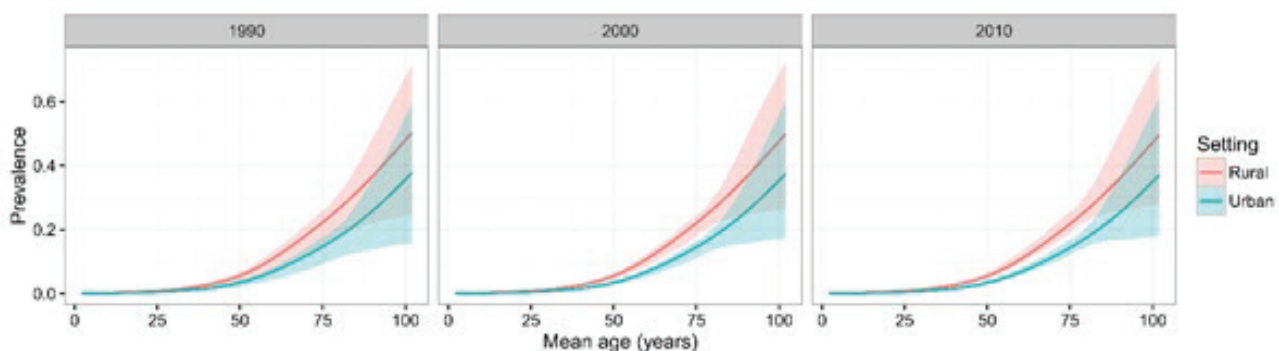
A total of 56 studies provided the setting-specific prevalence information. The comparison of COPD prevalence between urban and rural dwellers in the years 1990, 2000 and 2010 is presented in **Figure 4**. Similar to the comparison of gender-specific prevalence, there was a marked difference in COPD prev-



**Figure 2.** Gender and age-specific prevalence of chronic obstructive pulmonary disease (COPD) in China based on the information from the systematic review. The size of each bubble is proportional to the sample size, where at younger (<20 years) and older (>80 years) groups, the regression lines are based on fewer data points



**Figure 3.** Chronic obstructive pulmonary disease (COPD) prevalence for males and female in the years 1990, 2000 and 2010, with 95% confidence intervals.



**Figure 4.** Chronic obstructive pulmonary disease (COPD) prevalence for urban and rural dwellers in the years 1990, 2000 and 2010, with 95% confidence intervals.

alence between urban and rural dwellers, where people in rural areas had higher COPD prevalence than those in urban areas.

The overall prevalence of COPD and the number of COPD cases were generated based on the UNPD demographic data, where no demographic information for urban and rural areas was provided, so the setting was not included in the final modelling process. The formula generated from the multilevel mixed-effect logistic regression is shown below:

$$\ln(\text{odds}) = -13.388 + 0.004 \times \text{year} + 0.763 \times \text{gender} + 0.058 \times \text{age1} + 0.022 \times \text{age2} - 0.083 \times \text{age3} + 0.026 \times \text{age4} + u_i$$

Where:

odds =  $p/(1-p)$ ,  $p$  indicates the prevalence of COPD

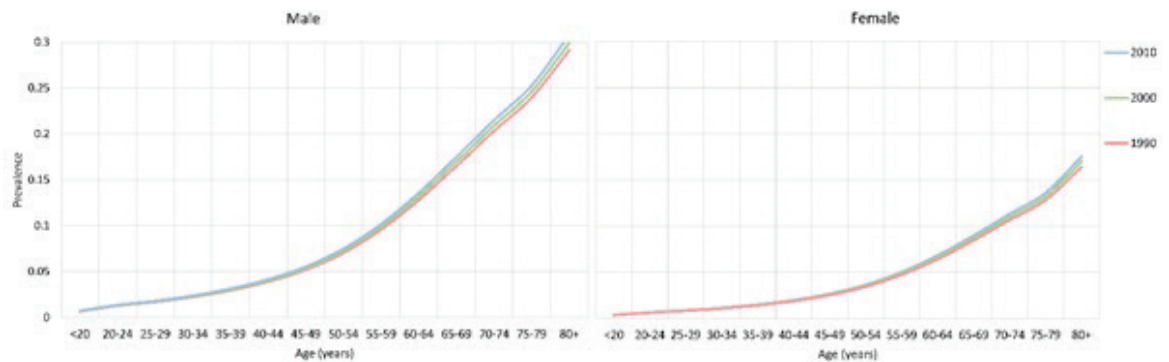
year = calendar year

gender = 1 for males and gender = 0 for females

age1–age4 are variables created in the process of fitting cubic splines (knots: 34.5, 49.5, 59.8, 71.7, 79.0)

$u_i$  = the study level random effect

After applying the estimates of age- and gender-specific COPD prevalence on UN Population Division's demographic data, the prevalence of COPD in populations aged <20 years and ≥80 years were calculated. From 1990 to 2010, the changes in the gender-specific COPD prevalence across all age groups are shown in Figure 5 and Table 1. Over the 20 years considered in this analysis, the prevalence of COPD increased



**Figure 5.** Estimated gender- and age-specific prevalence of chronic obstructive pulmonary disease (COPD) in China in the years 1990, 2000 and 2010.

**Table 1.** Estimated gender- and age-specific prevalence of chronic obstructive pulmonary disease (COPD) in China in the years 1990 and 2010

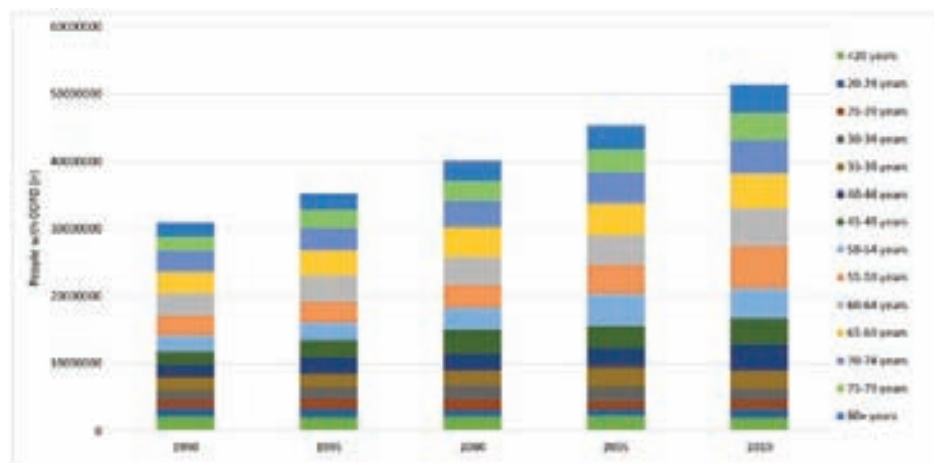
Age (years)	PREVALENCE OF COPD IN 1990 (%)			PREVALENCE OF COPD IN 2010 (%)			RATE OF CHANGE (%; 1990–2010)		
	Male	Female	Overall	Male	Female	Overall	Male	Female	Overall
<20	0.65 (0.38–1.14)	0.31 (0.18–0.54)	0.49 (0.29–0.85)	0.72 (0.49–1.36)	0.34 (0.23–0.65)	0.55 (0.37–1.04)	10.77	9.68	12.24
20–24	1.26 (0.83–1.86)	0.59 (0.40–0.87)	0.94 (0.62–1.38)	1.35 (1.06–2.11)	0.64 (0.50–1.02)	1.01 (0.79–1.59)	7.14	8.47	7.45
25–29	1.67 (1.15–2.33)	0.79 (0.55–1.09)	1.25 (0.86–1.73)	1.79 (1.46–2.59)	0.84 (0.69–1.25)	1.33 (1.09–1.93)	7.19	6.33	6.40
30–34	2.22 (1.56–2.93)	1.05 (0.74–1.39)	1.66 (1.17–2.19)	2.37 (2.00–3.20)	1.12 (0.95–1.53)	1.76 (1.49–2.38)	6.76	6.67	6.02
35–39	2.93 (2.09–3.79)	1.39 (0.99–1.80)	2.20 (1.57–2.84)	3.14 (2.73–4.00)	1.49 (1.31–1.93)	2.33 (2.04–2.98)	7.17	7.19	5.91
40–44	3.88 (2.71–5.01)	1.85 (1.29–2.37)	2.90 (2.03–3.75)	4.15 (3.60–5.15)	1.98 (1.75–2.48)	3.09 (2.70–3.85)	6.96	7.03	6.55
45–49	5.18 (3.60–6.53)	2.48 (1.71–3.17)	3.90 (2.71–4.94)	5.54 (4.83–6.71)	2.66 (2.35–3.26)	4.13 (3.62–5.03)	6.95	7.26	5.90
50–54	7.02 (4.92–8.75)	3.40 (2.33–4.29)	5.31 (3.70–6.65)	7.49 (6.66–8.84)	3.64 (3.26–4.30)	5.63 (5.02–6.65)	6.70	7.06	6.03
55–59	9.55 (6.71–11.92)	4.69 (3.25–5.97)	7.23 (5.06–9.08)	10.18 (8.88–11.97)	5.02 (4.41–5.98)	7.68 (6.72–9.07)	6.60	7.04	6.22
60–64	12.77 (9.00–15.64)	6.39 (4.36–7.98)	9.66 (6.74–11.91)	13.58 (11.98–15.74)	6.82 (5.99–8.06)	10.24 (9.02–11.95)	6.34	6.73	6.00
65–69	16.48 (11.69–19.91)	8.42 (5.73–10.45)	12.38 (8.66–15.09)	17.47 (15.40–19.97)	8.98 (7.83–10.52)	13.24 (11.64–15.27)	6.01	6.65	6.95
70–74	20.32 (14.30–24.44)	10.62 (7.16–13.33)	15.23 (10.56–18.62)	21.48 (18.59–24.84)	11.31 (9.65–13.47)	16.29 (14.03–19.04)	5.71	6.50	6.96
75–79	23.95 (17.27–28.67)	12.80 (8.91–15.90)	17.68 (12.57–21.49)	25.26 (22.67–28.42)	13.61 (12.02–15.65)	18.99 (16.94–21.55)	5.47	6.33	7.41
80+	29.10 (20.17–36.28)	16.38 (10.60–21.86)	20.95 (14.04–27.04)	30.80 (24.85–37.70)	17.51 (13.57–23.02)	22.89 (18.13–28.96)	5.84	6.90	9.26
Total	3.47 (2.40–4.49)	1.88 (1.28–2.47)	2.70 (1.86–3.51)	4.91 (4.22–6.06)	2.69 (2.31–3.37)	3.84 (3.30–4.77)	41.50	43.09	42.22

slightly in both males and females, which was also indicated in the model formula, where the annual rate of change in prevalence on a log odds scale was only 0.004. In 1990, the prevalence of COPD ranged from 0.49% (95% CI=0.29–0.85) in younger people aged under 20 years to 20.95% (95% CI=14.04–27.04) for those who were 80 years of age or older, and the overall prevalence was 2.70% (95% CI=1.86–3.51). In 2010, the prevalence of COPD ranged from 0.55% (95% CI=0.37–1.04) in people aged under 20 years to 22.89% (95% CI=18.13–28.96) in people aged 80 years and above, with the overall prevalence in the Chinese population increasing to the level of 3.84% (95% CI=3.30–4.77). The COPD prevalence between males and females differed consistently from 1990 to 2010, where the prevalence of COPD in males was around two-fold higher than in females, and this difference increased with increasing age. During 1990 to 2010, the overall prevalence of COPD increased by 42.22%. The most significant increases were in younger population who were under 20 years (12.24%) and older population aged 80 years and above (9.26%). The increasing rates were similar in males and females.

The estimates of the number of people living with COPD in China are shown in **Table 2** and **Figure 6**. With the aging of the Chinese population during 1990–2010, the total number of Chinese people living

**Table 2.** Estimated numbers of people living with COPD in China in the years 1990 and 2010, and the rate of change from 1990 to 2010, by gender and age group

Age (years)	PEOPLE LIVING WITH COPD IN 1990 (MILLION)			PEOPLE LIVING WITH COPD IN 2010 (MILLION)			RATE OF CHANGE (% <sub>1990–2010</sub> )		
	Male	Female	Overall	Male	Female	Overall	Male	Female	Overall
<20	1.50 (0.88–2.61)	0.66 (0.39–1.15)	2.16 (1.28–3.77)	1.44 (0.98–2.72)	0.57 (0.39–1.09)	2.00 (1.37–3.80)	-4.00	-13.64	-7.41
20–24	0.83 (0.55–1.22)	0.36 (0.24–0.53)	1.19 (0.79–1.75)	0.85 (0.66–1.32)	0.36 (0.28–0.58)	1.21 (0.95–1.91)	2.41	0.00	1.68
25–29	0.86 (0.59–1.19)	0.38 (0.26–0.52)	1.23 (0.85–1.71)	0.93 (0.76–1.34)	0.41 (0.34–0.61)	1.34 (1.10–1.95)	8.14	7.89	8.94
30–34	1.00 (0.70–1.32)	0.43 (0.30–0.57)	1.42 (1.00–1.88)	1.12 (0.94–1.51)	0.50 (0.43–0.68)	1.62 (1.37–2.19)	12.00	16.28	14.08
35–39	1.36 (0.97–1.75)	0.58 (0.41–0.75)	1.94 (1.38–2.50)	1.91 (1.66–2.43)	0.87 (0.77–1.13)	2.78 (2.43–3.56)	40.44	50.00	43.30
40–44	1.28 (0.90–1.66)	0.56 (0.39–0.72)	1.85 (1.29–2.38)	2.60 (2.26–3.22)	1.17 (1.04–1.47)	3.77 (3.29–4.70)	103.13	108.93	103.78
45–49	1.34 (0.93–1.69)	0.58 (0.40–0.74)	1.92 (1.33–2.43)	2.68 (2.33–3.24)	1.23 (1.08–1.50)	3.90 (3.42–4.74)	100.00	112.07	103.13
50–54	1.69 (1.19–2.11)	0.73 (0.50–0.92)	2.42 (1.69–3.03)	3.13 (2.79–3.70)	1.42 (1.27–1.67)	4.55 (4.06–5.37)	85.21	94.52	88.02
55–59	2.08 (1.46–2.60)	0.93 (0.65–1.19)	3.01 (2.11–3.78)	4.24 (3.70–4.98)	1.96 (1.72–2.33)	6.19 (5.42–7.31)	103.85	110.75	105.65
60–64	2.21 (1.56–2.71)	1.05 (0.72–1.32)	3.27 (2.28–4.03)	3.80 (3.35–4.40)	1.86 (1.64–2.20)	5.66 (4.99–6.61)	71.95	77.14	73.09
65–69	2.10 (1.49–2.53)	1.11 (0.76–1.38)	3.21 (2.24–3.91)	3.42 (3.02–3.91)	1.74 (1.52–2.04)	5.16 (4.54–5.95)	62.86	56.76	60.75
70–74	1.92 (1.35–2.31)	1.11 (0.75–1.39)	3.03 (2.10–3.71)	3.25 (2.81–3.75)	1.78 (1.52–2.12)	5.03 (4.33–5.87)	69.27	60.36	66.01
75–79	1.30 (0.94–1.55)	0.89 (0.62–1.11)	2.19 (1.56–2.66)	2.54 (2.28–2.86)	1.59 (1.41–1.83)	4.14 (3.69–4.70)	95.38	78.65	89.04
80+	1.03 (0.71–1.28)	1.03 (0.67–1.38)	2.06 (1.38–2.66)	2.27 (1.83–2.78)	1.90 (1.47–2.50)	4.17 (3.30–5.27)	120.39	84.47	102.43
Total	20.50 (14.22–26.53)	10.40 (7.06–13.67)	30.90 (21.28–40.02)	34.18 (29.37–42.16)	17.36 (14.88–21.75)	51.52 (44.26–63.93)	66.73	66.92	66.73



**Figure 6.** Estimated numbers of people living with chronic obstructive pulmonary disease (COPD) in China by year and age group.

with COPD increased by 66.73%, from 30.90 (95% CI=21.28–40.02) million in 1990 to 51.52 (95% CI=44.26–63.93) million in 2010. Throughout this period, the number of COPD cases decreased by 7.41% in younger people aged under 20 years. The most significant increases were in those who were middle-aged (40–49 years and 55–59 years) and at older ages (80 years and above), with rates of change in the absolute number of cases in those age-groups above 100%. This increase was greater in females than in males from 30 years up to 64 years. In 2010, around two-thirds (66.34%) of the people living with COPD were males in China.

## DISCUSSION

In this study, by systematically reviewing all published evidence of COPD prevalence databased on spirometry, and applying strict inclusion and exclusion criteria, we presented comprehensive and data-driven estimates of the COPD prevalence in China across the entire age range, and the differences between males and females, and urban and rural dwellers in China. As suggested by the Global Initiative for Chronic Obstructive Lung Disease (GOLD), a fixed post-bronchodilator ratio of FEV1/FVC less than 0.7 was used as the primary indicator of COPD diagnosis in our study [1]. In the present study, we estimated a national GOLD-defined COPD prevalence of 2.70% in 1990 and 3.84% in 2010, corresponding to 30.90 million and 51.52 million people living with COPD in 1990 and 2010 respectively. Our estimate for the year 2010 was similar to the estimate presented in the GBD 2013 study, which reported 32.44 million of cases in 1990 and 54.79 million of cases in 2013 [23]. Given that the estimates presented in GBD 2013 study were based on a limited number of studies that were mostly published after 2000, any differences in compared COPD number of cases for the year 1990 between these two studies – amounting to 1.54 million (or about 5% difference) – may be explained through this lack of data in the GBD 2013 study. Our estimate in 2010 is relevant to a point in time 3 years earlier than the GBD 2013 estimate, but based on our time trends we would expect our estimate for 2013 to be nearly identical to the GBD estimate, ie, just over 54 million cases (with difference of about 1%). In comparison with the results of the GBD 2013 study, where a slight decrease of a standardized COPD prevalence was reported between 1990 and 2013 (from 3.7% to 3.6% overall prevalence in the population), our study observed an overall increase trend of COPD prevalence during the two decades between 1990 to 2010 – from 2.70% to 3.84%. Our estimates were much in line with the increasing trend of COPD prevalence globally and in the Western Pacific region [4,11]. Further analysis from an epidemiological perspective will remain to be required to explore the temporal distribution of COPD prevalence in China. The detailed and systematic estimates of COPD prevalence and the number of cases in this study constitute the best currently available basis for policymaking, planning, and allocation of health and welfare resources related to the burden of COPD in China. The strength of this study arises from several measures that we took to derive the estimates. First, we did a comprehensive systematic review and thus included all available studies. Second, of all the measurement methods, we chose spirometry method as the diagnosis criteria, which served to avoid potential bias arising from the variation of measurement tools. Third, the estimate of prevalence was based on study points that came from many different locations China, which should ensure representativeness for the whole nation.

In view of the relation between COPD prevalence and age, our study confirms that COPD is a progressive and degenerative disease [11,33]. By estimating the prevalence of COPD in the Chinese population over the full age-span, we enabled a public health approach to prevention and treatment of COPD in younger population. It was important to note that COPD can also exist in early life, as suggested by previous studies [24,25]. Younger populations should therefore not be excluded from COPD screening programs. In our study, although the number of COPD people in <20 years decreased slightly during 1990 to 2010, the prevalence rate kept increase. Early diagnosis is very important, and disease-modifying therapies that can delay onset could have considerable potential for reducing age-specific prevalence [34].

As expected, the prevalence in males in China was much higher than in females. The excess was most likely to be explained by historic patterns of smoking and occupational exposures in Chinese men [4,20,35]. From a physiological perspective, it is also widely hypothesized that women are more vulnerable to the lung-damaging effects of cigarette smoking and biomass fuels, and more prone to develop COPD than men when they are at the same level of hazard exposure [35–37]. Our study showed a slightly higher increase of COPD prevalence and of the number of affected individuals in females than in males during 1990 and 2010. Based on this trend, it should be expected that prevalence of COPD in Chinese women is set to rise in the coming years, partly because of a markedly increasing rate of cigarette smoking in Chinese women [38,39].

The prevalence of COPD was consistently higher in rural areas than in urban areas in our study. This was in contrast with the global estimates, which show that COPD is more frequent in urban areas [4,40]. The possible explanation is the variation in classification of rural and urban settings across the world. In China, rural areas are generally those with less (or lower-quality) health resources and worse health outcomes in comparison to urban areas [41–44]. This urban–rural disparity of COPD prevalence may be associated with higher smoking rates, but also with prevalent indoor air pollution resulting from the burning of wood and other biomass fuels in rural populations [10,26,45]. In addition, lower socioeconomic status, lower health resources quality, and worse quality of cigarettes may also have contributed [20,21,26].

A striking increase of 66.73% in the absolute number of estimated COPD cases between 1990 and 2010 in China was indicated by our analysis. This increase was most prominent in middle-aged and older populations. The primary drivers of this increase were longer life expectancy in China, which led to an increase of the number of individuals in older age groups, and continuous exposure to risk factors, which led to an increase in prevalence among the middle-aged. The estimated number of 51.52 million people living with COPD in China in 2010 represents a very large burden for both China and the world, where an estimated number of 328 million people have COPD [46,47]. With the current trends of ageing in China [14,15], the numbers of individuals affected with COPD will continue to increase, particularly in older population. Elderly with medical care needs will constitute a burden on the health care system, and support for family caregivers may also be needed [14,44]. In young and middle-aged populations, the limitations in their ability to work and their overall mobility may not only lead to a reduced quality of life, but also bring an economic challenge to both individuals and the society [18–20]. Even among the affected individuals who only have mild symptoms, the effects of COPD may result in impaired quality of life. These effects might be profound in many poor settings, where continuous exposure to risk factors could accelerate the progress of the disease process [48,49].

Our study also had several potential shortcomings. Significant heterogeneity was seen between all of our included studies. This is not surprising, given that China is a diverse country with highly varied culture, levels of development and demographic characteristics. Although only studies that adopted the spirometry method and the GOLD criterion were included, in order to reduce the methodological variability, heterogeneity may still exist because of the differences among investigators in operating protocols, levels of training, adherence to guidelines and variations in implementation of the guidelines. Another possible cause for concern is that the results of meta-regression only took into account a limited set of covariates (ie, gender, age and urbanicity) because these were the only covariates that were broadly available. The scarcity of studies adopting unified definition of other covariates limited our ability to explore more possible effects. Our estimates, therefore, do not take into account the role of special sub-groups exposed to risk factors, such as smokers, household members who are exposed to biomass fuels used for home heating and cooking, workers with occupational exposures to dust, and people with tuberculosis. This is a shortcoming that needs to be addressed in future epidemiological studies of COPD in China. In addition, the estimates of prevalence in younger and older age groups were clearly less certain due to scarce data and require further refinement. Finally, although GOLD criteria are widely accepted as a diagnostic tool for COPD, they are prone to false-negative results among younger and false-positive among older adults, which should be taken into account [1,26]. We were also unable to account for the severity of COPD, particularly to characterise the burden of the early-stage COPD.

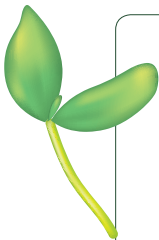
A growing amount of evidence suggests that ambient air pollution is a risk factor for COPD [10,50]. This hypothesis may be of particular interest in future COPD epidemiology studies in China, where a dramatic increase in emissions of ambient air pollutants has become a large societal concern [51]. Concern is often focused on observed air pollution levels in very large cities, but our analysis didn't find support for COPD prevalence being a larger problem in urban as opposed to rural areas.

Our work has implications in both the academic and public health areas. Epidemiological studies that are typically available in China on COPD could be designed better, with standardized diagnostic criteria and much better presentation of results, taking into account different covariates. Present gaps would ideally be filled by a national disease surveillance system, but this would need considerable and sustainable funding and training. Perhaps a more realistic approach would include high-quality epidemiology studies in large samples, such as Kadoorie study – which provided a large portion of all the data available for our study in terms of sample size [52,53]. Adopting standardized measures of COPD, with accurate assessments of the stage and severity of COPD, effectiveness of treatment and clear definition of study population sub-groups in relation to their risk exposure would all help. Currently, most people with COPD can only receive a diagnosis in the late stage of disease when symptoms are rather obvious [18]. Although

COPD is not a fully reversible disease, the benefits of earlier diagnosis and treatment have been well established [11,18]. Efforts should be taken to improve the quality and availability of health care resources to address COPD both from the point of prevention, diagnosis and treatment, especially in rural areas where the rates of COPD prevalence are higher. In the meanwhile, the importance of primary prevention should also be addressed, especially in primary care settings. Health education should focus on targeted populations based on the existing evidence [21]. Advocated measures include implementation of tobacco-free policies, adequate treatment of asthma and comprehensive strategies on reducing indoor pollution. [11,18,21].

## CONCLUSIONS

In conclusion, results from this study have shown that COPD is a highly prevalent disease in China and its importance is growing steadily. The number of people living with COPD has increased substantially and nearly doubled between 1990 and 2010. COPD is more frequent in males and in rural areas. Optimal intervention delivery – both in primary and secondary prevention and treatment – is urgently needed to counter this growing trend. Improved epidemiological studies will be required to assist development of more effective strategies of prevention and treatment of COPD in China in the next decade and beyond.



**Funding:** None.

**Authorship declaration:** KYC and IR designed the study. KYC, XL, WC and ANP conducted literature searches in Chinese and English databases. PS, SC and IYS designed the analyses and analysed the data. NWKW, WJ, DA, AS and HC provided important intellectual content to multiple drafts of the paper. PS, KYC and IR wrote the final draft of the paper.

**Competing interests:** The authors completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author), and declare no conflict of interest. IR and HC are Editors-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 organisations.

## REFERENCES

- 1 Vestbo J, Hurd SS, Agustí AG, Jones PW, Vogelmeier C, Anzueto A, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med*. 2013;187:347-65. Medline:22878278 doi:10.1164/rccm.201204-0596PP
- 2 Decramer M, Janssens W. Chronic obstructive pulmonary disease and comorbidities. *Lancet Respir Med*. 2013;1:73-83. Medline:24321806 doi:10.1016/S2213-2600(12)70060-7
- 3 Agustí AG, Noguera A, Sauleda J, Sala E, Pons J, Busquets X. Systemic effects of chronic obstructive pulmonary disease. *Eur Respir J*. 2003;21:347-60. Medline:12608452 doi:10.1183/09031936.03.00405703
- 4 Adeloye D, Chua S, Lee C, Basquill C, Papan A, Theodoratou E, et al. Global and regional estimates of COPD prevalence: Systematic review and meta-analysis. *J Glob Health*. 2015;5:020415. Medline:26755942 doi:10.7189/jogh.05.020415
- 5 Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *Lancet*. 1997;349:1498-504. Medline:9167458 doi:10.1016/S0140-6736(96)07492-2
- 6 Raheison C, Girodet P. Epidemiology of COPD. *Eur Respir Rev*. 2009;18:213-21. Medline:20956146 doi:10.1183/09059180.00003609
- 7 Soriano JB, Ancochea J, Miravittles M, García-Río F, Duran-Tauleria E, Muñoz L, et al. Recent trends in COPD prevalence in Spain: a repeated cross-sectional survey 1997–2007. *Eur Respir J*. 2010;36:758-65. Medline:19996189 doi:10.1183/09031936.00138409
- 8 Mannino DM, Braman S. The epidemiology and economics of chronic obstructive pulmonary disease. *Proc Am Thorac Soc*. 2007;4:502-6. Medline:17878461 doi:10.1513/pats.200701-001FM
- 9 Trupin L, Earnest G, San Pedro M, Balmes J, Eisner M, Yelin E, et al. The occupational burden of chronic obstructive pulmonary disease. *Eur Respir J*. 2003;22:462-9. Medline:14516136 doi:10.1183/09031936.03.00094203
- 10 Salvi SS, Barnes PJ. Chronic obstructive pulmonary disease in non-smokers. *Lancet*. 2009;374:733-43. Medline:19716966 doi:10.1016/S0140-6736(09)61303-9
- 11 Mannino DM, Buist AS. Global burden of COPD: risk factors, prevalence, and future trends. *Lancet*. 2007;370:765-73. Medline:17765526 doi:10.1016/S0140-6736(07)61380-4
- 12 Chapman KR, Mannino DM, Soriano J, Vermeire P, Buist AS, Thun M, et al. Epidemiology and costs of chronic obstructive pulmonary disease. *Eur Respir J*. 2006;27:188-207. Medline:16387952 doi:10.1183/09031936.06.00024505

- 13 Buist AS, McBurnie MA, Vollmer WM, Gillespie S, Burney P, Mannino DM, et al. International variation in the prevalence of COPD (the BOLD Study): a population-based prevalence study. *Lancet*. 2007;370:741-50. Medline:17765523 doi:10.1016/S0140-6736(07)61377-4
- 14 Woo J, Kwok T, Sze F, Yuan H. Ageing in China: health and social consequences and responses. *Int J Epidemiol*. 2002;31:772-5. Medline:12177017 doi:10.1093/ije/31.4.772
- 15 Flaherty JH, Liu ML, Ding L, Dong B, Ding Q, Li X, et al. China: the aging giant. *J Am Geriatr Soc*. 2007;55:1295-300. Medline:17661972 doi:10.1111/j.1532-5415.2007.01273.x
- 16 Gao J, Prasad N. Chronic obstructive pulmonary disease in China: the potential role of indacaterol. *J Thorac Dis*. 2013;5:549. Medline:23991315
- 17 Barnes PJ. Chronic obstructive pulmonary disease: a growing but neglected global epidemic. *PLoS Med*. 2007;4:e112. Medline:17503959 doi:10.1371/journal.pmed.0040112
- 18 Pauwels RA, Rabe KF. Burden and clinical features of chronic obstructive pulmonary disease (COPD). *Lancet*. 2004;364:613-20. Medline:15313363 doi:10.1016/S0140-6736(04)16855-4
- 19 Halpin DM, Miravittles M. Chronic obstructive pulmonary disease: the disease and its burden to society. *Proc Am Thorac Soc*. 2006;3:619-23. Medline:16963544 doi:10.1513/pats.200603-0935S
- 20 Fang X, Wang X, Bai C. COPD in China: the burden and importance of proper management. *Chest*. 2011;139:920-9. Medline:21467059 doi:10.1378/chest.10-1393
- 21 Lin H-H, Murray M, Cohen T, Colijn C, Ezzati M. Effects of smoking and solid-fuel use on COPD, lung cancer, and tuberculosis in China: a time-based, multiple risk factor, modelling study. *Lancet*. 2008;372:1473-83. Medline:18835640 doi:10.1016/S0140-6736(08)61345-8
- 22 Yach D, Hawkes C, Gould CL, Hofman KJ. The global burden of chronic diseases: overcoming impediments to prevention and control. *JAMA*. 2004;291:2616-22. Medline:15173153 doi:10.1001/jama.291.21.2616
- 23 Yin P, Wang H, Vos T, Li Y, Liu S, Liu Y, et al. A subnational analysis for mortality and prevalence of chronic obstructive pulmonary disease in China 1990–2013: Findings from Global Burden of Disease Study (GBD) 2013. *CHEST Journal*. 2016.
- 24 Postma DS, Bush A, van den Berge M. Risk factors and early origins of chronic obstructive pulmonary disease. *Lancet*. 2015;385:899-909. Medline:25123778 doi:10.1016/S0140-6736(14)60446-3
- 25 Svanes C, Sunyer J, Plana E, Dharmage S, Heinrich J, Jarvis D, et al. Early life origins of chronic obstructive pulmonary disease. *Thorax*. 2010;65:14-20. Medline:19729360 doi:10.1136/thx.2008.112136
- 26 Zhong N, Wang C, Yao W, Chen P, Kang J, Huang S, et al. Prevalence of chronic obstructive pulmonary disease in China: a large, population-based survey. *Am J Respir Crit Care Med*. 2007;176:753-60. Medline:17575095 doi:10.1164/rccm.200612-1749OC
- 27 Xia J, Wright J, Adams CE. Five large Chinese biomedical bibliographic databases: accessibility and coverage. *Health Info Libr J*. 2008;25:55-61. Medline:18251914 doi:10.1111/j.1471-1842.2007.00734.x
- 28 Fung IC. Chinese journals: a guide for epidemiologists. *Emerg Themes Epidemiol*. 2008;5:20. Medline:18826604 doi:10.1186/1742-7622-5-20
- 29 Celli BR. The importance of spirometry in COPD and asthma: effect on approach to management. *Chest*. 2000;117 Suppl:15S-19S. Medline:10673468 doi:10.1378/chest.117.2\_suppl.15S
- 30 Thompson SG, Higgins J. How should meta-regression analyses be undertaken and interpreted? *Stat Med*. 2002;21:1559-73. Medline:12111920 doi:10.1002/sim.1187
- 31 Desquilbet L, Mariotti F. Dose-response analyses using restricted cubic spline functions in public health research. *Stat Med*. 2010;29:1037-57. Medline:20087875
- 32 Barr DJ. Random effects structure for testing interactions in linear mixed-effects models. *Front Psychol*. 2013;4:328. Medline:23761778 doi:10.3389/fpsyg.2013.00328
- 33 Ito K, Barnes PJ. COPD as a disease of accelerated lung aging. *Chest*. 2009;135:173-80. Medline:19136405 doi:10.1378/chest.08-1419
- 34 Wilkinson TM, Donaldson GC, Hurst JR, Seemungal TA, Wedzicha JA. Early therapy improves outcomes of exacerbations of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2004;169:1298-303. Medline:14990395 doi:10.1164/rccm.200310-1443OC
- 35 Silverman EK, Weiss ST, Drazen JM, Chapman HA, Carey V, Campbell EJ, et al. Gender-related differences in severe, early-onset chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2000;162:2152-8. Medline:11112130 doi:10.1164/ajrccm.162.6.2003112
- 36 Han MK, Postma D, Mannino DM, Giardino ND, Buist S, Curtis JL, et al. Gender and chronic obstructive pulmonary disease: why it matters. *Am J Respir Crit Care Med*. 2007;176:1179-84. Medline:17673696 doi:10.1164/rccm.200704-553CC
- 37 S rheim IC, Johannessen A, Gulsvik A, Bakke PS, Silverman EK, DeMeo DL. Gender differences in COPD: are women more susceptible to smoking effects than men? *Thorax*. 2010;65:480-5. Medline:20522842 doi:10.1136/thx.2009.122002
- 38 Zhang J, OU JX, BAI CX. Tobacco smoking in China: prevalence, disease burden, challenges and future strategies. *Respirology*. 2011;16:1165-72. Medline:21910781 doi:10.1111/j.1440-1843.2011.02062.x
- 39 Qian J, Cai M, Gao J, Tang S, Xu L, Critchley JA. Trends in smoking and quitting in China from 1993 to 2003: National Health Service Survey data. *Bull World Health Organ*. 2010;88:769-76. Medline:20931062 doi:10.2471/BLT.09.064709
- 40 Halbert RJ, Natoli J, Gano A, Badamgarav E, Buist AS, Mannino DM. Global burden of COPD: systematic review and meta-analysis. *Eur Respir J*. 2006;28:523-32. Medline:16611654 doi:10.1183/09031936.06.00124605

- 41 Yu J. Multidimensional poverty in China: Findings based on the CHNS. *Soc Indic Res.* 2013;112:315-36. doi:10.1007/s11205-013-0250-x
- 42 Meng Q, Xu L, Zhang Y, Qian J, Cai M, Xin Y, et al. Trends in access to health services and financial protection in China between 2003 and 2011: a cross-sectional study. *Lancet.* 2012;379:805-14. Medline:22386034 doi:10.1016/S0140-6736(12)60278-5
- 43 Gong P, Liang S, Carlton EJ, Jiang Q, Wu J, Wang L, et al. Urbanisation and health in China. *Lancet.* 2012;379:843-52. Medline:22386037 doi:10.1016/S0140-6736(11)61878-3
- 44 Teng H, Cao Z, Liu J, Liu P, Hua W, Yang Y, et al. Health status and burden of health care costs among urban elderly in China. *Asia Pac J Public Health.* 2015;27(suppl):61S-8S. Medline:25673281 doi:10.1177/1010539515571579
- 45 Zhou Y, Zou Y, Li X, Chen S, Zhao Z, He F, et al. Lung function and incidence of chronic obstructive pulmonary disease after improved cooking fuels and kitchen ventilation: a 9-year prospective cohort study. *PLoS Med.* 2014;11:e1001621. Medline:24667834 doi:10.1371/journal.pmed.1001621
- 46 Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet.* 2012;380:2163-96. Medline:23245607 doi:10.1016/S0140-6736(12)61729-2
- 47 López-Campos JL, Tan W, Soriano JB. Global burden of COPD. *Respirology.* 2016;21:14-23. Medline:26494423 doi:10.1111/resp.12660
- 48 Jones PW. Issues concerning health-related quality of life in COPD. *Chest.* 1995;107(5 Suppl):187S-93S. Medline:7743825 doi:10.1378/chest.107.5\_Supplement.187S
- 49 Miravittles M, Soriano JB, Garcia-Rio F, Muñoz L, Duran-Tauleria E, Sanchez G, et al. Prevalence of COPD in Spain: impact of undiagnosed COPD on quality of life and daily life activities. *Thorax.* 2009;64:863-8. Medline:19553233 doi:10.1136/thx.2009.115725
- 50 Gan WQ, FitzGerald JM, Carlsten C, Sadatsafavi M, Brauer M. Associations of ambient air pollution with chronic obstructive pulmonary disease hospitalization and mortality. *Am J Respir Crit Care Med.* 2013;187:721-7. Medline:23392442 doi:10.1164/rccm.201211-2004OC
- 51 Kan H, Chen R, Tong S. Ambient air pollution, climate change, and population health in China. *Environ Int.* 2012;42:10-9. Medline:21440303 doi:10.1016/j.envint.2011.03.003
- 52 Chen Z, Lee L, Chen J, Collins R, Wu F, Guo Y, et al. Cohort profile: the Kadoorie Study of chronic disease in China (KSCDC). *Int J Epidemiol.* 2005;34:1243-9. Medline:16131516 doi:10.1093/ije/dyi174
- 53 Chen Z, Chen J, Collins R, Guo Y, Peto R, Wu F, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol.* 2011;40:1652-66. Medline:22158673 doi:10.1093/ije/dyr120

# National and subnational prevalence and burden of glaucoma in China: A systematic analysis

Peige Song<sup>1</sup>, Jiawen Wang<sup>2</sup>, Kajo Bucan<sup>3</sup>, Evropi Theodoratou<sup>1</sup>, Igor Rudan<sup>1\*</sup>, Kit Yee Chan<sup>1\*</sup>; on behalf of the Global Health Epidemiology Research Group (GHERG)

<sup>1</sup> Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, Scotland, UK

<sup>2</sup> Institute of Medical Humanities, Peking University, Beijing, China

<sup>3</sup> Department of Ophthalmology, University of Split Hospital Centre, Split, Croatia

\* Joint last authors

**Background** Glaucoma, the second leading cause of blindness, affects approximately 64.3 million individuals worldwide. In China, demographic ageing is in rapid progress. Yet detailed and up-to-date estimates of the scale of glaucoma are rare. We aimed to quantify and understand the prevalence and burden of glaucoma in China from 1990 to 2015, with projections until 2050.

**Methods** For this systematic review and meta-analysis, we searched China National Knowledge Infrastructure (CNKI), Wanfang, Chinese Biomedicine Literature Database (CBM-SinoMed), PubMed, Embase and Medline using comprehensive search strategies to identify all relevant articles that have reported the prevalence of glaucoma in the general Chinese population. We used a multilevel mixed-effect meta-regression to estimate the prevalence rates of primary open-angle glaucoma (POAG) and primary angle-closure glaucoma (PACG), and a random-effects meta-analysis to pool the overall prevalence of secondary glaucoma. United Nations population data were used to estimate and project the number of people with glaucoma from 1990 to 2050. Univariable and multivariable meta-regressions were conducted to assess the association between the prevalence of POAG and PACG and relevant demographic and geographic factors. The national burden of POAG and PACG in the years 2000 and 2010 were distributed to six geographic regions accordingly.

**Results** From 1990 to 2015, the prevalence of all glaucoma ranged from 2.59% (95% CI=1.96–3.49) to 2.58% (95% CI=1.94–3.47). For different subtypes of glaucoma, the overall prevalence of POAG ranged from 1.03% (95% CI=0.67–1.58) in 1990 to 1.02% (95% CI=0.67–1.57) in 2015, PACG from 1.41% (95% CI=1.18–1.68) to 1.40% (95% CI=1.17–1.68). The overall prevalence of secondary glaucoma was 0.15% (95% CI=0.10–0.23) during this period. The number of people with all glaucoma in China was 5.92 million (95% CI=4.47–7.97) in 1990, and 13.12 million (95% CI=9.88–17.68) in 2015. This increasing trend was also witnessed in different subtypes of glaucoma. The number of people affected by POAG increased from 2.35 million (95% CI=1.54–3.60) in 1990 to 5.22 million (95% CI=3.40–7.98) in 2015, PACG from 3.22 million (95% CI=2.70–3.84) to 7.14 million (95% CI=5.97–8.53), and secondary glaucoma from 0.34 million (95% CI=0.23–0.53) to 0.76 million (95% CI=0.51–1.17). In 2015, more than half (54.42%) of the glaucoma cases were PACG, followed by POAG (39.79%) and secondary glaucoma (5.79%). By 2050, the number of all glaucoma cases in China will be 25.16 million (95% CI=18.96–33.86). In the multivariable meta-regressions, the odds ratio (OR) for each decade's increase in age was 1.43 (95% CI=1.33–1.55) for POAG, and 1.65 (95% CI=1.51–1.80) for PACG; males were more likely to have POAG (OR 1.36, 95% CI=1.17–1.59), but less likely to have PACG (OR 0.53, 95% CI=0.46–0.60) compared with females. After adjustment of age and gender, people living in urban areas were more likely to have POAG compared with those in rural areas (OR 1.54, 95% CI=1.02–2.35). People in Northeast China were at a higher risk (OR 1.77, 95% CI=1.07–2.94) of having PACG than people in East China. Among the six regions, East China owed the most POAG and PACG cases, whereas Northwest China owed the least.

**Conclusions** This systematic review and meta-analysis suggests a substantial burden of glaucoma in China, with great variances among the different age groups, genders, settings and geographic regions. With the dramatic ageing trend in the next three decades, the prevalence and burden of glaucoma will continue to increase. More elaborate epidemiological studies are needed to optimise public health strategies for mitigating this important health problem.

## Correspondence to:

Peige Song  
Centre for Global Health Research  
Usher Institute of Population Health Sciences and Informatics  
University of Edinburgh  
Edinburgh EH8 9AG  
Scotland, UK  
p.song@sms.ed.ac.uk



Glaucoma, the second leading cause of blindness, is an optic neuropathy characterised by progressive structural and functional changes of the optic nerve, leading to a typical appearance of the optic disc and visual field damage if untreated [1–5]. People with glaucoma-induced visual impairment generally suffer from decreased vision-related quality of life (including reduced vision-dependent mobility, increased incidence of falls), and place a huge burden on caregivers and communities [6–9]. Glaucoma is often associated with a long and asymptomatic initial phase, and is usually unnoticed until its later stages, when extensive and irreversible damage has occurred [10,11]. In the late stage of the disease, the effects of medical and surgical treatment can be unsatisfactory, underscoring the importance of early detection and treatment [1,2,12]. As glaucoma has an uncertain prognosis, it requires lifelong management and follow-up to prevent further loss of vision [1,13,14]. The recognition of glaucoma's pervasive nature and adverse impact on both individuals and society, and the documentation of the magnitude and distribution of glaucoma is of pronounced importance to inform clinicians and researchers, and will guide policymakers in health services allocation [9,15,16].

Globally, 64.3 million individuals, or 3.5% of the world's population, have glaucoma; of these, about 5.7 million people are visually impaired and 3.1 million are blind [5,16]. Of the many subtypes of glaucoma, primary open-angle glaucoma (POAG) is the most common in nearly all regions, accounting for more than two-thirds (68.6%) of all glaucoma cases [15,16]. Geographically, POAG is believed to be particularly prevalent in Africa (4.2%), and least prevalent in Asia (2.3%); however, more than half (53.4%) of the global POAG cases are in Asia due to the relatively large population size of this region [16,17]. Compared with POAG, primary angle-closure glaucoma (PACG) is a less common subtype, but is more visually damaging [18]. PACG also disproportionately affects the global population; it is least common in North America (0.3%), but is the predominant type of glaucoma in Asian populations (1.1%) [16]. More than three quarters (76.7%) of the global PACG cases are in Asia [16,17]. Given the positive association of glaucoma prevalence and advanced age, glaucoma is expected to become an even larger public health concern in the coming decades [18–20]. This dramatic increase of glaucoma burden is also expected to be the case for the largest developing country – China – where rapid ageing of the population is under way [21–23].

The last three decades have seen a proliferation of population-based studies in China. The mounting volume of data on the prevalence of glaucoma in Chinese bibliographical databases allows us to explore the burden of glaucoma in China from a modelling approach [24–26]. However, epidemiological studies on glaucoma to date have been restricted to specific demographic and geographic features in China, and are therefore not generalisable to the overall Chinese population [27–30]. Because of the uncertainty and variation surrounding the epidemiological surveys on glaucoma, a systematic synthesis of the prevalence of glaucoma in China is particularly needed. The first study to pool the prevalence of primary glaucoma in China was published in 2013, which revealed an overall prevalence of 0.7% for POAG, and 1.4% for PACG [30]. The meta-analysis was based on 14 articles from 12 population-based studies published before 2009. No study has yet systematically appraised research into the prevalence of primary glaucoma in China published over the last nine years. Furthermore, considering that different subtypes of glaucoma may require different strategies for screening, prevention and treatment, there is a pressing need for a more updated effort to provide finer quantification of the relative magnitude across the main types of glaucoma, namely, POAG, PACG and secondary glaucoma, which is missing in the 2013 study [17].

To fill the gaps in the evidence matrix, in this study we used a comprehensive systematic review to synthesise the best available evidence from 1990 onwards. Based on the retained evidence, we assessed the prevalence and burden of glaucoma and its subtypes at both the national and subnational levels. The aims of this present study were: 1) to estimate glaucoma prevalence in China by using epidemiological modelling; 2) to estimate and project the overall prevalence and number of people with glaucoma at the national level from 1990 to 2050; and 3) to estimate the regional prevalence and number of people with glaucoma from 2000 to 2010.

## METHODS

### Systematic review

Our comprehensive systematic review was conducted and reported in line with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines and the Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER) statement [31,32].

### **Search strategy**

A systematic literature search was performed to identify all relevant articles that have reported the prevalence of glaucoma in the general Chinese population. The searched databases included three Chinese and three English electronic databases: China National Knowledge Infrastructure (CNKI), Wanfang, Chinese Biomedicine Literature Database (CBM–SinoMed), PubMed, Embase, and Medline. The search strategy combined controlled vocabularies (eg, Medical Subject Heading terms) and free text terms of prevalence (prevalence, incidence, mortality, morbidity, epidemiology), glaucoma and China (China, Chinese, Hong Kong, Macau, Taiwan); the specific search strategies for each database were adapted to fit their specific features (Table S1 in **Online Supplementary Document**). We restricted our searches to studies that were published between January 1990 and August 2017. No language restrictions were applied to the searches or search results. The reference lists of all included full-text articles were also scrutinised in detail to identify additional data sources.

### **Inclusion and exclusion criteria**

The inclusion and exclusion criteria adopted in this study were developed based on the examination guidelines for glaucoma-related population-based studies [3,16,33]. To be included in this systematic review, studies had to be population-based and report the prevalence of glaucoma. We excluded studies that were hospital-based or conducted in a population that was not representative of the general population. Reviews, commentaries, studies that only adopted qualitative methods and studies that reported the number of eyes with glaucoma instead of the number of individuals were also excluded because they were not able to provide numerical estimates of glaucoma prevalence. Studies that did not include clear assessment methods of glaucoma or relied on self-reported diagnosis were also excluded. Although different case definitions and examination methods exist in identifying glaucoma cases, a remarkable similarity of glaucoma prevalence was noted across surveys despite variations in survey methodology and glaucoma definition [16,34–36]. In this systematic review, we did not exclude studies on the basis of their specific definitions of glaucoma or adopted instrumentation; the assessment of glaucoma should be independent of intraocular pressure (IOP) measurements, but rely on structural or functional evidence of glaucomatous optic neuropathy evaluated by optic disc evaluation or visual field testing [3,16]. Therefore, studies were eligible to contribute data if the following standardised assessments were carried out in suspected cases of glaucoma: anterior chamber angle/depth evaluation by slit-lamp examination or gonioscopy, optic disc evaluation by ophthalmologists using slit-lamp biomicroscopy or fundus photography and visual field testing with automated static perimetry.

### **Study selection and data extraction**

Search results from the six bibliographic databases were merged together and duplicate references were removed within and between the databases. All records were independently screened by two researchers (PS and JW) in two stages: screening of titles and abstracts, followed by the retrieval and screening of full-text articles. For multiple articles that reported results of the same individual study, those with the most comprehensive or most recent data were kept. Disagreements were resolved by consensus through discussion.

For the purpose of this study, glaucoma was classified into three main types: POAG, PACG and secondary glaucoma. Relevant data on different subtypes of glaucoma were separately extracted from the studies included. The pilot tested and refined extraction table included three modules:

- 1) Characteristics of the study: author(s), publication year, study setting (urban, rural or mixed), study location, geographic region, survey year, sampling method, study design (cross-sectional or cohort), whether anterior chamber angle/depth evaluation, IOP measurement, optic disc evaluation and visual field testing were conducted;
- 2) Characteristics of the investigated population: number of the sample, gender (male, female or mixed), and age (age range, mean or median age, or midpoint of the age range);
- 3) Prevalence estimates: the number of participants who had been tested and the number of people with glaucoma, by age group, gender, setting and glaucoma subtype, where available.

We classified the sites where the studies were conducted into six geographic regions following definitions of National Bureau of Statistics of China: East China, North China, Northeast China, Northwest China, South Central China, and Southwest China (**Table 1**) [37,38]. Missing data on survey years were imputed for two studies by subtracting three years from the published year, which was based on the average

**Table 1.** The six geographic regions in China

REGION	INCLUDED PROVINCES
North China	Beijing Municipality, Hebei province, Inner Mongolia Autonomous Region, Shanxi province, Tianjin Municipality;
Northeast China	Heilongjiang province, Jilin province, Liaoning province;
East China	Anhui province, Fujian province, Jiangsu province, Jiangxi province, Shandong province, Shanghai Municipality, Zhejiang province;
South Central China	Guangdong province, Guangxi Zhuang Autonomous Region, Hainan province, Henan province, Hubei province, Hunan province;
Southwest China	Chongqing Municipality, Guizhou province, Sichuan province, Tibet Autonomous Region, Yunnan province;
Northwest China	Gansu province, Ningxia Hui Autonomous Region, Qinghai province, Shaanxi province, Xinjiang Uyghur Autonomous Region;

time from survey to publication in studies with available data. In case of censoring age groups, eg, older than 80 years, the same width as other age groups in the same study was used to impute the missing age band.

## Statistical analysis

### Epidemiological modelling of glaucoma prevalence

Prevalence of POAG, PACG and secondary glaucoma was stabilised by using the logit transformation [39]. In this study, random-effects models were used throughout because of significant heterogeneity in the reported prevalence of POAG, PACG and secondary glaucoma between studies (Table S2 in **Online Supplementary Document**). The overall prevalence of secondary glaucoma was derived from the study-specific estimates using a random-effects meta-analysis model (DerSimonian and Laird method) [40]. For POAG and PACG, one individual study might have contributed multiple outcome measurements in the data extraction stage; to take into account the occurrence of different data points from the same study, a multilevel mixed-effect logistic regression approach was adopted [41,42]. Before constructing epidemiological models of the prevalence for POAG and PACG, the association of prevalence estimates and each individual variable, ie, age, gender (male and female), setting (urban, rural and mixed), geographic region, and survey year, was explored using a univariable meta-regression; this was done for POAG and PACG separately. Age and gender were found to be the only common factors that were significantly associated with prevalence estimates of both POAG and PACG. For the purpose of estimating the national prevalence of POAG and PACG, an age- and gender-adjusted model was developed. Given that:

$$\text{prevalence} = p = \frac{\text{glaucoma cases}}{\text{number of participants}}$$

Then, the prevalence estimates were stabilised by the logit link,

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \ln(\text{odds}) = \alpha + \beta_1 * x_1 + \beta_2 * x_2 + \dots$$

The prevalence of glaucoma was established as a function of age and gender:

$$\text{logit}(p) = \alpha + \beta_1 * \text{Age} + \beta_2 * \text{Gender}$$

Therefore,

$$\text{odds} = \frac{p}{1-p} = e^{(\alpha + \beta_1 * \text{Age} + \beta_2 * \text{Gender})}$$

And,

$$\text{prevalence} = p = \frac{e^{(\alpha + \beta_1 * \text{Age} + \beta_2 * \text{Gender})}}{1 + e^{(\alpha + \beta_1 * \text{Age} + \beta_2 * \text{Gender})}}$$

Finally, the age- and gender-specific prevalence of POAG and PACG was generated based on the above-mentioned model. The lower bound of age range was set as 45 years and the upper bound as 89 years because enough data were available for model construction in this broad age range.

### Estimation of national population with glaucoma from 1990 to 2015

The national number of people with glaucoma (glaucoma “envelopes”) from 1990 to 2015 was derived by multiplying the prevalence of glaucoma with the population in China, available from the United Na-

tions Population Division (UNPD) [43]. For POAG and PACG, the numbers of age- and gender-specific cases were calculated by using their age- and gender-specific prevalence for each 5-year age group estimated in the above models. In view of the limited data availability, the age- and gender-specific prevalence of secondary glaucoma was not estimated, thus the overall prevalence pooled from 12 studies were used (Figure S3 in **Online Supplementary Document**). This was performed for the years 1990, 2000, 2010 and 2015 consecutively.

### ***Projection of national population with glaucoma from 2020 to 2050***

For our projection of people with glaucoma to the year 2050, the prevalence of POAG, PACG and secondary glaucoma was assumed to be constant over the next 33 years. This assumption was partly supported by our multivariable meta-regression model, where no significant changes of POAG and PACG prevalence with survey year were observed after adjusting the effects of age and gender. Bases on the same procedures adopted in the estimation of national population with glaucoma from 1990 to 2015, the numbers of people with glaucoma from 2020 to 2050 were projected by taking UNDP Prospects data, which took into account mortality rates and fertility rates in its population projection [43].

### ***Effects of demographic and geographic factors on the prevalence of POAG and PACG***

To investigate whether the prevalence of POAG and PACG varied across different subgroups of the population, the associations of prevalence estimates and variables of interest were assessed by multivariable meta-regression, adjusting the effects of age and gender. Before model fitting, all variables were tested for correlation to avoid multicollinearity. The variables were selected based on our knowledge, previous studies and the availability of data in this present study, which included setting (urban, rural and mixed), geographic region, and survey year. As a rule, at least seven data points should be available for each variable [44].

### ***Estimation of regional population with POAG and PACG from 2000 to 2010***

The regional population with glaucoma was estimated at an envelope condition. This method was initially proposed by the Child Health Epidemiology Reference Group (CHERG) and has been widely adopted in disease burden research [45–47]. The national glaucoma cases were set as the glaucoma envelope, for POAG and PACG separately. Then the “POAG envelope” and the “PACG envelope” were split into the six subnational regions according to the different distributions of risk factors identified in the multivariable meta-regression models. This was conducted for the years 2000 and 2010, where regional population data were available from the fifth and sixth censuses of China [37,38].

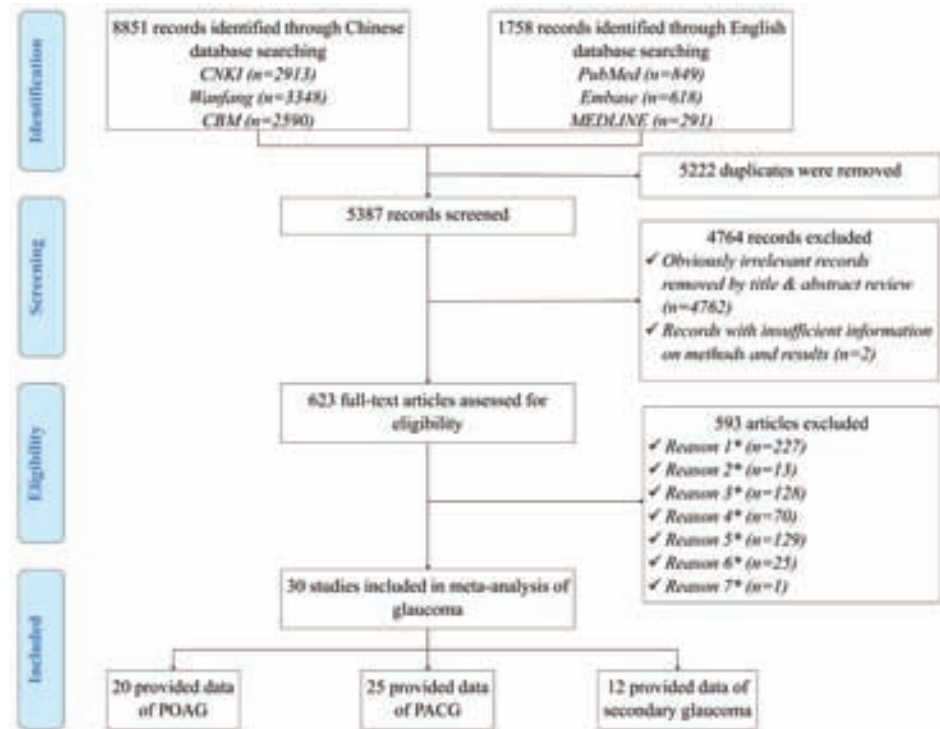
All analyses were performed using R, version 3.3.0 (R Foundation for Statistical Computing, Vienna, Austria). The China base map was obtained as a shapefile from the Global Administrative Areas (GADM) database (GADM, 2015, version 2.0; www.gadm.org) and all maps were drawn by ArcMap version 10.1 (Environmental Systems Research Institute, Redlands, CA). A two-sided *p*-value of less than 0.05 indicated statistically significant difference for all analyses.

## **RESULTS**

### **Summary of systematic review**

Our initial search identified a total of 10 609 citations for screening, after elimination of duplicates, 5387 records remained. After screening titles and abstracts, 623 potentially relevant full-text articles were reviewed for eligibility, of which 30 reported the prevalence of glaucoma and were included in the systematic review (**Figure 1**). A full list of the included studies is shown in Table S3 in **Online Supplementary Document**.

The 30 studies, published between 1995 and 2016, reported the prevalence of glaucoma with a geographical distribution covering all the six regions in China (**Figure 2**). The included studies were all cross-sectional, of which 20 studies reported the prevalence of POAG, 25 focused on PACG and 12 on secondary glaucoma. The detailed characteristics of every study are listed in Table S4 in **Online Supplementary Document**, and the main characteristics of the 30 studies are summarised in **Table 2**. More than half of the 30 studies were published in the past seven years, underlining the necessity for conducting our revision of the estimate. The included studies were generally larger, with the majority (60%, *n* = 18) being conducted in rural areas. Anterior chamber angle/depth evaluation, IOP measurement and



**Figure 1.** Systematic review flow diagram of studies on glaucoma prevalence in China. Note: \*Reason 1—Studies that were not population-based; \*Reason 2—Studies that were not based in China; \*Reason 3—Papers with no numerical prevalence measure of glaucoma; \*Reason 4—Studies that relied on self-reported diagnoses or didn't conduct standardised assessments (anterior chamber angle/depth evaluation by slit-lamp examination or gonioscopy, optic disc evaluation by ophthalmologists using slit-lamp biomicroscopy or fundus photography and visual field testing with automated static perimetry) in at least glaucoma suspects; \*Reason 5—Studies that were conducted in a population with unrepresentative characteristics (diabetic patients, people with reduced vision, etc.); \*Reason 6—Multiple publications of the same study; \*Reason 7—Papers with inconsistency between reported methods and presented results.



**Figure 2.** Geographical distribution of the included studies on glaucoma prevalence in China (n=30).

**Table 2.** Main characteristics of the included studies on glaucoma prevalence in China (n = 30)

CHARACTERISTICS OF STUDY	NUMBER OF STUDIES (%)
<b>Year published:</b>	
1990–1999	2 (6.7)
2000–2009	12 (40.0)
2010–2017	16 (53.3)
<b>Setting:</b>	
Urban	6 (20.0)
Rural	18 (60.0)
Mixed	6 (20.0)
<b>Sample size:</b>	
500–1500	6 (20.0)
1501–2500	7 (23.3)
2501–5000	9 (30.0)
>5000	8 (26.7)
<b>Geographic regions:</b>	
North China	12 (40.0)
Northeast China	5 (16.7)
East China	4 (13.3)
South Central China	4 (13.3)
Southwest China	3 (10.0)
Northwest China	2 (6.7)
<b>Anterior chamber angle/depth evaluation:</b>	
In all participants	28 (93.3)
In glaucoma suspects	2 (6.7)
<b>IOP measurement:</b>	
In all participants	29 (96.7)
In glaucoma suspects	1 (3.3)
<b>Optic disc evaluation:</b>	
In all participants	28 (93.3)
In glaucoma suspects	2 (6.7)
<b>Visual field testing:</b>	
In all participants	4 (13.3)
In glaucoma suspects	26 (86.7)

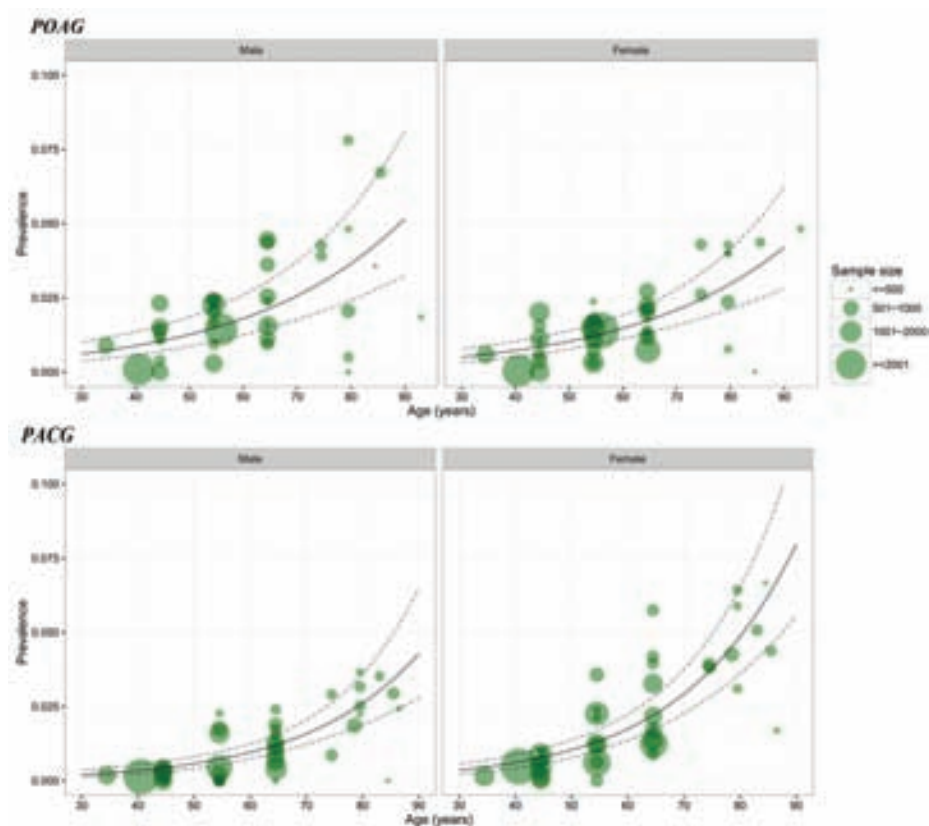
optic disc evaluation were mostly undertaken in all participants, whereas visual field testing was largely used in glaucoma suspects.

### Age- and gender-specific prevalence of POAG and PACG

Based on a substantial number of data points from the included studies, the gender-specific relationship between age and the prevalence of POAG and PACG was constructed (Figure 3). The informative data points covered a wide age spectrum from the mid-30s to the 9th decade, with the majority concentrating between the mid-40s to the mid-80s. Therefore, in the estimation of age- and gender-specific prevalence of POAG and PACG, the age range was set as from 45 years to 89 years.

Table 3 and Figure 4 show the estimated age- and gender-specific prevalence of POAG and PACG respectively. Generally, the prevalence of POAG and PACG both increased steadily with advanced age, and this positive relationship between age and prevalence rate was similar between genders, but more pronounced for PACG. In males, the prevalence of POAG ranged from 0.74% (95% CI = 0.48–1.14) in individuals aged 45–49 years to 3.02% (95% CI = 1.92–4.73) in those aged 85–89 years. The prevalence of POAG in females was slightly lower than that in males across the whole age spectrum from 45 to 89 years, ranging from 0.54% (95% CI = 0.35–0.84) to 2.24% (95% CI = 1.41–3.53). In contrast, the prevalence of PACG was consistently higher in females than in males. In females, the prevalence of PACG ranged from 0.91% (95% CI = 0.74–1.11) in those aged 45–49 years to 6.33% (95% CI = 4.98–8.02) in those aged 85–89 years. In males, the prevalence of PACG increased from 0.48% (95% CI = 0.39–0.60) in people aged 45–49 years to 3.44% (95% CI = 2.66–4.45) in elderly aged 85–89 years.

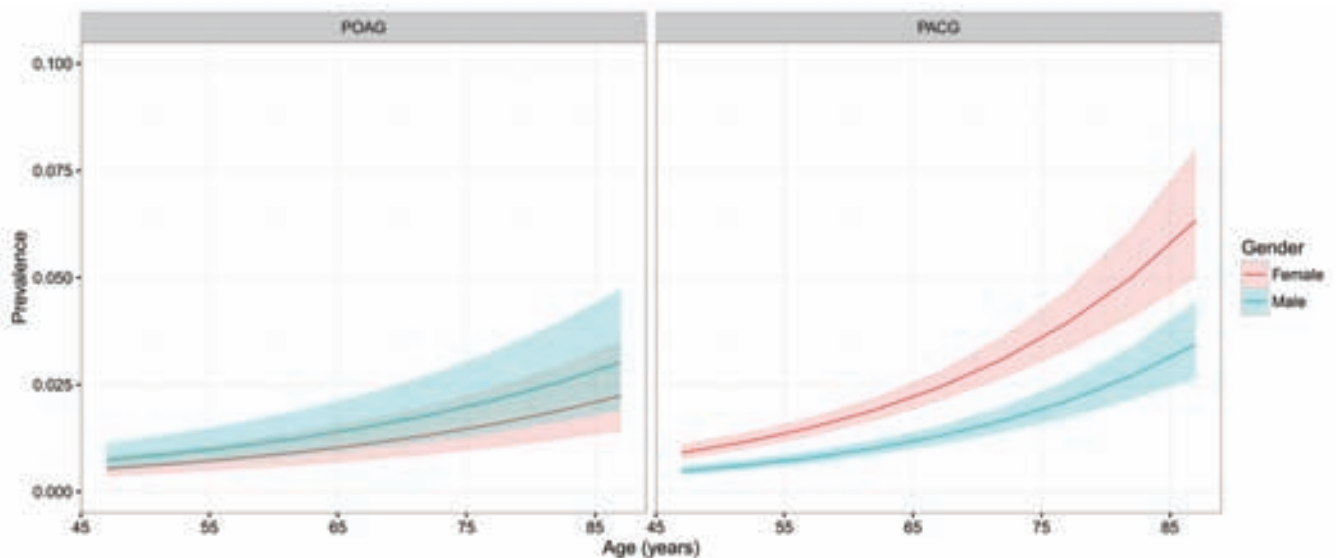
**Figure 3.** Age- and gender-specific prevalence of primary open-angle glaucoma (POAG) and primary angle-closure glaucoma (PACG) based on the informative data points from the included studies. Note: The size of each bubble is proportional to the sample size. Overall, there were 86 data points for constructing the gender-specific relation between age and prevalence for POAG, and 103 for PACG.



**Table 3.** Estimated gender-specific prevalence of POAG and PACG in China, by age group

Age (years)	Prevalence of POAG (%; 95% CI)		Prevalence of PACG (%; 95% CI)	
	Male	Female	Male	Female
45–49	0.74 (0.48–1.14)	0.54 (0.35–0.84)	0.48 (0.39–0.60)	0.91 (0.74–1.11)
50–54	0.88 (0.57–1.34)	0.65 (0.42–0.99)	0.62 (0.51–0.75)	1.16 (0.98–1.38)
55–59	1.05 (0.69–1.60)	0.77 (0.51–1.18)	0.79 (0.66–0.94)	1.49 (1.28–1.74)
60–64	1.25 (0.83–1.90)	0.92 (0.61–1.41)	1.01 (0.86–1.20)	1.90 (1.65–2.20)
65–69	1.50 (0.98–2.27)	1.10 (0.72–1.68)	1.30 (1.09–1.54)	2.43 (2.10–2.81)
70–74	1.79 (1.17–2.72)	1.32 (0.86–2.02)	1.66 (1.38–1.99)	3.10 (2.64–3.64)
75–79	2.13 (1.38–3.27)	1.57 (1.02–2.43)	2.12 (1.73–2.59)	3.94 (3.28–4.73)
80–84	2.54 (1.63–3.93)	1.87 (1.20–2.92)	2.70 (2.15–3.39)	5.00 (4.06–6.16)
85–89	3.02 (1.92–4.73)	2.24 (1.41–3.53)	3.44 (2.66–4.45)	6.33 (4.98–8.02)

POAG – primary open-angle glaucoma, PACG – primary angle-closure glaucoma, CI – confidence interval

**Figure 4.** Estimated age- and gender-specific prevalence of primary open-angle glaucoma (POAG) and primary angle-closure glaucoma (PACG) in China, with 95% confidence intervals.

### National population with glaucoma from 1990 to 2015

By extrapolating the estimated age- and gender-specific prevalence of POAG and PACG to UNPD data, the numbers of people with POAG and PACG were generated (Table S5 in **Online Supplementary Document**). At the national level, the overall prevalence of glaucoma was listed in **Table 4**. From 1990 to 2015, the prevalence of all glaucoma ranged from 2.59% (95% CI = 1.96–3.49) to 2.58% (95% CI = 1.94–3.47), indicating a slightly relative decreasing rate of 0.39%. For different subtypes of glaucoma, the overall prevalence of POAG ranged from 1.03% (95% CI = 0.67–1.58) in 1990 to 1.02% (95% CI = 0.67–1.57) in 2015, which yielded a relative decreasing rate of 0.97%. Similarly, the prevalence of PACG decreased by 0.71%, from 1.41% (95% CI = 1.18–1.68) in 1990 to 1.40% (95% CI = 1.17–1.68) in 2015. Based on

**Table 4.** Estimated prevalence and number of people with glaucoma in China from 1990 to 2015, by glaucoma type

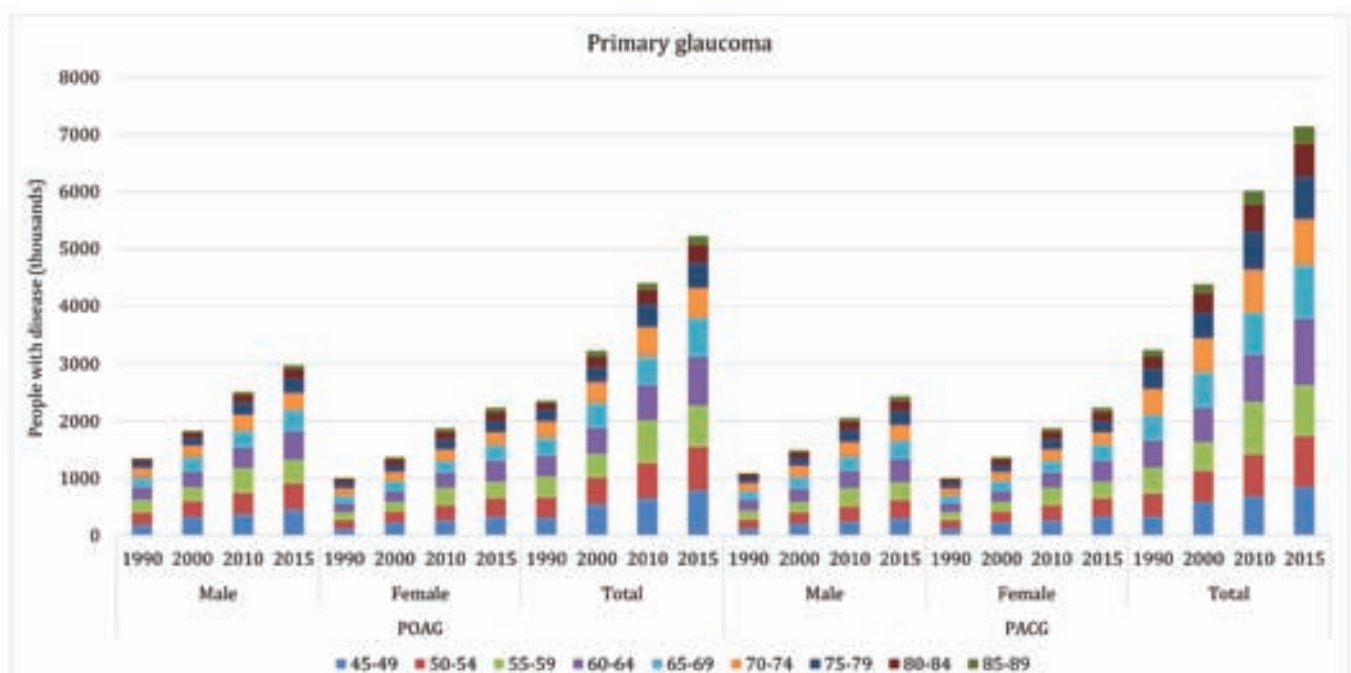
GLAUCOMA TYPE	PREVALENCE OF GLAUCOMA (%, 95% CI)		NUMBER OF PEOPLE WITH GLAUCOMA (MILLION, 95% CI)		RELATIVE RATE OF CHANGE (%, 1990–2015)	
	1990	2015	1990	2015	Prevalence	Cases
POAG	1.03 (0.67–1.58)	1.02 (0.67–1.57)	2.35 (1.54–3.60)	5.22 (3.40–7.98)	–0.97	122.13
PACG	1.41 (1.18–1.68)	1.40 (1.17–1.68)	3.22 (2.70–3.84)	7.14 (5.97–8.53)	–0.71	121.74
Secondary glaucoma	0.15 (0.10–0.23)		0.34 (0.23–0.53)		–	123.53
All glaucoma*	2.59 (1.96–3.49)	2.58 (1.94–3.47)	5.92 (4.47–7.97)	13.12 (9.88–17.68)	–0.39	121.62

POAG – primary open-angle glaucoma, PACG – primary angle-closure glaucoma, CI – confidence interval

\*All glaucoma includes POAG, PACG and secondary glaucoma.

12 individual studies, the overall prevalence of secondary glaucoma was pooled as 0.15% (95% CI=0.10–0.23) and assumed as constant over the time frame of this research (Figure S3 in **Online Supplementary Document**).

With the ageing of Chinese population during 1990–2015, the total number of people living with glaucoma increased dramatically (**Table 4**). The number of people with all glaucoma in China was 5.92 million (95% CI=4.47–7.97) in 1990 and 13.12 million (95% CI=9.88–17.68) in 2015, indicating an overall increasing rate of 121.62% throughout this period. This increasing trend was also witnessed in different subtypes of glaucoma. For POAG, the affected cases increased by 122.13%, from 2.35 million (95% CI=1.54–3.60) in 1990 to 5.22 million (95% CI=3.40–7.98) in 2015. Similarly, the number of people with PACG increased by 121.74%, ranging from 3.22 million (95% CI=2.70–3.84) in 1990 to 7.14 million (95% CI=5.97–8.53) in 2015. Even for secondary glaucoma, whose overall prevalence was assumed as constant, the number of affected people also increased by 123.53% during 1990 to 2015, from 0.34 million (95% CI=0.23–0.53) to 0.76 million (95% CI=0.51–1.17). In 2015, more than half (54.42%) of the glaucoma cases were PACG, followed by POAG (39.79%) and secondary glaucoma (5.79%). From 1990 to 2015, the age group that contributed the most cases shifted from 55–59 years to 60–64 years for both POAG and PACG (**Figure 5**).



**Figure 5.** Estimated gender-specific number of people with primary open-angle glaucoma (POAG) and primary angle-closure glaucoma (PACG) in China from 1990 to 2015, with contributing age groups.

## Projection of national population with glaucoma from 2020 to 2050

In our projection analysis, the age- and gender-specific prevalence estimates of POAG and PACG, and the overall prevalence estimate of secondary glaucoma were all assumed as constant. By applying these estimates to the UNPD data up to the year 2050, the numbers of people with glaucoma were projected (Table S5 in **Online Supplementary Document** and **Table 5**). Unlike the slight declining trend of glaucoma prevalence between 1990 and 2015, from 2020 to 2050, the overall prevalence of all glaucoma is expected to increase from 2.64% (95% CI=1.99–3.55) to 3.48% (95% CI=2.63–4.69), which is a 32% increase. For different subtypes of glaucoma, the prevalence of POAG will also increase during this period, but at a lower rate (27%). In 2020, the prevalence of POAG is projected to be 1.05% (95% CI=0.68–1.60), and then reach 1.33% (95% CI=0.86–2.04) in 2050. The prevalence of PACG will show a greater increasing rate between 2020 and 2050, from 1.44% (95% CI=1.21–1.72) to 2.01% (95% CI=1.66–2.42), ie, by 40%.

The projected number of people living with glaucoma in China is also shown in **Table 5**. Between 2020 and 2050, the number of all glaucoma cases in China is expected to increase from 15.28 million (95% CI=11.53–20.58) to 25.16 million (95% CI=18.96–33.86), ie, by 65%. The increasing rates for POAG and PACG will also be notable within the same period. The number of people with POAG is expected to increase from 6.06 million (95% CI=3.95–9.27) to 9.59 million (95% CI=6.23–14.72), and those with PACG from 8.36 million (95% CI=7.00–9.98) to 14.49 million (95% CI=12.01–17.48), which will translate into the rates of increase between 2020 and 2050 of 58% and 73%, respectively. Due to the forecasted trend in population ageing over the next three decades, the number of secondary glaucoma cases is anticipated to also increase slightly, from 0.87 million (95% CI=0.58–1.33) in 2020 to 1.08 million (95% CI=0.72–1.66) in 2050; ie, an increase by 24.14%. During this period, PACG will remain the predominant subtype of glaucoma in China, followed by POAG and secondary glaucoma. From 2020 to 2050, the age group in which most POAG cases will be concentrated will shift from 65–69 years to 75–79 years, and the age group for PACG will shift from 65–69 years to 80–84 years (**Figure 6**).

## Effects of demographic and geographic factors on the prevalence of POAG and PACG

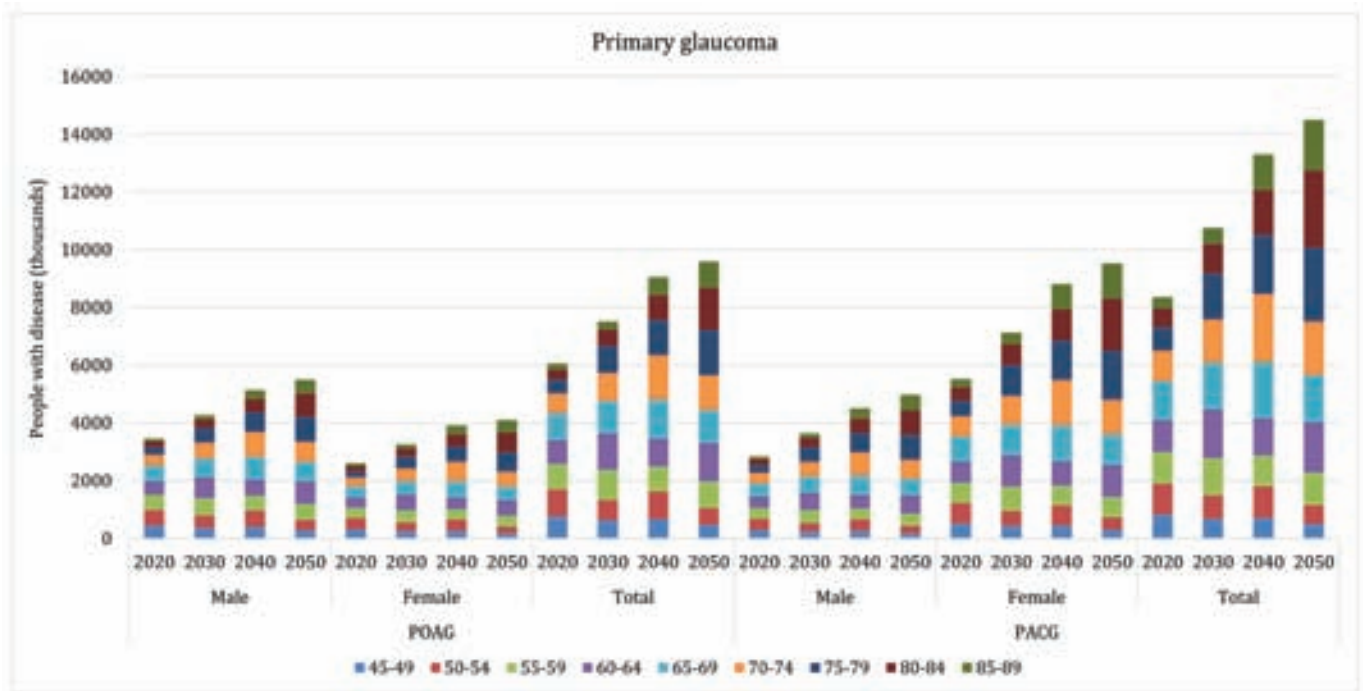
Based on univariable and multivariable meta-regression models (**Table 6**), the effects of selected demographic and geographic factors on the risk of POAG and PACG were assessed. The univariable meta-regression indicated that age, gender and study setting were significantly associated with POAG, while PACG was associated with age, gender and geographic region. No evidence for secular trends was observed. After adjusting for age and gender in a multivariable meta-regression, the odds ratio (OR) for each increase in age by 10 years was 1.43 (95% CI=1.33–1.55) for POAG, and 1.65 (95% CI=1.51–1.80) for PACG. Males still showed a higher risk of POAG (1.36, 95% CI=1.17–1.59), but a lower risk of PACG (0.53, 95% CI=0.46–0.60) in comparison with females. People living in urban areas were more likely to have POAG compared to those in rural areas, with an OR of 1.54 (95% CI=1.02–2.35). Among the six geographic regions, people in Northeast China were at a higher risk (1.77, 95% CI=1.07–2.94) of having PACG than people in East China.

**Table 5.** Projected prevalence and number of people with glaucoma in China from 2020 to 2050, by glaucoma type

GLAUCOMA TYPE	PREVALENCE OF GLAUCOMA (%; 95% CI)		NUMBER OF PEOPLE WITH GLAUCOMA (MILLION; 95% CI)		RELATIVE RATE OF CHANGE (%; 2020–2050)	
	2020	2050	2020	2050	Prevalence	Cases
POAG	1.05 (0.68–1.60)	1.33 (0.86–2.04)	6.06 (3.95–9.27)	9.59 (6.23–14.72)	26.67	58.25
PACG	1.44 (1.21–1.72)	2.01 (1.66–2.42)	8.36 (7.00–9.98)	14.49 (12.01–17.48)	39.58	73.33
Secondary glaucoma	0.15 (0.10–0.23)	0.87 (0.58–1.33)	0.87 (0.58–1.33)	1.08 (0.72–1.66)	–	24.14
All glaucoma*	2.64 (1.99–3.55)	3.48 (2.63–4.69)	15.28 (11.53–20.58)	25.16 (18.96–33.86)	31.82	64.66

POAG – primary open-angle glaucoma, PACG – primary angle-closure glaucoma, CI – confidence interval

\*All glaucoma includes both primary glaucoma and secondary glaucoma.



**Figure 6.** Projected gender-specific number of people with primary open-angle glaucoma (POAG) and primary angle-closure glaucoma (PACG) in China from 2020 to 2050, with contributing age groups.

**Table 6.** Odds ratios for POAG and PACG in terms of age, gender, setting and geographic region from multilevel univariable and multivariable meta-regression models, with 95% confidence intervals

VARIABLE	UNADJUSTED		AGE AND GENDER-ADJUSTED	
	POAG	PACG	POAG	PACG
Age (per decade)	1.38 (1.30–1.47)*	1.58 (1.49–1.67)*	1.43 (1.33–1.55)*	1.65 (1.51–1.80)*
<b>Gender: †</b>				
Female	Reference	Reference	Reference	Reference
Male	1.39 (1.19–1.62)*	0.53 (0.46–0.60)*	1.36 (1.17–1.59)*	0.53 (0.46–0.60)*
<b>Setting:</b>				
Rural	Reference	Reference	Reference	Reference
Urban	1.68 (1.13–2.51)*	0.90 (0.64–1.28)	1.54 (1.02–2.35)*	0.82 (0.54–1.23)
Mixed	0.93 (0.37–2.33)	1.13 (0.67–1.90)	2.18 (0.68–6.99)	0.64 (0.25–1.62)
<b>Geographic region:</b>				
East	Reference	Reference	Reference	Reference
North	1.44 (0.45–4.61)	1.81 (0.93–3.52)	1.29 (0.36–4.69)	1.23 (0.77–1.97)
Northeast	1.48 (0.31–7.02)	2.87 (1.37–6.01)*	1.41 (0.26–7.74)	1.77 (1.07–2.94)*
Northwest	0.80 (0.10–6.33)	1.87 (0.79–4.41)	0.44 (0.05–4.07)	1.38 (0.76–2.49)
South Central	2.14 (0.45–10.14)	1.54 (0.72–3.28)	–	0.89 (0.51–1.53)
Southwest	1.48 (0.37–6.00)	1.51 (0.69–3.29)	0.94 (0.17–5.16)	1.24 (0.70–2.22)
<b>Investigation year (per decade) ‡</b>	1.07 (0.58–1.98)	0.76 (0.50–1.15)	1.47 (0.51–4.25)	1.18 (0.92–1.50)

POAG – primary open-angle glaucoma, PACG – primary angle-closure glaucoma, CI – confidence interval

\*Statistically significant.

†The effect of gender was estimated based on studies that reported gender-specific glaucoma prevalence.

‡The secular trend was evaluated based on studies that were conducted after the year 2000.

### Regional population with POAG and PACG from 2000 to 2010

By taking the effects of age, gender and setting, the national POAG cases were distributed to the six geographic regions in China (Table 7). In 2000, the overall prevalence of POAG was 1.01% (95% CI=0.66–1.55) in China, ranging from 0.96% (95% CI=0.60–1.53) in Southwest China to 1.08% (95% CI=0.75–1.54) in Northeast China. In 2010, the overall prevalence of POAG in China rose to 1.03% (95% CI=0.67–1.57), and the regions with the highest prevalence of POAG were Northeast China (1.05%,

**Table 7.** Estimated prevalence and number of people with POAG in China from 2000 to 2010, by geographic region

REGION	PREVALENCE OF POAG (% , 95% CI)		NUMBER OF PEOPLE WITH POAG (MILLION, 95% CI)		RELATIVE RATE OF CHANGE (% , 2000–2010)	
	2000	2010	2000	2010	Prevalence	Cases
North China	1.01 (0.67–1.53)	1.02 (0.68–1.54)	0.38 (0.25–0.57)	0.55 (0.37–0.83)	1.00	46.13
Northeast China	1.08 (0.75–1.54)	1.05 (0.71–1.54)	0.30 (0.21–0.43)	0.44 (0.30–0.64)	-2.47	45.03
East China	1.04 (0.69–1.58)	1.05 (0.69–1.59)	1.02 (0.67–1.54)	1.39 (0.91–2.10)	1.00	36.34
South Central China	1.01 (0.65–1.56)	1.01 (0.65–1.57)	0.83 (0.53–1.29)	1.13 (0.73–1.75)	0.72	35.65
Southwest China	0.96 (0.60–1.53)	1.00 (0.63–1.59)	0.48 (0.30–0.77)	0.61 (0.38–0.97)	3.68	26.51
Northwest China	0.97 (0.62–1.51)	0.98 (0.62–1.53)	0.19 (0.12–0.30)	0.27 (0.18–0.43)	1.00	41.49
China	1.01 (0.66–1.55)	1.03 (0.67–1.57)	3.21 (2.09–4.91)	4.39 (2.86–6.72)	1.05	36.97

POAG – primary open-angle glaucoma, PACG – primary angle-closure glaucoma, CI – confidence interval

95% CI = 0.71–1.54) and East China (1.05%, 95% CI = 0.69–1.59), and that with the lowest POAG prevalence was Northwest China (0.98%, 95% CI = 0.62–1.53). From 2000 to 2010, the prevalence of POAG has risen in China, with an exception of Northeast China, where the prevalence of POAG decreased by 2.47%. The most marked increasing rate was observed in Southwest China (3.68%). In both 2000 and 2010, the distribution of POAG cases across the six geographic regions was similar, with the most cases in East China and the least in Northwest China. From 2000 to 2010, the greatest increase rate was in North China (46.13%), and the least in Southwest China (26.51%).

The number of people with PACG in China was distributed based on the multivariable meta-regression model of PACG and features of the six geographic regions (Table 8). In 2000, the overall prevalence of PACG in China was 1.38% (95% CI = 1.16–1.65), with the highest regional prevalence estimate in Northeast China (2.04%, 95% CI = 1.82–2.26) and lowest in South Central China (1.12%, 95% CI = 0.95–1.32). In 2010, the national prevalence of PACG rose by 1.30%, reaching to 1.40% (95% CI = 1.17–1.68), and the region with the highest prevalence of PACG was still Northeast China (2.06%, 95% CI = 1.84–2.30)

**Table 8.** Estimated prevalence and number of people with PACG in China from 2000 to 2010, by geographic region

REGION	PREVALENCE OF PACG (% , 95% CI)		NUMBER OF PEOPLE WITH PACG (MILLION, 95% CI)		RELATIVE RATE OF CHANGE (% , 2000–2010)	
	2000	2010	2000	2010	Prevalence	Cases
North China	1.48 (1.42–1.53)	1.47 (1.41–1.52)	0.55 (0.53–0.57)	0.79 (0.76–0.82)	-0.47	44.00
Northeast China	2.04 (1.82–2.26)	2.06 (1.84–2.30)	0.57 (0.51–0.64)	0.86 (0.77–0.96)	1.27	50.59
East China	1.27 (0.99–1.63)	1.27 (0.99–1.63)	1.24 (0.97–1.59)	1.68 (1.30–2.15)	-0.03	34.94
South Central China	1.12 (0.95–1.32)	1.11 (0.94–1.31)	0.93 (0.79–1.09)	1.24 (1.05–1.46)	-0.52	33.99
Southwest China	1.51 (1.22–1.87)	1.59 (1.28–1.97)	0.76 (0.61–0.94)	0.97 (0.78–1.20)	5.06	28.20
Northwest China	1.60 (1.26–2.03)	1.63 (1.27–2.07)	0.32 (0.25–0.41)	0.46 (0.36–0.58)	1.58	42.31
China	1.38 (1.16–1.65)	1.40 (1.17–1.68)	4.37 (3.66–5.23)	6.01 (5.03–7.18)	1.30	37.31

POAG – primary open-angle glaucoma, PACG – primary angle-closure glaucoma, CI – confidence interval

and that with the lowest PACG prevalence was still South Central China (1.11%, 95% CI=0.94–1.31). Within this time frame, the prevalence of PACG decreased in North China, East China and South Central China, but increased in Northeast China, Southwest China and Northwest China. The region with the greatest increasing rate of PACG prevalence was Southwest China (5.06%), and that with the greatest decreasing rate was South Central China (−0.52%). From 2000 to 2010, East China consistently harboured the largest share of PACG cases, while Northwest China has had the smallest share. Overall, the number of people with PACG increased from 4.37 million (95% CI=3.66–5.23) to 6.01 million (95% CI=5.03–7.18), which is a 37% increase over this period. This increase was also witnessed in every region, being the most marked in Northeast China (51%) and the least in Southwest China (28%).

## DISCUSSION

Based on rigorous systematic review of existing evidence on glaucoma prevalence in China, this study offers a comprehensive estimate of the prevalence and burden of glaucoma in China at both national and subnational levels, and compares the relative magnitude of three main subtypes of glaucoma, ie, POAG, PACG and secondary glaucoma, in the general mainland Chinese population. From 1990 to 2015, the prevalence of glaucoma fluctuated at around 2.6%, corresponding to 5.92 million and 13.12 million people with glaucoma in the years 1990 and 2015, respectively. By 2050, the prevalence of glaucoma will rise to 3.48%, which equivalents to a total of 25.16 million affected people. Substantial evidence demonstrated that PACG was the predominant subtype of glaucoma in the general Chinese population, followed by POAG and secondary glaucoma. The geographic variations in the prevalence of POAG and PACG were also assessed, with urban dwellers at a higher risk of developing POAG than rural dwellers, and people living in Northeast China being more prone to PACG than people in East China. Because of the uneven population distribution in China, from 2000 to 2010, East China consistently had the largest share of both POAG and PACG cases, and Northwest China the least.

To the best of our knowledge, this study is the most up-to-date and comprehensive systematic review and meta-analysis to explore and present the national and subnational prevalence and burden of glaucoma in China. The principal strengths of this study are a reasonable coverage of the Chinese population, a comprehensive literature search, and a stringent approach to selecting studies for inclusion. Ultimately, this systematic review was built upon 30 individuals studies, which was more than double the number of studies included in the first systematic review on glaucoma in China [30]. With a wide geographical scope covering all the six geographic regions of China, the included studies ensured a sufficient power to conduct the estimates for both the whole nation and subnational regions. Furthermore, with the aim of limiting between-study heterogeneity due to methodological variations, the assessments of glaucoma in the studies included were based on structural or functional evidence of glaucomatous optic neuropathy, rather than IOP measurements, which in part guaranteed a very good detection ability of early-stage glaucoma. In addition, POAG in this study included persons with IOP at all levels [3,16]. Moreover, the prevalence and burden of secondary glaucoma in China was developed for the first time, which added new evidence to the epidemiology of glaucoma both domestically and globally.

Despite the strengths of this study, there are also multiple limitations. First, given the diversity of studies included in study design, targeted population, methods and settings, a relatively high degree of heterogeneity among studies included was observed. Although the estimates of POAG and PACG prevalence were generated based on meta-regression, by taking the effects of age, gender and geographic factors together, some factors other than chance may also be attributable to the observed variance, but could not be fully controlled. In this study, a key issue was that we didn't choose to exclude studies based on consensus criteria for the definitions and grading systems of glaucoma, but rather relied on the examinations. This is because previous studies suggested a remarkable similarity among surveys with different survey methods and glaucoma definitions [16,34–36]. This approach for defining eligible studies has been widely adopted in previous systematic reviews on glaucoma prevalence, but it might still be influenced by the inherent subjectivity of interpreting ophthalmic images [16,17,34,35]. Second, compared with primary glaucoma, secondary glaucoma has been under-examined in epidemiological studies [16,17,48]. In this study, despite our extensive efforts to identify all the available evidence without language restrictions, the number of eligible studies that provided the estimate of secondary glaucoma prevalence was still not sufficient. Given that there was moderate heterogeneity between studies that reported the prevalence of secondary glaucoma, we acknowledge issues about the appropriateness of roughly reporting an overall prevalence of secondary glaucoma. Third, the projections of glaucoma were only based on the assumption

that the prevalence estimates will be constant, thus, changes in the number of people with glaucoma only reflect changes in demographic features of the next three decades. Although this assumption has been commonly adopted in the projection of disease burden, the power of projection analysis beyond the period of studies conducted is limited [16,17,49]. Forth, only the effects of age, gender, setting and geographic region were assessed in subgroup analyses by using both univariable and multivariable meta-regression; however, other relevant factors that were not obtained from the included studies may have also had a role. In addition, all these factors were aggregate level data, thus hampering the opportunity to explore the differences in effects at the individual level, or interaction between factors [50,51]. Fifth, the estimates of glaucoma prevalence were generated at the regional level at best. Any estimates at the provincial level were not possible, owing to the limited availability of data in each province. Taken collectively, the results presented in this study should be interpreted with caution.

The overall estimated prevalence of POAG in this study was slightly higher than that in the previous systematic review (1.0% vs 0.7%) [30]. The disparity between these two estimates might be explained by the combined effect of the different age and gender structures, and the different geographic features of the participants included in these two systematic reviews. Surprisingly, despite the substantial variation in the studies that were included as the basis for both reports, and further differences in adopted methods of meta-analysis, the prevalence of PACG in these two studies was almost identical – both at the level of 1.4%. This similarity in PACG prevalence supports the current understanding of the magnitude of PACG in China.

The prevalence estimates of POAG and PACG were notably associated with advanced age in both sexes; this strong positive relationship matches the natural history of primary glaucoma, which was described in many previous studies. This association confirms the commonly accepted notion that primary glaucoma is an age-related disease [36,52–56]. With increasing longevity, a striking increase in the prevalence and burden of glaucoma is likely, especially for primary glaucoma.

The distribution patterns of POAG and PACG by gender were opposite, with POAG being the predominant subtype of glaucoma in males, and PACG in females. The female predilection for PACG has been widely acknowledged in previous studies, and could be linked to the aetiology of disease, differences in biological factors and environmental exposures between sexes [30,55–58]. However, the evidence of gender effect on POAG is still conflicting [16,34,35,57]. The findings in the present study disagree with the first systematic review of glaucoma prevalence in China, where male predilection for POAG was not reported [30]. The discrepancy between these two studies might be explained by the inadequate study power to confirm associations. Further studies are still needed to explore the different effects of gender on the development of glaucoma, especially POAG, and for deciding different public health policies.

In view of the general understanding that secondary glaucoma is caused by other ocular or systematic disorders that may lead to an increase in intraocular pressure, rather than a normal degenerative process with ageing, it was expected that no effects of age, gender or geographic factors would be seen [48,53]. Only a pooled overall prevalence was generated for secondary glaucoma, with no separate subgroup analysis. In this study, the prevalence of secondary glaucoma was largely lower than that for East Asia (0.15% vs 0.39%) [17]. However, this relatively lower prevalence of secondary glaucoma, presented in this study, still needs to be confirmed with new data. There is little doubt that secondary glaucoma is less frequent in comparison with primary glaucoma; however, the disease burden caused by secondary glaucoma should never be underestimated or neglected, bearing in mind its visually destructive effects [53,54].

In China, PACG was estimated to be responsible for the largest share of glaucoma, followed by POAG and secondary glaucoma. This finding confirms previous studies, which concluded that Chinese people might be more likely to develop PACG than any other ethnic groups in the world [36,59]. Mechanisms underlying this phenomenon are controversial, but may be associated with the difference of anterior chamber and angle anatomy among races [18,60,61]. In addition, given that PACG is more common in females than males, and females have a longer life expectancy than males, the burden of PACG is considerably concerning [43]. The visual damages of PACG are more severe than of the other main subtypes of glaucoma, presenting an even greater public health challenge with a considerable social and economic impact [18,21].

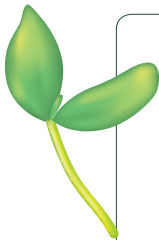
The higher prevalence of POAG in urban than in rural settings is in agreement with findings from studies conducted in China, and also with many other regions across the world [16,17,30]. The reasons for this are not certain, but may be partly related to a higher myopia prevalence in urban areas, and to other potential risk factors for POAG that vary greatly between urban and rural areas. Hypertension, diabetes, diet,

physical activity and air pollution may also play a role [16,17,62,63]. With rapid urbanisation, the prevalence and burden of POAG may continue to increase in China [64,65]. For PACG, people in Northeast China had the highest prevalence. An explanation for this geographic variation is an evolutionary modification of shallower anterior chambers that resist corneal freezing [66,67]. However, these geographic variations might also be a product of differences in the studies included among regions. Future studies should be undertaken to assess geographical risk factors for glaucoma in more detail, to improve locally relevant policy-making on glaucoma.

The findings of this study add insight to our knowledge of the epidemiology of glaucoma and have clear policy implications for China. Together with an ageing demography, glaucoma, especially primary glaucoma, will place an ever-increasing burden on the already stretched health care services in China, unless proactive preventive strategies are put in place. Despite advances in medical treatment, a cost-effective approach for detecting and diagnosing glaucoma is still lacking [68–71]. Indeed, a strong need remains for the development of an appropriate prevention and treatment framework to counter the growing burden of glaucoma in China, especially in rural and poor areas where medical resources are unevenly distributed. National and local efforts are also needed for the formulation of better medical systems and effective public health strategies informed by evidence, such as reallocating medical resources, improving access to health care, and health education on the importance of early examination.

In the meantime, the need to scale up reliable data on glaucoma epidemiology in places where primary data have never been available has been highlighted in the present study. This is essential for both researchers and policymakers to improve understanding of the magnitude and distribution of this problem and the main risk factors. For a comprehensive assessment of glaucoma epidemiology in China, more robust evidence from studies using consistent methods across populations and further reviews of the prevalence of glaucoma is needed to corroborate the statements in the present study more reliably.

In conclusion, this contemporary systematic review and meta-analysis suggests a substantial burden of glaucoma in China, with considerable variation among the different age groups, genders, study settings and geographic regions. PACG is the predominant subtype of glaucoma in the general Chinese population, followed by POAG and secondary glaucoma. In the next three decades, the prevalence and burden of glaucoma will continue to increase with the current ageing trend. More elaborate epidemiological studies are needed to optimise public health strategies for mitigating this important health problem.



**Acknowledgements:** The authors would like to thank the China Scholarship Council for the scholarship to Peige Song.

**Funding:** Peige Song is supported by the China Scholarship Council.

**Authorship declaration:** PS and IR conceptualised and designed the study, PS and JW conducted the systematic review. PS conducted the analysis and prepared the first draft. KB, IR, KC and ET critically reviewed the manuscript. All authors have read and approved the final manuscript as submitted.

**Competing interests:** IR is co-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 organisations. The author completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author). The authors declare no competing interests.

## REFERENCES

- Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: a review. *JAMA*. 2014;311:1901-11. Medline:24825645 doi:10.1001/jama.2014.3192
- Weinreb RN, Khaw PT. Primary open-angle glaucoma. *Lancet*. 2004;363:1711-20. Medline:15158634 doi:10.1016/S0140-6736(04)16257-0
- Foster PJ, Buhrmann R, Quigley HA, Johnson GJ. The definition and classification of glaucoma in prevalence surveys. *Br J Ophthalmol*. 2002;86:238-42. Medline:11815354 doi:10.1136/bjo.86.2.238
- Congdon NG, Friedman DS. Angle-closure glaucoma: impact, etiology, diagnosis, and treatment. *Curr Opin Ophthalmol*. 2003;14:70-3. Medline:12698044 doi:10.1097/00055735-200304000-00002
- Pascolini D, Mariotti SP. Global estimates of visual impairment: 2010. *Br J Ophthalmol*. 2012;96:614-8. Medline:22133988 doi:10.1136/bjophthalmol-2011-300539

- 6 Ramulu P. Glaucoma and disability: which tasks are affected, and at what stage of disease? *Curr Opin Ophthalmol.* 2009;20:92. Medline:19240541 doi:10.1097/ICU.0b013e32832401a9
- 7 Rouland J-F, Berdeaux G, Lafuma A. The economic burden of glaucoma and ocular hypertension. *Drugs Aging.* 2005;22:315-21. Medline:15839720 doi:10.2165/00002512-200522040-00004
- 8 Varma R, Lee PP, Goldberg I, Kotak S. An assessment of the health and economic burdens of glaucoma. *Am J Ophthalmol.* 2011;152:515-22. Medline:21961848 doi:10.1016/j.ajo.2011.06.004
- 9 Traverso CE, Walt J, Kelly S, Hommer A, Bron A, Denis P, et al. Direct costs of glaucoma and severity of the disease: a multinational long term study of resource utilisation in Europe. *Br J Ophthalmol.* 2005;89:1245-9. Medline:16170109 doi:10.1136/bjo.2005.067355
- 10 Quigley HA. Number of people with glaucoma worldwide. *Br J Ophthalmol.* 1996;80:389-93. Medline:8695555 doi:10.1136/bjo.80.5.389
- 11 Topouzis F, Coleman AL, Harris A, Koskosas A, Founti P, Gong G, et al. Factors associated with undiagnosed open-angle glaucoma: the Thessaloniki Eye Study. *Am J Ophthalmol.* 2008;145:327-35. Medline:18045565 doi:10.1016/j.ajo.2007.09.013
- 12 Nduaguba C, Lee RK. Glaucoma screening: current trends, economic issues, technology, and challenges. *Curr Opin Ophthalmol.* 2006;17:142-52. Medline:16552249 doi:10.1097/01.icu.0000193088.75432.c9
- 13 Rulli E, Biagioli E, Riva I, Gambirasio G, De Simone I, Floriani I, et al. Efficacy and safety of trabeculectomy vs nonpenetrating surgical procedures: a systematic review and meta-analysis. *JAMA Ophthalmol.* 2013;131:1573-82. Medline:24158640 doi:10.1001/jamaophthalmol.2013.5059
- 14 Schwartz GF. Compliance and persistency in glaucoma follow-up treatment. *Curr Opin Ophthalmol.* 2005;16:114-21. Medline:15744142 doi:10.1097/01.icu.0000156139.05323.26
- 15 Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol.* 2006;90:262-7. Medline:16488940 doi:10.1136/bjo.2005.081224
- 16 Tham Y-C, Li X, Wong TY, Quigley HA, Aung T, Cheng C-Y. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology.* 2014;121:2081-90. Medline:24974815 doi:10.1016/j.ophtha.2014.05.013
- 17 Chan EW, Li X, Tham Y-C, Liao J, Wong TY, Aung T, et al. Glaucoma in Asia: regional prevalence variations and future projections. *Br J Ophthalmol.* 2016;100:78-85. Medline:26112871 doi:10.1136/bjophthalmol-2014-306102
- 18 Wong TY, Loon S, Saw S. The epidemiology of age related eye diseases in Asia. *Br J Ophthalmol.* 2006;90:506-11. Medline:16547337 doi:10.1136/bjo.2005.083733
- 19 United Nations. World population ageing 2015. New York: Department of Economic and Social Affairs, United Nations; 2015.
- 20 Shetty P. Grey matter: ageing in developing countries. *Lancet.* 2012;379:1285-7. Medline:22489326 doi:10.1016/S0140-6736(12)60541-8
- 21 Foster PJ, Johnson GJ. Glaucoma in China: how big is the problem? *Br J Ophthalmol.* 2001;85:1277-82. Medline:11673287 doi:10.1136/bjo.85.11.1277
- 22 World Health Organization. World report on ageing and health. Geneva: World Health Organization; 2015.
- 23 Woo J, Kwok T, Sze F, Yuan H. Ageing in China: health and social consequences and responses. *Int J Epidemiol.* 2002;31:772-5. Medline:12177017 doi:10.1093/ije/31.4.772
- 24 Xia J, Wright J, Adams CE. Five large Chinese biomedical bibliographic databases: accessibility and coverage. *Health Info Libr J.* 2008;25:55-61. Medline:18251914 doi:10.1111/j.1471-1842.2007.00734.x
- 25 Fung IC. Chinese journals: a guide for epidemiologists. *Emerg Themes Epidemiol.* 2008;5:20. Medline:18826604 doi:10.1186/1742-7622-5-20
- 26 Song P, Theodoratou E, Li X, Liu L, Chu Y, Black RE, et al. Causes of death in children younger than five years in China in 2015: an updated analysis. *J Glob Health.* 2016;6:020802. Medline:28028436 doi:10.7189/jogh.06.020802
- 27 Wang YX, Xu L, Yang H, Jonas JB. Prevalence of glaucoma in North China: the Beijing eye study. *Am J Ophthalmol.* 2010;150:917-24. Medline:20970107 doi:10.1016/j.ajo.2010.06.037
- 28 Song W, Shan L, Cheng F, Fan P, Zhang L, Qu W, et al. Prevalence of glaucoma in a rural northern China adult population: a population-based survey in Kailu County, Inner Mongolia. *Ophthalmology.* 2011;118:1982-8. Medline:21684607 doi:10.1016/j.ophtha.2011.02.050
- 29 Sun J, Zhou X, Kang Y, Yan L, Sun X, Sui H, et al. Prevalence and risk factors for primary open-angle glaucoma in a rural northeast China population: a population-based survey in Bin County, Harbin. *Eye (Lond).* 2012;26:283. Medline:22157917 doi:10.1038/eye.2011.243
- 30 Cheng J-W, Cheng S-W, Ma X-Y, Cai J-P, Li Y, Wei R-L. The prevalence of primary glaucoma in mainland China: a systematic review and meta-analysis. *J Glaucoma.* 2013;22:301-6. Medline:22134352 doi:10.1097/IJG.0b013e31824083ca
- 31 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med.* 2009;151:264-9. Medline:19622511 doi:10.7326/0003-4819-151-4-200908180-00135
- 32 Stevens GA, Alkema L, Black RE, Boerma JT, Collins GS, Ezzati M, et al. Guidelines for accurate and transparent health estimates reporting: the GATHER statement. *PLoS Med.* 2016;13:e1002056. Medline:27351744 doi:10.1371/journal.pmed.1002056
- 33 Quigley HA, West SK, Munoz B, Mmbaga B, Glovinsky Y. Examination methods for glaucoma prevalence surveys. *Arch Ophthalmol.* 1993;111:1409-15. Medline:8216023 doi:10.1001/archophth.1993.01090100117038

- 34 Rudnicka AR, Mt-Isa S, Owen CG, Cook DG, Ashby D. Variations in primary open-angle glaucoma prevalence by age, gender, and race: a Bayesian meta-analysis. *Invest Ophthalmol Vis Sci.* 2006;47:4254-61. Medline:17003413 doi:10.1167/iovs.06-0299
- 35 Kapetanakis VV, Chan MP, Foster PJ, Cook DG, Owen CG, Rudnicka AR. Global variations and time trends in the prevalence of primary open angle glaucoma (POAG): a systematic review and meta-analysis. *Br J Ophthalmol.* 2016;100:86-93. Medline:26286821 doi:10.1136/bjophthalmol-2015-307223
- 36 Friedman DS, Wolfs RC, O'Colmain BJ, Klein BE, Taylor HR, West S, et al; Eye Diseases Prevalence Research Group. Prevalence of open-angle glaucoma among adults in the United States. *Arch Ophthalmol.* 2004;122:532. Medline:15078671 doi:10.1001/archophth.122.4.532
- 37 National Bureau of Statistics. Tabulation on the 2000 population census of the People's Republic of China. Beijing: China Statistics Press; 2002.
- 38 National Bureau of Statistics. Tabulation on the 2010 population census of the People's Republic of China. Beijing: China Statistics Press; 2012.
- 39 Barendregt JJ, Doi SA, Lee YY, Norman RE, Vos T. Meta-analysis of prevalence. *J Epidemiol Community Health.* 2013;67:974-8. Medline:23963506 doi:10.1136/jech-2013-203104
- 40 Higgins JP, Green S. *Cochrane handbook for systematic reviews of interventions.* Vol 5: Wiley Online Library; 2008.
- 41 Hox JJ, Moerbeek M, van de Schoot R. *Multilevel analysis: Techniques and applications.* Routledge; 2010.
- 42 Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw.* 2010;36:1-48. doi:10.18637/jss.v036.i03
- 43 United Nations, Department of Economic and Social Affairs, Population Division. *World Population Prospects, the 2015 Revision.* 2015. Available: <https://esa.un.org/unpd/wpp/>. Accessed: 8 January 2017.
- 44 Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am J Epidemiol.* 2007;165:710-8. Medline:17182981 doi:10.1093/aje/kwk052
- 45 Rudan I, Tomaskovic L, Boschi-Pinto C, Campbell H. Global estimate of the incidence of clinical pneumonia among children under five years of age. *Bull World Health Organ.* 2004;82:895-903. Medline:15654403
- 46 Fowkes FGR, Rudan D, Rudan I, Aboyans V, Denenberg JO, McDermott MM, et al. Comparison of global estimates of prevalence and risk factors for peripheral artery disease in 2000 and 2010: a systematic review and analysis. *Lancet.* 2013;382:1329-40. Medline:23915883 doi:10.1016/S0140-6736(13)61249-0
- 47 Adeloye D, Chua S, Lee C, Basquill C, Papana A, Theodoratou E, et al. Global and regional estimates of COPD prevalence: Systematic review and meta-analysis. *J Glob Health.* 2015;5:020415. Medline:26755942 doi:10.7189/jogh.05.020415
- 48 Yamamoto T, Iwase A, Araie M, Suzuki Y, Abe H, Shirato S, et al. The Tajimi Study report 2: prevalence of primary angle closure and secondary glaucoma in a Japanese population. *Ophthalmology.* 2005;112:1661-9. Medline:16111758 doi:10.1016/j.ophtha.2005.05.012
- 49 Wong WL, Su X, Li X, Cheung CMG, Klein R, Cheng C-Y, et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *Lancet Glob Health.* 2014;2:e106-16. Medline:25104651 doi:10.1016/S2214-109X(13)70145-1
- 50 Welch VA, Ghogomu E, Hossain A, Awasthi S, Bhutta ZA, Cumberbatch C, et al. Mass deworming to improve developmental health and wellbeing of children in low-income and middle-income countries: a systematic review and network meta-analysis. *Lancet Glob Health.* 2017;5:e40-50. Medline:27955788 doi:10.1016/S2214-109X(16)30242-X
- 51 Song P, Chang X, Wang M, An L. Variations of pterygium prevalence by age, gender and geographic characteristics in China: A systematic review and meta-analysis. *PLoS One.* 2017;12:e0174587. Medline:28355301 doi:10.1371/journal.pone.0174587
- 52 Gabelt BT, Kaufman PL. Changes in aqueous humor dynamics with age and glaucoma. *Prog Retin Eye Res.* 2005;24:612-37. Medline:15919228 doi:10.1016/j.preteyeres.2004.10.003
- 53 Quigley HA. Glaucoma. *Lancet.* 2011;377:1367-77. Medline:21453963 doi:10.1016/S0140-6736(10)61423-7
- 54 Foster PJ, Oen FT, Machin D, Ng T-P, Devereux JG, Johnson GJ, et al. The prevalence of glaucoma in Chinese residents of Singapore: a cross-sectional population survey of the Tanjong Pagar district. *Arch Ophthalmol.* 2000;118:1105-11. Medline:10922206 doi:10.1001/archophth.118.8.1105
- 55 Friedman DS, Gazzard G, Min CB, Broman AT, Quigley H, Tielsch J, et al. Age and sex variation in angle findings among normal Chinese subjects: a comparison of UBM, Scheimpflug, and gonioscopic assessment of the anterior chamber angle. *J Glaucoma.* 2008;17:5-10. Medline:18303376 doi:10.1097/IJG.0b013e31806ab327
- 56 Xu L, Cao WF, Wang YX, Chen CX, Jonas JB. Anterior chamber depth and chamber angle and their associations with ocular and general parameters: the Beijing Eye Study. *Am J Ophthalmol.* 2008;145:929-36. Medline:18336789 doi:10.1016/j.ajo.2008.01.004
- 57 Vajaranant TS, Nayak S, Wilensky JT, Joslin CE. Gender and glaucoma: what we know and what we need to know. *Curr Opin Ophthalmol.* 2010;21:91. Medline:20051857 doi:10.1097/ICU.0b013e3283360b7e
- 58 Cheng J-W, Zong Y, Zeng Y-Y, Wei R-L. The prevalence of primary angle closure glaucoma in adult Asians: a systematic review and meta-analysis. *PLoS One.* 2014;9:e103222. Medline:25057993 doi:10.1371/journal.pone.0103222
- 59 He M, Foster P, Johnson G, Khaw P. Angle-closure glaucoma in East Asian and European people. Different diseases? *Eye (Lond).* 2006;20:3-12. Medline:15688051 doi:10.1038/sj.eye.6701797
- 60 Congdon NG, Foster P, Wamsley S, Gutmark J, Nolan W, Seah S, et al. Biometric gonioscopy and the effects of age, race, and sex on the anterior chamber angle. *Br J Ophthalmol.* 2002;86:18-22. Medline:11801496 doi:10.1136/bjo.86.1.18

- 61 Leung CK, Palmiero P, Weinreb R, Li H, Sbeity Z, Dorairaj S, et al. Comparisons of anterior segment biometry between Chinese and Caucasians using anterior segment optical coherence tomography. *Br J Ophthalmol*. 2010;94:1184-9. Medline:20610476 doi:10.1136/bjo.2009.167296
- 62 Kwon YH, Fingert JH, Kuehn MH, Alward WL. Primary open-angle glaucoma. *N Engl J Med*. 2009;360:1113-24. Medline:19279343 doi:10.1056/NEJMra0804630
- 63 Bonovas S, Peponis V, Filioussi K. Diabetes mellitus as a risk factor for primary open-angle glaucoma: a meta-analysis. *Diabet Med*. 2004;21:609-14. Medline:15154948 doi:10.1111/j.1464-5491.2004.01173.x
- 64 Gong P, Liang S, Carlton EJ, Jiang Q, Wu J, Wang L, et al. Urbanisation and health in China. *Lancet*. 2012;379:843-52. Medline:22386037 doi:10.1016/S0140-6736(11)61878-3
- 65 Zhang KH, Shunfeng S. Rural-urban migration and urbanization in China: Evidence from time-series and cross-section analyses. *China Econ Rev*. 2003;14:386-400. doi:10.1016/j.chieco.2003.09.018
- 66 Qu W, Li Y, Song W, Zhou X, Kang Y, Yan L, et al. Prevalence and risk factors for angle-closure disease in a rural Northeast China population: a population-based survey in Bin County, Harbin. *Acta Ophthalmol*. 2011;89:e515-20. Medline:21834920 doi:10.1111/j.1755-3768.2011.02146.x
- 67 Casson RJ, Newland H, Muecke J, McGovern S, Abraham L, Shein W, et al. Gonioscopy findings and prevalence of occludable angles in a Burmese population: the Meiktila Eye Study. *Br J Ophthalmol*. 2007;91:856-9. Medline:17576708 doi:10.1136/bjo.2006.108688
- 68 Barton K, Hitchings RA. Medical management of glaucoma. Springer; 2013:71-100.
- 69 Hernández R, Rabindranath K, Fraser C, Vale L, Blanco AA, Burr JM. Screening for open angle glaucoma: systematic review of cost-effectiveness studies. *J Glaucoma*. 2008;17:159-68. Medline:18414099 doi:10.1097/IJG.0b013e31814b9693
- 70 Hernández RA, Burr JM, Vale LD. Economic evaluation of screening for open-angle glaucoma. *Int J Technol Assess Health Care*. 2008;24:203-11. Medline:18400124 doi:10.1017/S0266462308080288
- 71 Vaahtoranta-Lehtonen H, Tuulonen A, Aronen P, Sintonen H, Suoranta L, Kovanen N, et al. Cost effectiveness and cost utility of an organized screening programme for glaucoma. *Acta Ophthalmol Scand*. 2007;85:508-18. Medline:17655612 doi:10.1111/j.1755-3768.2007.00947.x

# Prevalence of epilepsy in China between 1990 and 2015: A systematic review and meta-analysis

Peige Song<sup>1\*</sup>, Yezhou Liu<sup>2,3\*</sup>, Xinwei Yu<sup>2,4\*</sup>, Jingjing Wu<sup>2,5</sup>, Adrienne N Poon<sup>1,6</sup>, Alessandro Demaio<sup>7,8</sup>, Wei Wang<sup>2,4\*\*</sup>, Igor Rudan<sup>1\*\*</sup>, Kit Yee Chan<sup>1,9\*\*</sup>; on behalf of the Global Health Epidemiology Research Group (GHERG)

<sup>1</sup> Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Scotland, UK

<sup>2</sup> Beijing Municipal Key Laboratory of Clinical Epidemiology, School of Public Health, Capital Medical University, Beijing, China

<sup>3</sup> Xijing Hospital of Digestive Diseases, the Fourth Military Medical University

<sup>4</sup> School of Medical and Health Sciences, Edith Cowan University, Perth, Western Australia, Australia

<sup>5</sup> Beijing Tiantan Hospital, Capital Medical University, Beijing, China

<sup>6</sup> Department of Internal Medicine, School of Medicine & Health Sciences, The George Washington University, Washington, D.C., USA

<sup>7</sup> Harvard Global Equity Initiative, Harvard Medical School, Boston, Massachusetts, USA

<sup>8</sup> Copenhagen School of Global Health, University of Copenhagen, Denmark

<sup>9</sup> Nossal Institute for Global Health, University of Melbourne, Australia

\* Joint-first authors

\*\* Joint-last authors

## Correspondence to:

Peige Song  
Kit Yee Chan  
Igor Rudan  
Centre for Global Health Research  
Usher Institute for Population Health Sciences and Informatics,  
The University of Edinburgh  
Teviot Place  
Edinburgh EH8 9AG  
Scotland, UK  
p.song@sms.ed.ac.uk  
k.chan@ed.ac.uk  
igor.rudan@ed.ac.uk

**Background** Epilepsy is a major neurological disorder that affects approximately 65 million people worldwide. Globally, the burden of epilepsy is not evenly distributed, with more than 80% of sufferers residing in low- and middle-income countries. This study estimates the burden of epilepsy in mainland China from 1990 to 2015 and explores the variations of burden by age and gender.

**Methods** We conducted a systematic review of the peer-reviewed literature from 1990 to 2015 using Chinese and English academic databases (CNKI, WanFang, VIP and PubMed) to identify population-based prospective studies on the prevalence of epilepsy in mainland Chinese. Multilevel mixed-effects logistic regression was used to estimate the prevalence of lifetime epilepsy (LTE), and restricted cubic regression splines were applied to model the functional forms of the non-linear effects of age and LTE prevalence. Random-effects meta-analysis was used to obtain the pooled prevalence of 1-year active epilepsy (AE), 2-year AE and 5-year AE separately. To estimate the number of people with LTE and AE in the years 1990, 2000, and 2015, LTE and AE prevalence were multiplied by the total population of mainland China of the corresponding year.

**Findings** Analyses were conducted using 39 prevalence studies that met the inclusion criteria and comprised 77 separate data points (37 on LTE, 16 on 1-year AE, 12 on 2-year AE and 12 on 5-year AE). In 1990, the prevalence of LTE ranged from 1.31‰ (95% CI=0.85–2.00) in the 0–4 age group to 2.42‰ (95% confidence interval CI=1.60–3.65) in the 30–34 age group. By 2015, the LTE prevalence had increased to 4.57‰ (95% CI=2.52–8.27) in the 0–4 group and 8.43‰ (95% CI=4.71–15.04) in the 30–34 group. Over the 25-year period, the overall prevalence of LTE had steadily increased by 259%, from 1.99‰ (95% CI=1.31–3.02) in 1990 to 7.15‰ (95% CI=3.98–12.82) in 2015. The rates of increase were similar across the whole age spectrum, fluctuating around 250%. Between 1990 and 2015, the total number of people with LTE in mainland China increased by 328%, from 2.30 million (95% CI=1.51–3.49) in 1990 to 9.84 million (95% CI=5.48–17.64) in 2015. The pooled 1-year, 2-year, and 5-year AE prevalence were 3.79‰ (95% CI=3.31–4.34), 4.08‰ (95% CI=3.41–4.89) and 4.19‰ (95% CI=3.42–5.15).

**Conclusions** The burden of LTE in China has increased substantially between 1990 and 2015, with the prevalence of LTE having more than doubled and the number of people with LTE more than tripled. The large amount of AE cases in China calls for optimal management and treatment. More high-quality epidemiological studies on LTE and AE prevalence are still needed.



Epilepsy is a disorder of neuronal excitability, characterised predominantly by unpredictable and recurrent seizures of cerebral origin [1]. As one major neurological disorder, epilepsy affects approximately 65 million people worldwide, ranging from neonates to elderly [2–5]. The burden of epilepsy is not only limited in neurological deficits but also includes devastating psychological and psychiatric problems [6,7], influencing the quality of personal, familial and social life significantly [6,8,9]. If left untreated, epilepsy would be incapacitating and sometimes fatal. People living with epilepsy generally have higher disability and mortality rates [5,10,11]. Worldwide, it is estimated that mortality in people with epilepsy is two to three times higher than in the general population [12], while the Global Burden of Disease (GBD) study for 2013 estimated the disability-adjusted life year for epilepsy to be 253 per 100 000 people [13,14]. Although effective and cost-effective medications exist for controlling seizures, many people with epilepsy are excluded from treatments due to cultural, economic and other factors [11,15]. The global burden of epilepsy is not distributed evenly, with more than 80% of people with epilepsy residing in low- and middle-income countries (LMICs) [4,16], where the majority of people with epilepsy receive inadequate treatment and management [9,16].

Quality epidemiological data are crucial for estimating the burden of epilepsy which in turn serve to inform policies on resource allocation and disease management [17,18]. Over the past decades, the number of prevalence studies on epilepsy has grown considerably, making it possible to synthesise prevalence of epilepsy at a regional and global level [10,19–24]. Such estimates are valuable despite potential confounders that arise with the difference in sampling, case definitions, case ascertainment, and screening tools [10,21,23]. The prevalence estimates of epilepsy in developed countries are generally consistent with each other, whereas those in developing countries are often made for isolated geographical areas and they vary widely [10,22,23]. It is estimated that the prevalence of epilepsy in LMICs is twice as high as in high-income countries (HICs), making it an even more important global health issue in these settings [16,25].

Nevertheless, the reported prevalence of epilepsy is likely to be conservative, because underdiagnosis and misdiagnosis are common in resource-poor areas [10,25]. Furthermore, as a disease marked with stigma and prejudice across the world and throughout history, epilepsy may be largely concealed because of profound cultural and social restrictions [7,8,26]. The number of high-quality studies regarding the epidemiology of epilepsy in the developing world is quite small, which makes the estimate of epilepsy prevalence very difficult [10].

In China, although substantial economic development and improvement of health services occurred in the past decades, the diversity of development and demographic structures across the whole country still limits the opportunity for an estimate of epilepsy prevalence at the national level [19,27]. However, the large volume of data on the prevalence of epilepsy in Chinese bibliographical databases makes it feasible to explore the burden of epilepsy from a modelling approach [28,29]. For instance, Lian Gu and colleagues revealed an overall epilepsy prevalence of 2.89‰ in Mainland China by using the meta-analysis method [19]. However, no temporal trend has been analysed in their study, and the prevalence estimates were not conducted in a large age span, which may have limited their study to some extent. Moreover, as new evidence continues to emerge, such evidence synthesis should ideally be updated using a more detailed approach. For these reasons, we conducted a systematic review of the literature in both Chinese and English databases to analyse the temporal distribution of epilepsy prevalence in China from 1990 to 2015. We also investigated the variations in prevalence by age and gender.

## METHODS

### Literature search

We conducted a parallel systematic review of the published literature from 1990 to 2015 using PubMed and three Chinese databases; China National Knowledge Infrastructure (CNKI), Wanfang Data, VIP in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Guidelines and the Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER) statement [30,31]. The search strategy for PubMed was ((epilepsy) AND (China OR Chinese) AND (inciden\* OR prevalen\* OR morbidity OR mortality)). The search terms for the Chinese databases were the term epilepsy in both China and English, and two versions of Chinese terms for ‘incidence’ and ‘prevalence’, attack rate, mortality, fatality, epi\*, burden, epidemiological survey and cross-sectional investigation. The precise combination of the search term for the four databases search is detailed in Table S1 in the **Online Supplementary Document**.

## Selection criteria

Our inclusion criteria were: (1) population-based studies; (2) studies of mainland Chinese populations in mainland China; (3) studies that provide prevalence of lifetime epilepsy (LTE) and/or active epilepsy (AE); (4) studies that include clear case definitions (the case definitions used to define a case of epilepsy are described in detail in Table S2 in **Online Supplementary Document**). Our exclusion criteria were: (1) case-control and hospital-based studies; (2) studies of populations outside of mainland China (including Hong Kong and Taiwan); (3) studies of specific areas; eg, mining or fishery districts; and (4) reviews and conference abstracts; (5) duplicate publications; (6) studies that included febrile convulsions and provoked seizures in their estimates of epilepsy prevalence; (7) studies with unclear case definitions; and (8) studies with inconsistent results.

## Data extraction

Data were independently extracted (PS, YZ and XY for data in Chinese; PS, KYC and AP for data in English). A database was set up to record the extracted information which included names of authors, published year, study setting (urban and/or rural), medium year of data collection, sampling method, case definition, sample size, number of epilepsy cases, and epilepsy prevalence.

In epidemiological studies of epilepsy, the prevalence estimates of LTE and AE are generally reported separately, where the LTE prevalence is the proportion of individuals manifesting a disorder anytime during the earlier period of their life up to the point of investigation [32], and AE prevalence represents the proportion of individuals who have experienced at least two unprovoked seizures within a certain period of time (1 year, 2 years, and 5 years) up to the point of investigation [2,10,33,34]. In data extraction process, prevalence estimates were classified into these two exclusive groups based on the definitions or investigation methodologies provided in each study.

## Statistical analysis

Our analysis of epilepsy prevalence was conducted for LTE, 1-year AE, 2-year AE and 5-year AE separately. For studies that provided estimates of LTE prevalence, multiple data points were available in each study to contribute to the overall database. To take into account the availability of different data points from the same study, meta-analysis via multilevel mixed-effects models was adopted [35]. To investigate whether LTE prevalence varied significantly according to different demographic factors (age and gender) or had a secular trend (study year), univariate meta-regression was adopted to test their significance consecutively. Restricted cubic regression splines were used to model the functional forms of the non-linear effects of age and LTE prevalence. Variables that significantly correlated with LTE prevalence in the univariate analyses were then included in the final multivariate regression model. Due to the paucity of studies that reported AE prevalence, effects of demographic factors and secular trends could not be explored using meta-regression. Instead, random-effects meta-analysis (DerSimonian Laird method) was applied to obtain the pooled prevalence [36].

To estimate the number of people with LTE and AE in the years 1990, 2000 and 2015, LTE and AE prevalence were multiplied by the population of China in the corresponding years using population data from the United Nations Population Division (UNPD) [37]. All the analyses were conducted in R v3.3.0 (R Development Core Team; <http://www.R-project.org>).

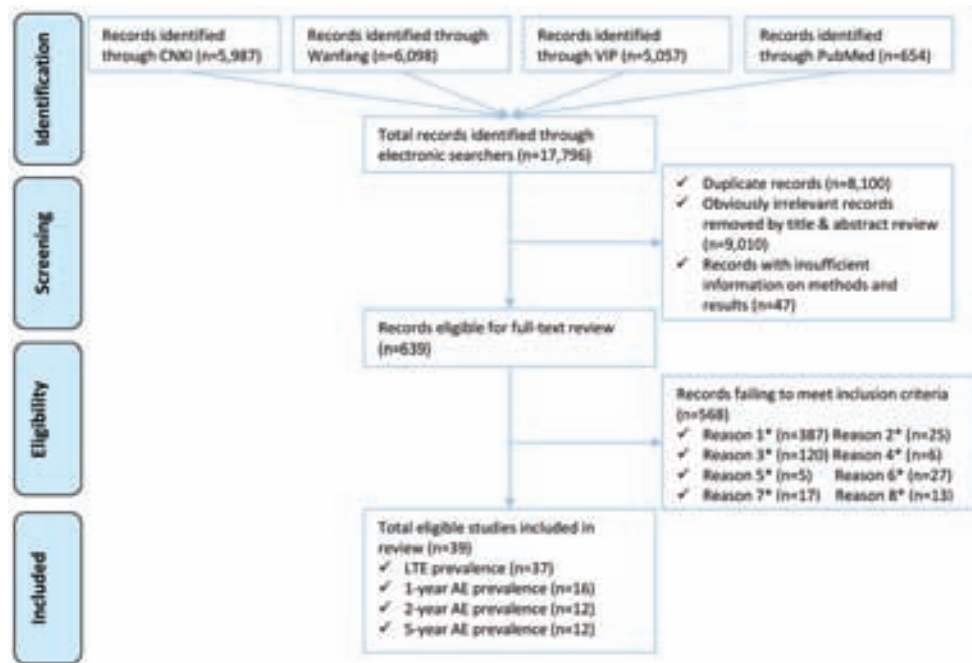
## RESULTS

### Systematic review

Our database searches returned 17 796 titles. After removing 8100 duplicates, and 9010 titles and abstracts that contained no information on epilepsy prevalence, and 47 papers with insufficient information on methods and results, 639 full-text papers were reviewed. Of the 39 full-text papers that met our inclusion criteria, 37 reported LTE prevalence, 16 reported 1-year AE prevalence, 12 reported 2-year AE prevalence and 12 reported 5-year AE prevalence (**Figure 1**).

### Study characteristics

**Table 1** summarises the key characteristics of the 39 studies. All studies were cross-sectional in design. Most of them were large studies published after 2000. Participants were typically investigated by neurologists.



**Figure 1.** PRISMA flow diagram. \*Note: \*Reason 1–Papers that were not population-based epilepsy study; \*Reason 2–Studies that were not based in Mainland China; \*Reason 3–Papers with no numerical prevalence measure of epilepsy; \*Reason 4–Studies with no clear time period; \*Reason 5–Studies only reported febrile convulsions or provoked seizures; \*Reason 6–Multiple publications of the same study; \*Reason 7–Studies only reported specific subtype of epilepsy; \*Reason 8–Unclear prevalence type.

**Table 1.** Main characteristics of the retained studies

CHARACTERISTICS OF STUDY	STUDIES WITH LTE PREVALENCE (N = 37, %)	STUDIES WITH 1-YEAR AE PREVALENCE (N = 16, %)	STUDIES WITH 2-YEAR AE PREVALENCE (N = 12, %)	STUDIES WITH 5-YEAR AE PREVALENCE (N = 12, %)
<b>Year published:</b>				
1990–1999	12 (32.4)	0 (0.0)	0 (0.0)	0 (0.0)
2000–2009	16 (43.2)	11 (68.8)	8 (66.7)	7 (58.3)
2010–2016	9 (24.3)	5 (31.3)	4 (33.3)	5 (41.7)
<b>Setting:</b>				
Urban	4 (10.8)	1 (6.3)	1 (8.3)	1 (8.3)
Rural	21 (56.8)	14 (87.5)	10 (83.3)	10 (83.3)
Mixed	9 (24.3)	1 (6.3)	1 (8.3)	1 (8.3)
Both	3 (8.1)	0 (0.0)	0 (0.0)	0 (0.0)
<b>Sample size:</b>				
4000–10000	9 (24.3)	2 (12.5)	1 (8.3)	1 (8.3)
10001–50000	17 (45.9)	11 (68.8)	9 (75.0)	10 (83.3)
50001–200000	8 (21.6)	2 (12.5)	2 (16.7)	1 (8.3)
200001–900000	3 (8.1)	1 (6.3)	0 (0.0)	0 (0.0)
<b>Screening tool:</b>				
WHO questionnaire	14 (37.8)	9 (56.3)	7 (58.3)	7 (58.3)
ICBERG questionnaire	5 (13.5)	4 (25.0)	4 (33.3)	4 (33.3)
Self-designed questionnaire	10 (27)	1 (6.3)	0 (0.0)	0 (0.0)
Questionnaire based on ILAE/CMA/ BNI diagnosis	7 (18.9)	2 (12.5)	1 (8.3)	1 (8.3)
Not specified	1 (2.7)	0 (0.0)	0 (0.0)	0 (0.0)
<b>Diagnosis of epilepsy:</b>				
By neurologists	27 (73.0)	12 (75.0)	9 (75.0)	10 (83.3)
By trained physicians	6 (16.2)	3 (18.8)	2 (16.7)	2 (16.7)
By trained investigators	2 (5.4)	0 (0.0)	0 (0.0)	0 (0.0)
Not specified	2 (5.4)	1 (6.3)	1 (8.3)	0 (0.0)

LTE – lifetime epilepsy, AE – active epilepsy, WHO – World Health Organization, ICBERG – International Community-based Epilepsy Research Group, ILAE – International League Against Epilepsy, CMA – Chinese Medical Association, BNI – Beijing Neurosurgical Institute

ogists using standard international questionnaires (eg, by World Health Organization [WHO] or International Community–based Epilepsy Research Group [ICBERG]). For more details of the studies, see Table S2 in the **Online Supplementary Document**.

### Estimates of LTE prevalence and number of cases in China

The 37 studies that reported LTE prevalence involved a combined total of 2 851 219 participants. Of these, 5813 met the criteria for LTE diagnosis, giving an LTE prevalence of 2.04‰. A total of 274 specific data points based on age, gender and location provided the information on LTE prevalence. Based on these informative data points, the gender–specific relationship between age and LTE prevalence was explored and it is shown in **Figure 2**. Generally, the larger studies yielded lower LTE prevalence in both males and females across most of the age spectrum.

In the univariate meta–regression analysis (Table S3 in **Online Supplementary Document**), no gender difference was found in LTE prevalence, whereas age and study year were all significantly associated with LTE prevalence. The final formula generated from the multilevel mixed–effects meta–analysis is shown below:

$$\ln(\text{odds}) = -106.597 + 0.050 * \text{year} + 0.032 * \text{age1} - 0.081 * \text{age2} + 0.146 * \text{age3} + 0.058 * \text{age4}$$

Where:

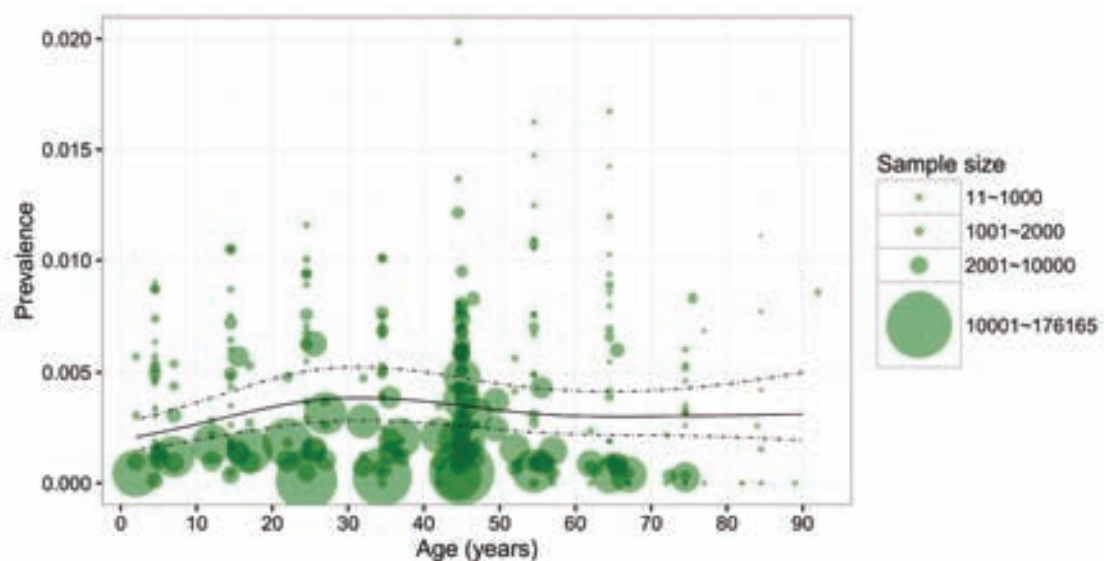
odds =  $p/(1-p)$ ,  $p$  indicates the prevalence of LTE

year = calendar year

age1–age4 are variables created in the process of fitting cubic splines (knots: 4.5, 24.5, 45.0, 54.5, 74.5)

Based on the final regression model, age–specific LTE prevalence in mainland China was calculated for the years 1990, 2000 and 2015 (**Table 2** and **Figure 3**). In 1990, the prevalence of LTE was lowest in the 0–4 age group (1.31‰; 95% CI=0.85–2.00) and highest in the 30–34 age group (2.42‰; 95% CI:1.60–3.65). By 2015, this prevalence had increased by three–fold to 4.57‰ (95% CI=2.52–8.27) in the 0–4 age group and 8.43‰ in the 30–34 age group (95% CI=4.71–15.04). Over 25 years, the overall LTE prevalence had steadily increased by 259%, from 1.99‰ (95% CI=1.31–3.02) in 1990 to 7.15‰ (95% CI=3.98–12.82) in 2015. This rate of increase is similar across the entire age spectrum, fluctuating around 250%.

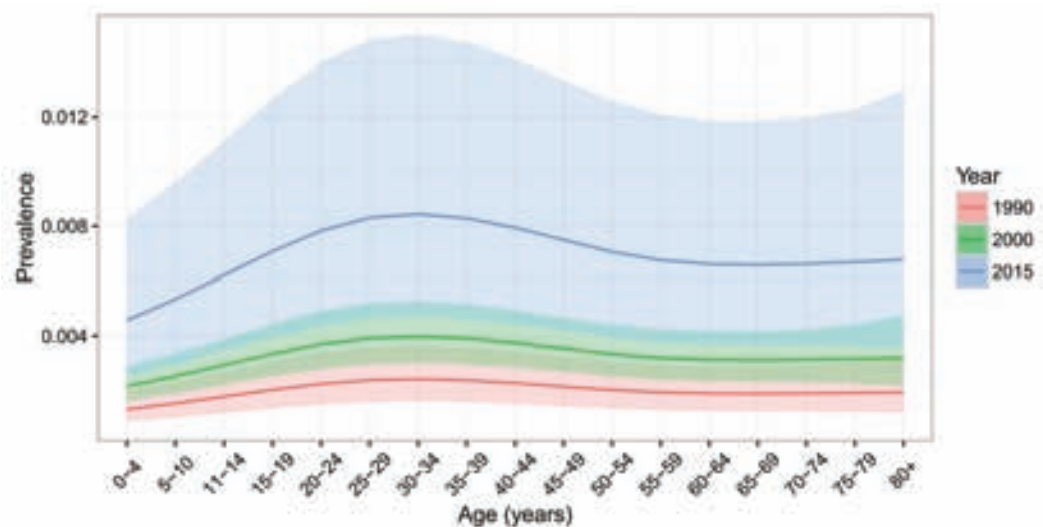
The estimated number of individuals with LTE in China was 2.30 million (95% CI=1.51–3.49) in 1990, and 9.84 million (95% CI=5.48–17.64) in 2015, which implied an overall increase of 328% throughout this period (**Table 2** and **Figure 4**). The most significant increase of LTE cases was observed among peo-

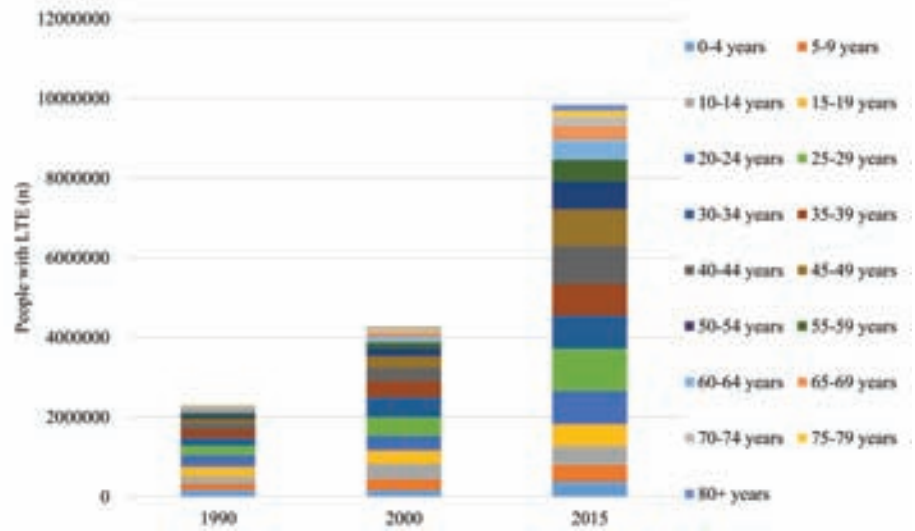


**Figure 2.** Age–specific prevalence of lifetime epilepsy (LTE) in China based on the data points from the included studies. Note: The size of each bubble is proportional to the sample size, the regression lines are based on the retained data points.

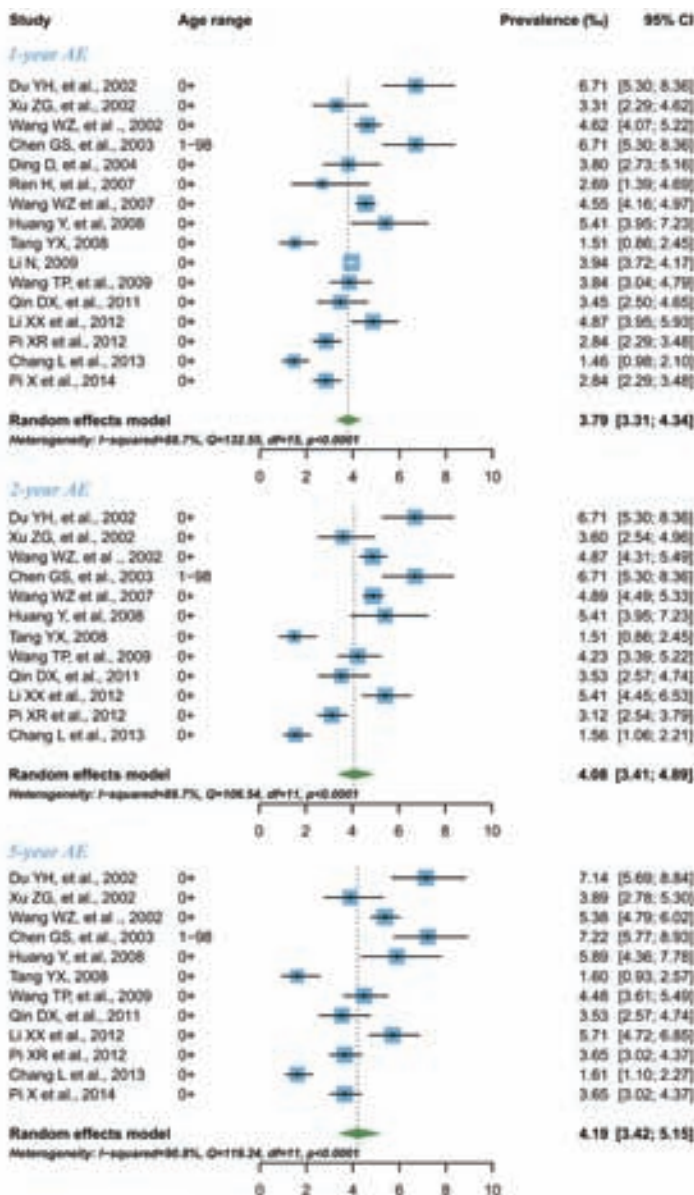
**Table 2.** Estimated age-specific prevalence of and numbers of people with lifetime epilepsy (LTE) in China in the years 1990, 2000 and 2015, and the rate of change from 1990 to 2015 by age group

Age (years)	PREVALENCE OF LTE (%)			NUMBER OF LTE CASES (THOUSANDS)			RATE OF CHANGE (1990–2015) FOR:	
	1990	2000	2015	1990	2000	2015	Prevalence (in %)	Number of cases (in %)
0–4	1.31 (0.85–2.00)	2.16 (1.60–2.90)	4.57 (2.52–8.27)	173.05 (113.00–264.94)	172.13 (127.74–231.91)	380.15 (209.75–687.84)	+250	+119
5–9	1.53 (1.00–2.32)	2.52 (1.89–3.36)	5.34 (2.96–9.60)	155.95 (102.60–236.98)	270.71 (203.15–360.67)	419.97 (233.14–755.10)	+250	+169
10–14	1.77 (1.17–2.69)	2.93 (2.21–3.88)	6.20 (3.45–11.12)	174.30 (114.97–264.15)	383.76 (289.21–509.10)	466.92 (259.90–837.01)	+250	+168
15–19	2.02 (1.33–3.07)	3.34 (2.52–4.43)	7.07 (3.94–12.67)	249.43 (164.44–378.22)	338.96 (255.33–449.87)	558.11 (310.78–999.76)	+249	+124
20–24	2.24 (1.48–3.40)	3.70 (2.78–4.91)	7.82 (4.36–14.00)	288.19 (189.85–437.31)	360.06 (271.02–478.21)	830.20 (462.40–1486.46)	+249	+188
25–29	2.38 (1.57–3.60)	3.92 (2.96–5.20)	8.30 (4.63–14.83)	249.27 (164.43–377.73)	478.13 (360.75–633.51)	1071.19 (597.59–1914.55)	+249	+329
30–34	2.42 (1.60–3.65)	3.99 (3.02–5.26)	8.43 (4.71–15.04)	205.79 (136.04–311.18)	506.37 (383.40–668.57)	837.98 (468.43–1494.68)	+249	+307
35–39	2.37 (1.57–3.58)	3.91 (2.97–5.16)	8.28 (4.63–14.76)	208.14 (137.67–314.55)	404.40 (306.56–533.29)	788.56 (441.14–1405.54)	+249	+278
40–44	2.27 (1.50–3.43)	3.75 (2.84–4.94)	7.93 (4.44–14.14)	142.97 (94.56–216.08)	313.54 (237.73–413.40)	944.93 (528.69–1684.21)	+249	+560
45–49	2.14 (1.42–3.24)	3.54 (2.68–4.66)	7.49 (4.19–13.34)	104.03 (68.84–157.15)	302.62 (229.67–398.62)	924.81 (517.65–1647.94)	+249	+789
50–54	2.02 (1.33–3.06)	3.33 (2.52–4.41)	7.05 (3.94–12.59)	91.45 (60.35–138.54)	201.68 (152.44–266.74)	700.89 (391.51–1251.68)	+249	+666
55–59	1.93 (1.27–2.94)	3.19 (2.39–4.25)	6.75 (3.76–12.10)	78.90 (51.83–120.07)	145.26 (109.05–193.44)	532.83 (296.59–954.95)	+249	+575
60–64	1.89 (1.24–2.88)	3.12 (2.34–4.17)	6.61 (3.67–11.86)	61.42 (40.29–93.58)	127.25 (95.33–169.81)	514.26 (285.89–922.86)	+249	+737
65–69	1.88 (1.23–2.87)	3.11 (2.32–4.16)	6.58 (3.65–11.83)	46.17 (30.22–70.50)	105.87 (79.04–141.76)	336.59 (186.75–605.22)	+250	+629
70–74	1.89 (1.23–2.92)	3.12 (2.30–4.24)	6.62 (3.64–11.98)	35.68 (23.13–55.01)	74.30 (54.68–100.93)	223.30 (122.97–404.51)	+250	+526
75–79	1.91 (1.21–3.01)	3.15 (2.25–4.41)	6.68 (3.62–12.29)	21.30 (13.52–33.53)	45.91 (32.82–64.19)	161.31 (87.46–296.77)	+250	+657
80+ years	1.94 (1.18–3.19)	3.20 (2.15–4.75)	6.77 (3.53–12.95)	13.83 (8.41–22.77)	38.51 (25.92–57.26)	151.44 (79.00–289.61)	+250	+995
Total	1.99 (1.31–3.02)	3.36 (2.53–4.47)	7.15 (3.98–12.82)	2299.87 (1514.15–3492.29)	4269.45 (3213.85–5671.27)	9843.44 (5479.62–17638.67)	+259	+328

**Figure 3.** Age-specific prevalence of lifetime epilepsy (LTE) in China in the years 1990, 2000 and 2015, with 95% confidence intervals.



**Figure 4.** Estimated numbers of people with lifetime epilepsy (LTE) in China by year and age group.



**Figure 5.** Pooled prevalence of 1-year, 2-year and 5-year active epilepsy (AE) in China by random-effects meta-analysis.

ple aged 40 years and above, where the rates of change were all above 500%, and the highest was noted in people aged 80+ years. Most LTE cases were in individuals aged 20–24 years in 1990, and then shifted to the group of 25–29 years in 2015.

### Estimates of AE prevalence and number of cases in China

As shown in Figure 5, the pooled 1-year AE prevalence across all time periods was 3.79% (95% CI=3.31–4.34), and the 2-year and 5-year AE prevalence were slightly higher, as expected, amounting to 4.08% (95% CI=3.41–4.89) and 4.19% (95% CI=3.42–5.15) respectively. When these estimates are applied to the Chinese population size in the year 2015, the numbers of individuals with 1-year, 2-year and 5-year AE cases were estimated to 5.22 million (95% CI=4.55–5.97), 5.61 million (95% CI=4.69–6.73) and 5.77 million (95% CI=4.71–7.09), accounting for more than half of all the contemporary LTE cases.

### DISCUSSION

This study describes a comprehensive estimate of the prevalence of epilepsy in China from 1990 to 2015. Although this is the second attempt to summarize the available data on epilepsy prevalence in China, our study is the first (to the best of our knowledge) that provides robust estimates of LTE prevalence across the entire age range. It also reveals the prevalence of 1-year, 2-year and 5-year AE simultaneously in Chinese population. This is of particular clinical and public health interest, because individuals with AE may benefit more from treatment compared to those with LTE [10,33]. In addition, we also conducted a temporal analysis of LTE burden in China from 1990 to 2015. This innovative exploration is of potential

value in studying the effect of previous national antiepileptic programs. Significant strengths of this study include exhaustive search strategy, attempts to reduce bias by rigorous study selection procedure, parallel review and extraction of data.

In the current study, LTE was defined as “at least two unprovoked (or reflex) seizures, except for febrile convulsion and seizure induced by acute brain diseases at some point during lifetime”. Our study revealed LTE prevalence of 1.99‰ in 1990 and 7.15‰ in 2015 in China. Our estimate of LTE prevalence for the year 1990 was in line with the estimates presented in the previous systematic review, which reported a prevalence of 1.84‰ for the period of 1986–1990, and 3.58‰ for the period of 1991–1995 respectively [19]. Our estimate of LTE prevalence in 2015 was higher than their estimate for the period from 2006–2010 (7.15 ‰ vs 6.62‰), which is plausible given that a dramatic increase of LTE prevalence was encountered in our study.

When we perform comparisons to the global estimate of LTE prevalence, our estimate of LTE prevalence for the year 2015 (7.15‰) was even lower than the lower band of the estimated range in LMICs (8.75‰, 95% CI=7.23–10.59) but equal to the upper band of that in HICs; 5.18‰, 95% CI=3.75–7.15) [16,24,38]. Similarly, our estimate of 5-year AE (4.19‰) is much more in line with the estimate in developed countries (4.9‰, 95% CI=2.3–10.3), but less than in developing countries [10]. Despite the higher rates of spontaneous remission of epilepsy cases, given that fact that epilepsy is still a highly stigmatising disorder in China, any concealment could make these estimates even higher [26,39].

Epilepsy is known to affect people of all ages, though more frequently affecting young people [40]. The estimates in our study indicated that the LTE prevalence peaked at individuals aged 30–34 years. This finding confirms the statement made in previous studies that the epilepsy is generally a disease of the young [23]. The peak of this disorder in young age could partly be explained by the accumulation of early-onset epilepsy cases. It is estimated that the incidence of epilepsy is the highest in young children and elderly, forming a characteristic U-shape [21,41]. In addition, individuals with epilepsy may at a higher risk of premature death. Two previous studies reported that the risk of premature death in individuals with epilepsy is 3–5 times higher than in general population, and especially among the young [42,43]. According to the survival effects theory, it is plausible that the lifetime prevalence of epilepsy showed a decreasing trend after the age of 30–34 years in our study. This phenomenon has also been seen in many other neurological diseases [44].

The gender difference is another interesting topic in epilepsy research, both relevant to the public health and clinical research. Previous investigations reported slightly higher prevalence estimates of epilepsy in males than in females, which was probably attributable to the inherent differences in healthy brain development between genders and marked social effects on disease risks and courses [19,45,46]. However, in our analysis, the difference of LTE prevalence between sexes was not statistically significant. This finding is in contrast with the previous systematic review of epilepsy prevalence in China, but in line with the synthesised results in Asia, Latin America and Europe and the whole globe [21,23,24,47]. These inconsistencies with other studies may be a result of the different proportional contribution of epilepsy subtypes to the included case series. According to previous evidence, localization-related symptomatic epilepsy was more prevalent in males, and cryptogenic localization-related epilepsy was more prevalent in females [46,48].

In this study, we confirmed the hypothesis that the prevalence of LTE was continually increasing by presenting a positive temporal trend between 1990 and 2015. The prevalence of LTE is determined by the incidence rate at which new cases arise and the mortality rate [10]. Given the increased life expectancy and rapid ageing process in China during the two decades, it is reasonable to expect a dramatic increase rate of LTE prevalence through cumulative effects [49,50].

In the current study, we developed an estimate of the 1-year, 2-year and 5-year AE prevalence in China, which provides a basis for future studies, especially in the research of treatment gap for epilepsy in China [33,34]. However, because of the scarcity of studies that reported age- or gender-specific AE prevalence, we were not able to undertake some more detailed studies. Compared to the global 1-year AE prevalence (4.6‰ in 2000 and 4.5‰ in 2000) [15] in the Campaign Against Epilepsy Demonstration Project conducted in rural China, our study revealed a much lower 1-year AE prevalence of 3.79‰ in general population.

The magnitude of epilepsy burden in China, estimated in this study, represents a huge and significant health and socioeconomic burden [34,51]. With proper antiepileptic medication, up to 70% of epileptic seizures can be well controlled [52,53]. However, previous studies suggested that more than half of the individuals with epilepsy in China had never been treated with appropriate antiepileptic medicines

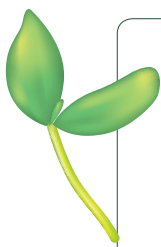
[15,39]. Many barriers may contribute to this situation. In China, due to the stigma-attached nature of the diagnosis of epilepsy, individuals with epilepsy are generally socially isolated and suffer in silence [15, 39]. Individuals with epilepsy are more likely to be under-educated or under-employed. A lack of knowledge about the nature of epilepsy and treatment may also influence the patients' personal health-seeking behaviours and compliance [15]. In addition, most individuals with epilepsy may be economically disadvantaged, which makes the free provision of antiepileptic medicines critically important for the management of epilepsy. However, this goal has not been universally achieved across the whole country, especially in resource-poor areas [52]. To make the situation worse, the lack of electroencephalogram and neuroimaging equipment, and personnel with neurologic expertise, may severely restrict the diagnosis of patients to a large extent.

Our study also had several potential limitations. First, our estimate of epilepsy prevalence was based on cross-sectional studies in the community. However, because of the uncertainty regarding each case definition, and the ratio between sensitivity and specificity of the diagnostic tools used, it has already been pointed that cross-sectional assessment may considerably underestimate the prevalence [44,54,55]. In addition, the initial suspicion on epilepsy in most of our included studies was established through the use of questionnaires or interviews. Although it was later also confirmed by neurologists, this investigatory approach may still be problematic due to recall bias and a high proportion of concealment [15,39]. Second, large variations were observed between the studies included in our systematic analysis. We tried to minimise these variations through strict inclusion and exclusion criteria. Still, variations in study methods, approaches to sampling, diagnostic criteria, availability of or access to appropriate treatment led to a considerable variation. Third, our modelled estimates were based on a limited number of covariates that were available in our included studies. Future attempts at evidence synthesis should include more of these covariates, such as the subtype of epilepsy, level of economic development and treatment gap. With these limitations in mind, the estimates presented in this study need to be interpreted cautiously.

The results of our meta-analysis have both academic and public health implications. An immense deficit in epidemiologic data regarding age-specific AE prevalence and LTE prevalence in China was identified. In particular, new studies should also adopt appropriate methods to reduce the variation in reported prevalence. Other important contributors include identifying subtypes of epilepsy and the current barriers in the society and health care systems. From the public health perspective, it is well-documented that most epilepsy cases can be prevented by effective measurements. Common preventable causes in children include infectious diseases, prenatal and perinatal central nervous system damage [2,16]. Birth asphyxia and febrile seizures are also well-documented risk factors for epilepsy. Programs targeted at reducing birth asphyxia and timely treatment of febrile convulsions can also contribute to the reduction of epilepsy burden [56,57]. In adults, the most prevalent causes of epilepsy include head injury, intracranial infection and cerebrovascular diseases, which account for 88.5% of all the adult-onset epilepsy cases [58,59], highlighting the importance of preventing injury.

## CONCLUSIONS

Our study provides a comprehensive evidence synthesis of LTE and AE prevalence in China to date. LTE prevalence is the highest in individuals aged 30–34 years. The burden of LTE in China has increased dramatically between 1990 and 2015, when the prevalence of LTE has more than doubled, and the number of people with LTE more than tripled. The large amount of AE cases in China calls for optimal management and treatment. More high-quality epidemiological studies on LTE and AE prevalence are still needed.



**Funding:** None.

**Authorship contribution:** KYC, IR and WW designed the study. PS, YZ, XY, KYC and AP conducted literature searches in Chinese and English databases. PS designed the analyses and analysed the data. PS and IR wrote the final draft of the paper. JW and AD provided important intellectual content to multiple drafts of the paper.

**Competing interests:** IR is co-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 organisations. The author completed the Unified Competing Interest declare no competing interests.

- 1 Fisher RS, Boas WvE, Blume W, Elger C, Genton P, Lee P, et al. Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia*. 2005;46:470-2. Medline:15816939 doi:10.1111/j.0013-9580.2005.66104.x
- 2 Duncan JS, Sander JW, Sisodiya SM, Walker MC. Adult epilepsy. *Lancet*. 2006;367:1087-100. Medline:16581409 doi:10.1016/S0140-6736(06)68477-8
- 3 Kwan P, Brodie MJ. Neuropsychological effects of epilepsy and antiepileptic drugs. *Lancet*. 2001;357:216-22. Medline:11213111 doi:10.1016/S0140-6736(00)03600-X
- 4 World Health Organization. International Classification of Functioning, Disability and Health. Geneva; World Health Organization: 2001.
- 5 Thurman DJ, Beghi E, Begley CE, Berg AT, Buchhalter JR, Ding D, et al. Standards for epidemiologic studies and surveillance of epilepsy. *Epilepsia*. 2011;52:2-26. Medline:21899536 doi:10.1111/j.1528-1167.2011.03121.x
- 6 Tellez-Zenteno JF, Patten SB, Jetté N, Williams J, Wiebe S. Psychiatric comorbidity in epilepsy: a population-based analysis. *Epilepsia*. 2007;48:2336-44. Medline:17662062
- 7 Hermann B, Seidenberg M, Jones J. The neurobehavioural comorbidities of epilepsy: can a natural history be developed? *Lancet Neurol*. 2008;7:151-60. Medline:18207113 doi:10.1016/S1474-4422(08)70018-8
- 8 Baker GA. The psychosocial burden of epilepsy. *Epilepsia*. 2002;43:26-30. Medline:12190975 doi:10.1046/j.1528-1157.43.s.6.12.x
- 9 Leonardi M, Ustun TB. The global burden of epilepsy. *Epilepsia*. 2002;43:21-5. Medline:12190974 doi:10.1046/j.1528-1157.43.s.6.11.x
- 10 Ngugi AK, Bottomley C, Kleinschmidt I, Sander JW, Newton CR. Estimation of the burden of active and life-time epilepsy: a meta-analytic approach. *Epilepsia*. 2010;51:883-90. Medline:20067507 doi:10.1111/j.1528-1167.2009.02481.x
- 11 Moshé SL, Perucca E, Ryvlin P, Tomson T. Epilepsy: new advances. *Lancet*. 2015;385:884-98. Medline:25260236 doi:10.1016/S0140-6736(14)60456-6
- 12 Neligan A, Bell GS. The mortality of epilepsy. *Epilepsy*. 2009.
- 13 Murray CJ. Quantifying the burden of disease: the technical basis for disability-adjusted life years. *Bull World Health Organ*. 1994;72:429. Medline:8062401
- 14 Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012;380:2197-223. Medline:23245608 doi:10.1016/S0140-6736(12)61689-4
- 15 World Health Organization. Epilepsy management at primary health level in rural China: WHO/ILAE/IBE-a Global Campaign Against Epilepsy Demonstration Project. Geneva: WHO; 2009.
- 16 Newton CR, Garcia HH. Epilepsy in poor regions of the world. *Lancet*. 2012;380:1193-201. Medline:23021288 doi:10.1016/S0140-6736(12)61381-6
- 17 Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet*. 2006;367:1747-57. Medline:16731270 doi:10.1016/S0140-6736(06)68770-9
- 18 Picot MC, Baldy-Moulinier M, Daurčs JP, Dujols P, Crespel A. The prevalence of epilepsy and pharmacoresistant epilepsy in adults: A population-based study in a Western European country. *Epilepsia*. 2008;49:1230-8. Medline:18363709 doi:10.1111/j.1528-1167.2008.01579.x
- 19 Gu L, Liang B, Chen Q, Long J, Xie J, Wu G, et al. Prevalence of epilepsy in the People's Republic of China: a systematic review. *Epilepsy Res*. 2013;105:195-205. Medline:23507331 doi:10.1016/j.epilepsyres.2013.02.002
- 20 Sayehmiri K, Tavan H, Sayehmiri F, Mohammadi I, Carson KV. Prevalence of epilepsy in Iran: a meta-analysis and systematic review. *Iran J Child Neurol*. 2014;8:9. Medline:25657765
- 21 Forsgren L, Beghi E, Oun A, Sillanpää M. The epidemiology of epilepsy in Europe—a systematic review. *Eur J Neurol*. 2005;12:245-53. Medline:15804240 doi:10.1111/j.1468-1331.2004.00992.x
- 22 Benamer HT, Grosset DG. A systematic review of the epidemiology of epilepsy in Arab countries. *Epilepsia*. 2009;50:2301-4. Medline:19389149 doi:10.1111/j.1528-1167.2009.02058.x
- 23 Mac TL, Tran D-S, Quet F, Odermatt P, Preux P-M, Tan CT. Epidemiology, aetiology, and clinical management of epilepsy in Asia: a systematic review. *Lancet Neurol*. 2007;6:533-43. Medline:17509488 doi:10.1016/S1474-4422(07)70127-8
- 24 Fiest KM, Sauro KM, Wiebe S, Patten SB, Kwon C-S, Dykeman J, et al. Prevalence and incidence of epilepsy: A systematic review and meta-analysis of international studies. *Neurology*. 2017;88:296-303. Medline:27986877 doi:10.1212/WNL.0000000000003509
- 25 Banerjee PN, Filippi D, Hauser WA. The descriptive epidemiology of epilepsy—a review. *Epilepsy Res*. 2009;85:31-45. Medline:19369037 doi:10.1016/j.epilepsyres.2009.03.003
- 26 de Boer HM, Mula M, Sander JW. The global burden and stigma of epilepsy. *Epilepsy Behav*. 2008;12:540-6. Medline:18280210 doi:10.1016/j.yebeh.2007.12.019
- 27 Yang G, Wang Y, Zeng Y, Gao GF, Liang X, Zhou M, et al. Rapid health transition in China, 1990–2010: findings from the Global Burden of Disease Study 2010. *Lancet*. 2013;381:1987-2015. Medline:23746901 doi:10.1016/S0140-6736(13)61097-1
- 28 Fung IC. Chinese journals: a guide for epidemiologists. *Emerg Themes Epidemiol*. 2008;5:20. Medline:18826604 doi:10.1186/1742-7622-5-20
- 29 Xia J, Wright J, Adams CE. Five large Chinese biomedical bibliographic databases: accessibility and coverage. *Health Info Libr J*. 2008;25:55-61. Medline:18251914 doi:10.1111/j.1471-1842.2007.00734.x

- 30 Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 2009;6:e1000097. Medline:19621072 doi:10.1371/journal.pmed.1000097
- 31 Stevens GA, Alkema L, Black RE, Boerma JT, Collins GS, Ezzati M, et al. Guidelines for accurate and transparent health estimates reporting: the GATHER statement. *PLoS Med.* 2016;13:e1002056. Medline:27351744 doi:10.1371/journal.pmed.1002056
- 32 MacDonald BK, Cockerell O, Sander J, Shorvon S. The incidence and lifetime prevalence of neurological disorders in a prospective community-based study in the UK. *Brain.* 2000;123:665-76. Medline:10733998 doi:10.1093/brain/123.4.665
- 33 Meyer A-C, Dua T, Ma J, Saxena S, Birbeck G. Global disparities in the epilepsy treatment gap: a systematic review. *Bull World Health Organ.* 2010;88:260-6. Medline:20431789 doi:10.2471/BLT.09.064147
- 34 Wang WZ, Wu J, Wang D, Dai X, Yang B, Wang T, et al. The prevalence and treatment gap in epilepsy in China: An ILAE/IBEW/WHO study. *Neurology.* 2003;60:1544-5. Medline:12743252 doi:10.1212/01.WNL.0000059867.35547.DE
- 35 Hox JJ, Moerbeek M, van de Schoot R. *Multilevel analysis: Techniques and applications.* Routledge; 2010.
- 36 Higgins JP, Green S. *Cochrane handbook for systematic reviews of interventions.* Vol 4: Oxford: John Wiley & Sons; 2011.
- 37 United Nations, Department of Economic and Social Affairs, Population Division. *World Population Prospects, the 2015 Revision.* 2015. Available: <https://esa.un.org/unpd/wpp/>. Accessed: 8 January 2017.
- 38 Ngugi AK, Kariuki S, Bottomley C, Kleinschmidt I, Sander J, Newton C. Incidence of epilepsy A systematic review and meta-analysis. *Neurology.* 2011;77:1005-12. Medline:21893672 doi:10.1212/WNL.0b013e31822cfc90
- 39 Kleinman A, Wang W-Z, Li S-C, Cheng X-M, Dai X-Y, Li K-T, et al. The social course of epilepsy: chronic illness as social experience in interior China. *Soc Sci Med.* 1995;40:1319-30. Medline:7638642 doi:10.1016/0277-9536(94)00254-Q
- 40 Sander JW. The epidemiology of epilepsy revisited. *Curr Opin Neurol.* 2003;16:165-70. Medline:12644744 doi:10.1097/00019052-200304000-00008
- 41 Kotsopoulos IA, Van Merode T, Kessels FG, De Krom MC, Knottnerus JA. Systematic review and meta-analysis of incidence studies of epilepsy and unprovoked seizures. *Epilepsia.* 2002;43:1402-9. Medline:12423392 doi:10.1046/j.1528-1157.2002.t01-1-26901.x
- 42 Mu J, Liu L, Zhang Q, Si Y, Hu J, Fang J, et al. Causes of death among people with convulsive epilepsy in rural West China: A prospective study. *Neurology.* 2011;77:132-7. Medline:21653888 doi:10.1212/WNL.0b013e318223c784
- 43 Ding D, Wang W, Wu J, Ma G, Dai X, Yang B, et al. Premature mortality in people with epilepsy in rural China: a prospective study. *Lancet Neurol.* 2006;5:823-7. Medline:16987728 doi:10.1016/S1474-4422(06)70528-2
- 44 Streiner DL, Patten SB, Anthony JC, Cairney J. Has 'lifetime prevalence' reached the end of its life? An examination of the concept. *Int J Methods Psychiatr Res.* 2009;18:221-8. Medline:20052690 doi:10.1002/mpr.296
- 45 Savic I. Sex differences in human epilepsy. *Exp Neurol.* 2014;259:38-43. Medline:24747359 doi:10.1016/j.expneurol.2014.04.009
- 46 Christensen J, Kjeldsen MJ, Andersen H, Friis ML, Sidenius P. Gender differences in epilepsy. *Epilepsia.* 2005;46:956-60. Medline:15946339 doi:10.1111/j.1528-1167.2005.51204.x
- 47 Burneo JG, Tellez-Zenteno J, Wiebe S. Understanding the burden of epilepsy in Latin America: a systematic review of its prevalence and incidence. *Epilepsy Res.* 2005;66:63-74. Medline:16125900 doi:10.1016/j.epilepsyres.2005.07.002
- 48 McHugh JC, Delanty N. Epidemiology and classification of epilepsy: gender comparisons. *Int Rev Neurobiol.* 2008;83:11-26. Medline:18929074 doi:10.1016/S0074-7742(08)00002-0
- 49 Chen S, Powell JL. *Aging in China: Implications to social policy of a changing economic state.* Vol 2: Springer Science & Business Media; 2012.
- 50 Zhang NJ, Guo M, Zheng X. China: awakening giant developing solutions to population aging. *Gerontologist.* 2012;52:589-96. Medline:22936537 doi:10.1093/geront/gns105
- 51 Ding D, Hong Z, Wang WZ, Wu JZ, De Boer HM, Prilipko L, et al. Assessing the disease burden due to epilepsy by disability adjusted life year in rural China. *Epilepsia.* 2006;47:2032-7. Medline:17201700 doi:10.1111/j.1528-1167.2006.00802.x
- 52 Scott RA, Lhatoo SD, Sander JW. The treatment of epilepsy in developing countries: where do we go from here? *Bull World Health Organ.* 2001;79:344-51. Medline:11357214
- 53 Sander JW. The use of antiepileptic drugs—principles and practice. *Epilepsia.* 2004;45:28-34. Medline:15315513 doi:10.1111/j.0013-9580.2004.455005.x
- 54 Moffitt TE, Caspi A, Taylor A, Kokaua J, Milne B, Polanczyk G, et al. How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. *Psychol Med.* 2010;40:899-909. Medline:19719899 doi:10.1017/S0033291709991036
- 55 Cockerell OC, Eckle I, Goodridge D, Sander J, Shorvon S. Epilepsy in a population of 6000 re-examined: secular trends in first attendance rates, prevalence, and prognosis. *J Neurol Neurosurg Psychiatry.* 1995;58:570-6. Medline:7745404 doi:10.1136/jnnp.58.5.570
- 56 Chawla S, Aneja S, Kashyap R, Mallika V. Etiology and clinical predictors of intractable epilepsy. *Pediatr Neurol.* 2002;27:186-91. Medline:12393128 doi:10.1016/S0887-8994(02)00416-2
- 57 Singhi P, Jagirdar S, Khandelwal N, Malhi P. Epilepsy in children with cerebral palsy. *J Child Neurol.* 2003;18:174-9. Medline:12731642 doi:10.1177/08830738030180030601
- 58 Chung T-T, Lin C-Y, Huang W-Y, Lin C-L, Sung F-C, Kao C-H. Risks of subsequent epilepsy among patients with hypertensive encephalopathy: a nationwide population-based study. *Epilepsy Behav.* 2013;29:374-8. Medline:24090775 doi:10.1016/j.yebeh.2013.08.013
- 59 Christensen J, Pedersen MG, Pedersen CB, Sidenius P, Olsen J, Vestergaard M. Long-term risk of epilepsy after traumatic brain injury in children and young adults: a population-based cohort study. *Lancet.* 2009;373:1105-10. Medline:19233461 doi:10.1016/S0140-6736(09)60214-2



# Consensus-based approach to develop a measurement framework and identify a core set of indicators to track implementation and progress towards effective coverage of facility-based Kangaroo Mother Care

Tanya Guenther<sup>1</sup>, Sarah Moxon<sup>2</sup>, Bina Valsangkar<sup>1</sup>, Greta Wetzel<sup>1</sup>, Juan Ruiz<sup>3,4</sup>, Kate Kerber<sup>1</sup>, Hannah Blencowe<sup>2</sup>, Queen Dube<sup>5</sup>, Shashi N Vani<sup>6</sup>, Donna Vivio<sup>7</sup>, Hema Magge<sup>8,9</sup>, Socorro De Leon-Mendoza<sup>10</sup>, Janna Patterson<sup>11</sup>, Goldy Mazia<sup>12,13</sup>

<sup>1</sup> Save the Children, Washington, D.C., USA

<sup>2</sup> London School of Hygiene and Tropical Medicine, London, UK

<sup>3</sup> Kangaroo Mother Care Foundation, Bogota, Colombia

<sup>4</sup> Herbert Wertheim College of Medicine, FIU, Miami, Florida, USA

<sup>5</sup> Queen Elizabeth Central Hospital, Blantyre, Malawi

<sup>6</sup> Kangaroo Mother Care Foundation, Ahmedabad, India

<sup>7</sup> USAID, Washington, D.C., USA

<sup>8</sup> Partners in Health, Kigali, Rwanda

<sup>9</sup> Brigham and Women's Hospital, Division of Global Health Equity, Boston, Massachusetts, USA

<sup>10</sup> Bless-Tetada Kangaroo Mother Care Foundation, Manila, Philippines

<sup>11</sup> Bill & Melinda Gates Foundation, Seattle WA, USA

<sup>12</sup> PATH, Washington, D.C., USA

<sup>13</sup> Maternal and Child Survival Programme, Washington, D.C., USA

## Correspondence to:

Tanya Guenther  
Save the Children  
899 North Capitol Street NE, Suite 900  
Washington, D.C. 20002  
USA  
tguenther@savechildren.org

**Background** As efforts to scale up the delivery of Kangaroo Mother Care (KMC) in facilities are increasing, a standardized approach to measure implementation and progress towards effective coverage is needed. Here, we describe a consensus-based approach to develop a measurement framework and identify a core set of indicators for monitoring facility-based KMC that would be feasible to measure within existing systems.

**Methods** The KMC measurement framework and core list of indicators were developed through: 1) scoping exercise to identify potential indicators through literature review and requests from researchers and program implementers; and 2) face-to-face consultations with KMC and measurement experts working at country and global levels to review candidate indicators and finalize selection and definitions.

**Results** The KMC measurement framework includes two main components: 1) service readiness, based on the WHO building blocks framework; and 2) service delivery action sequence covering identification, service initiation, continuation to discharge, and follow-up to graduation. Consensus was reached on 10 core indicators for KMC, which were organized according to the measurement framework. We identified 4 service readiness indicators, capturing national level policy for KMC, availability of KMC indicators in HMIS, costed operational plans for KMC and availability of KMC services at health facilities with inpatient maternity services. Six indicators were defined for service delivery, including weighing of babies at birth, identification of those  $\leq 2000$  g, initiation of facility-based KMC, monitoring the quality of KMC, status of babies at discharge from the facility and levels of follow-up (according to country-specific protocol).

**Conclusions** These core KMC indicators, identified with input from a wide range of global and country-level KMC and measurement experts, can aid efforts to strengthen monitoring systems and facilitate global tracking of KMC implementation. As data collection systems advance, we encourage program managers and evaluators to document their experiences using this framework to measure progress and allow indicator refinement, with the overall aim of working towards sustainable, country-led data systems.

An estimated 15 million babies are born prematurely each year, accounting for about 1 in 10 births worldwide [1]. Preterm birth, defined as birth before 37 completed weeks of gestation, is the leading direct cause of newborn mortality and morbidity [2,3]. Complications of prematurity are the primary cause of child death worldwide and also a risk factor for neonatal deaths from other causes, especially infections [3]. Globally, preterm birth complications contribute 3% of disability-adjusted life years (DALYs) for all ages and account for 38% of DALYs attributed to neonatal conditions [4]. The burden of mortality and morbidity due to preterm birth is heavily concentrated in south Asia and sub-Saharan Africa, where more than 60% of preterm births take place and health systems face multiple challenges to deliver high quality care [1,2,5].

The Every Newborn Action Plan (ENAP), a global multi-partner movement with the goal of ending preventable newborn deaths has set national targets of  $\leq 12$  neonatal deaths per 1000 live births by 2030 [6]. As the leading cause of newborn deaths, a focus on preterm birth and the associated complications are essential to achieving these ambitious goals. There are evidence-based, cost-effective interventions to prevent preterm birth and manage complications. As part of the evidence base for the ENAP, an epidemiological analysis estimated that up to 70% of preterm deaths could be averted through the provision of quality inpatient care [7]. Kangaroo Mother Care (KMC) is a critical part of inpatient care for preterm newborns, and also provides the foundation for improved outpatient and follow-up care of small babies [5,8]. In July 2015, the World Health Organization (WHO) released guidelines on interventions to improve preterm birth outcomes, which strongly recommend KMC for the routine care of neonates born weighing  $\leq 2000$  g as soon as they are clinically stable [9]. Birthweight is used as an indication for KMC initiation and a proxy for preterm birth given the challenges of accurate gestational age measurement in many low-resource settings.

Kangaroo Mother Care is defined by WHO as early, continuous and prolonged skin-to-skin contact between the mother (or other caregiver) and the baby, and exclusive breastfeeding (ideally) or feeding with expressed breastmilk [9]. Provision of KMC is embedded within a broader package of inpatient care for premature babies that involves supportive care (eg, infection prevention and management, respiratory support, etc.) and requires referral for higher level care when necessary and ongoing follow-up post-discharge [8]. In some more developed settings (eg, certain Latin American countries), KMC may be initiated at the facility and continued on an ambulatory basis with mothers returning to the facility frequently (as needed). Such an approach is only feasible in settings where health facilities are easily accessible and the appropriate infrastructure is in place. Studies show that continuous KMC implemented at health facilities can prevent up to 50% of deaths among babies  $\leq 2000$  g [10]. The practice of facility-based KMC also offers benefits beyond reduced mortality; compared with conventional neonatal care for small babies (incubator care), KMC reduces infections, hypothermia, and length of hospital stay and improves breastfeeding, weight gain and maternal-infant bonding [10]. Intermittent KMC, as tolerated, is increasingly being used for babies that are less stable to support clinical and developmental outcomes [9].

Despite the strong evidence base for KMC, progress in taking KMC implementation to scale has been slow [5,11,12]. While more than half of the 75 Countdown to 2015 countries report national policies recommending KMC, availability of KMC services is limited to a small number of central or teaching hospitals in all but a handful of countries ([13] and our unpublished results). A multi-country assessment of health systems bottlenecks to scale up of KMC in 12 African and Asian countries found that health financing, community ownership and partnership, health service delivery, leadership and governance and health workforce were perceived as major or significant barriers by nine or more countries [5]. One of the cross-cutting challenges underpinning these barriers was effective information systems and data on KMC coverage and quality [5,12].

In an effort to accelerate and support the uptake of KMC, the Bill & Melinda Gates Foundation and partners released a call to action in 2013 for the global adoption of facility-based KMC and formed the KMC Acceleration Partnership (KAP) [11]. The call to action set an ambitious target of 50% coverage of KMC among preterm newborns by 2020 and emphasized the importance of measuring progress using robust metrics and indicators [11]. Similarly, to meet its ambitious goals, ENAP recognized the critical need for improved data on preterm, small and sick newborns to support the scale up of high impact interventions [6]. At the time of its launch in 2014, ENAP published a core set of indicators needed for tracking progress in reaching their goals [6]. Coverage of KMC is one of the core ENAP indicators, and also one of the indicators with some of the greatest identified data gaps [6,14]. At the time, there was no existing definition for a KMC coverage indicator. To achieve scale up of KMC, there is a need for consensus on a common set of indicators to track KMC implementation and progress to effective coverage. The ENAP metrics

stream, therefore, prioritized work on defining and testing a measureable coverage indicator, but also emphasized the importance of developing process indicators to track content and quality. In conjunction with ENAP, the KAP initiated a consensus-based process to identify a core set of standardized indicators for KMC to facilitate country and global monitoring and evaluation of KMC efforts and inform the integration of data on KMC into national health management information systems (HMIS). Regular monitoring and reporting of these indicators will strengthen the global evidence base for KMC and inform approaches to strengthen scale-up of KMC [11,12]. Further, careful facility-level measurement of KMC service delivery is important for improving the quality of KMC services and can help avoid the phenomenon of “empty” scale-up.

The purpose of this paper is to describe the approach to develop a measurement framework and select and refine a set of indicators for monitoring implementation of facility-based KMC. The aim was to develop a focused list of indicators that would be relevant across settings and could be measured within existing health systems at scale. The challenges to establish an appropriate denominator for measuring coverage of KMC and options for testing are also discussed.

## METHODS

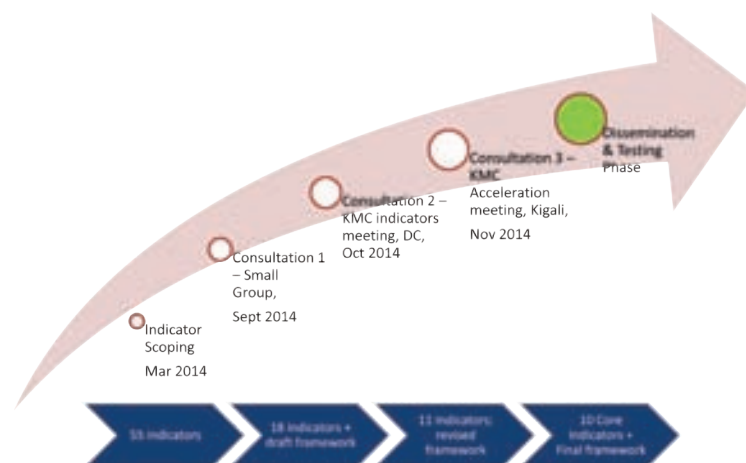
The KMC measurement framework and core list of indicators were developed through: 1) scoping exercise to identify potential indicators; and 2) face-to-face consultations with measurement and KMC experts to review candidate indicators and finalize selection (**Figure 1**).

### Scoping exercise to identify potential indicators

An initial list of candidate indicators was developed through a review of the grey literature (program documents and surveys) and consultations with KMC and measurement experts. We circulated a request for existing KMC indicators to members of the Newborn Indicators Technical Working Group (NITWG), an inter-agency working group convened by Saving Newborn Lives, and the KMC Acceleration Partnership. A total of 79 candidate indicators and data elements were extracted and summarized in an excel spreadsheet. This list was refined to 55 through sorting and removal of duplicates and organized by a standard results framework (impact, coverage, access, quality, demand, policy/enabling environment). The main sources of indicators in this initial list included the Fundacion Cangaru, Maternal and Child Health Integrated Program (MCHIP) KMC implementation Guide, various facility and household surveys conducted by programs implementing KMC (SNL Malawi Facility Assessment and household survey, Uganda Newborn Study (UNEST) survey, Ghana Newhints survey, Ethiopia household survey, South Africa Facility Assessment) and the Malawi HMIS.

### Consultations to review and finalize indicator selection

A series of face-to-face meetings were convened with KMC and measurement experts working at country and global level over a three month period. A full list of participants, their affiliations and area of expertise (measurement, KMC or both) are included in Table S1 in **Online Supplementary Document**.



**Figure 1.** Overview of Kangaroo Mother Care (KMC) framework development and indicator selection process.

## Initial scoring and development of a measurement framework

A small group meeting with 12 members of the KAP and the NITWG was held September 5, 2014 to review and score the raw list of 55 indicators. A focused set of five criterion for initial scoring of the indicators was developed, which reflects commonly applied indicator selection criteria: feasibility (data can be collected with reasonable and affordable effort in low resource settings), reliability (data can be collected consistently over time), usefulness for decision-making (data are relevant and will help guide KMC programming), sensitivity (responsive to change), and specificity (focused on specific aspect, not overly broad) [15–17]. The group broke into smaller groups for in-depth discussion and scored each candidate indicator as high, medium, or low for each of the five criteria. The group recommended that a measurement framework specific to KMC should be developed to better organize the indicators and assist with prioritizing selection. Following the meeting, a core team representing the KAP, NITWG and ENAP metrics stream extracted the strongest indicators based on the scoring criteria for further development (preparation of full definitions, data source, and methods) and drafted a KMC measurement framework. The resulting 18 candidate indicators were then organized according to the draft framework.

## Refinement of framework and indicator list

In October, consultations were held with a broader group of newborn programmatic and measurement experts on October 6–7, 2014 in Washington DC. The 24 attendees, representing implementing agencies, donors and researchers, formed three small groups (national level/service readiness; facility-level/service delivery; and coverage) to review the KMC framework and each of the 18 candidate indicators in detail. Groups were provided with a series of questions for each indicator to guide their discussion and decision-making process. The group made recommendations about which indicators to retain, which to drop and areas for further research; further details on the discussion and outputs is available in the meeting report [18]. Following the meeting, the core team consolidated feedback and updated the measurement framework and refined the indicator list down to 11 candidate indicators. A smaller task team was delegated to work specifically on defining a feasible coverage indicator for KMC that could be tested as part of the ENAP metrics measurement improvement plan. This task team, alongside the core group, also undertook a preliminary mapping exercise to see what data were available, with a focus on assessing denominator options for generating a potential coverage indicator for KMC that could be tested as part of the ENAP measurement improvement plan.

## Finalization of framework and indicators

The final consultation took place on November 15, 2014 in Kigali, Rwanda as part of the KMC Acceleration meeting and focused on country level input. Eighteen participants, including individuals supporting KMC implementation in nine countries (Bangladesh, Malawi, Nigeria, Rwanda, India, Indonesia, Philippines, Uganda and South Africa) gave feedback on the measurement framework and reviewed each candidate indicator to assess availability, feasibility and usefulness considering their country context. Participants were split into two groups. One group worked specifically on the coverage indicator, and the other group focused on readiness indicators and facility level data for tracking service delivery and quality of care. In each group, a presentation was made to provide an overview of progress to date, review each indicator in detail and identify priority areas for discussion. Participants in the service readiness and facility data group were asked to use post-it notes to record information on availability/data source, data users, collection methods, and challenges for each indicator in their setting and then vote whether the indicator was ready to go, needed more work/unsure or should be dropped. Participants in the coverage group, reviewed the work carried out by the ENAP metrics KMC task team (**Box 1**) and discussed a feasible a measurable coverage indicator. In view of the challenges in measuring a denominator, the group reached consensus through placing individual votes between use of <2500 g, total facility births or estimated live births. Based on the feedback, the core team finalized the framework and list of indicators.

## RESULTS

### KMC measurement framework

**Figure 2** shows the KMC measurement framework, which was developed to guide the identification and prioritization of core indicators. The framework includes two main components: 1) KMC service readiness and 2) action sequence of KMC service delivery. The seven service readiness elements are based on the WHO

**Box 1.** The challenge of measuring KMC coverage**What is coverage measurement and why is it challenging for KMC?**

A coverage indicator aims to measure the number of individuals that receive a specific intervention or treatment within a given population in need of the intervention. The numerator is measured as the total number of individuals that received the intervention and the denominator is the total population, usually those that could have benefitted from that specific intervention or treatment. For KMC, neither the numerator nor the denominator are easy to define or measure. KMC is not a one-off contact with the health system; many of the components of KMC are processes (eg, continuous skin-to-skin contact, follow up care). And measurement of specific interventions is a challenge when only by a small group or sub population benefit from that intervention. Defining whether or not an infant could benefit from KMC requires a level of clinical judgement and more precise metrics than those reported by most routine information systems in LMIC.

**The ENAP metrics KMC task team**

ENAP metrics assembled a KMC task team with experts in measurement and programme implementation drawing on expertise from the KMC acceleration partnership and wider groups. Different numerators and denominators were proposed and discussed based on their definition and the feasibility of measurement.

**Numerator challenges**

The evidence base for mortality impact of KMC is currently for infants weighing 2000 g or less. However, in some low and middle income countries where programmes have been extended, eligibility criteria for entry to KMC may be for babies up to 2500 g. Coverage of most maternal and newborn interventions in many settings is still measured through household surveys and relies on maternal recall up to five years after the birth in question. Even though mothers can accurately recall KMC, even years after the event, the sample size needed to gather representative data through a household survey may be prohibitively large<sup>15</sup>. Typically, facility based assessments capture information on infrastructure, processes and service readiness, and are best suited to measure the number of facilities that are prepared to provide components of the service (eg, sufficient trained staff, space, and equipment). In most settings, the number of newborns initiated on facility-based KMC is measured either through hospital admission or care records, but currently these data are rarely reported into national health information systems.

**Denominator challenges**

The denominator was the most technically challenging and a list of options were proposed. A large proportion of newborns do not have their weight recorded at birth and even where birthweight is recorded, there is a known tendency for “heaping” of data, especially at measures of 2500 g and 2000 g. Given the difficulty in accurately capturing all those babies in need of KMC, especially through existing data collection systems, using total live births as the denominator to give a proxy was considered. This has been done with other interventions where the aim is not for 100% coverage, such as C-section, to generate a rate that is benchmarked against a target threshold. Recent estimates suggest a variation in preterm birth rates of between 4–18% of total live births in different countries. This means that the KMC rate in each country may indicate a different unmet need and target thresholds would need to vary between settings to reflect these differences as well as variation in numbers of full-term LBW and pre-term babies. As an important limitation, if total live births is used as a denominator, it does not reflect whether the babies that received it were drawn from the population that could have benefitted from KMC.

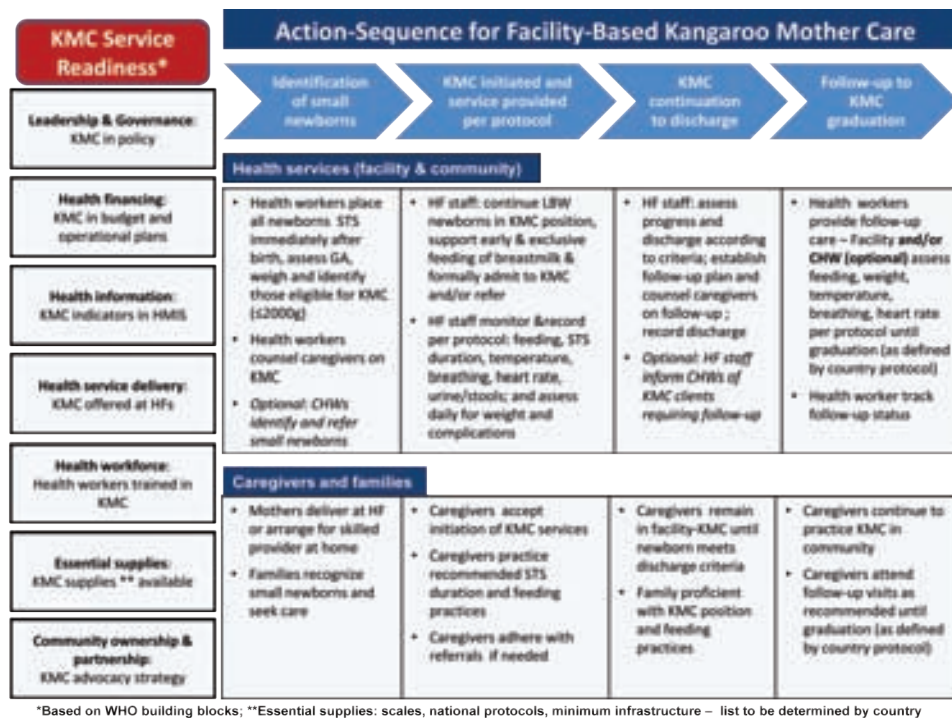
**Proposed indicator**

The ENAP KMC task team established that it is not possible to capture all of the components of KMC in one coverage indicator as many of these refer to processes that happen over a period of time. Household surveys are unlikely to be a feasible approach to measure KMC coverage and increasingly, health facility assessments are starting to measure key components of KMC care. Of all the available options, **the number of newborns initiated on facility based KMC** gives a representation of the number of newborns initiating the care. Task teams agreed the indicators would need rigorous testing for validity and feasibility with a variety of different denominator options including, **live births in the facility, estimated live births and eventually target population for coverage (total number of newborns  $\leq$ 2000 g)**.

As a preliminary exercise, the task team approached a select few LMIC countries for data on the KMC numerator, which is available through a limited number of HMIS and many hospital registration systems. To demonstrate the numerator with different denominator options, task teams present three graphs showing the proposed numerator over total reported live births, total reported live births <2500 g and estimated live births for two countries, Malawi and Dominican Republic (**Figure 4**).

**What are the next steps?**

As national facility based data and health information systems become more advanced, the ideal is to develop more precise indicators, but these are not currently available in most of the countries where the unmet need for KMC is arguably the greatest and there are the most data gaps. It is critical to improve the recording and reporting of birth weight in facilities. Given the importance of prematurity as a direct cause of death and as a risk factor for morbidities and death from other causes (eg, infections), developing simplified tools for measuring gestational age is critical to plan for programmes, to improve the evidence base and to develop more precise indicators of unmet need. If such data were available in more settings, indicators based on specific weight or gestational age criteria could be measured. Existing data sets from countries with established KMC programmes and accurate assessment of gestational age and birthweight could be used for testing the denominators and proposed numerators. The ENAP metrics measurement improvement plan has a five year plan set out to test the validity and feasibility of a number of numerator and denominator options for all the ENAP core indicators with the objective of institutionalizing a KMC coverage indicator in global accountability mechanisms by 2020.



**Figure 2.** Measurement framework for Kangaroo Mother Care (KMC).

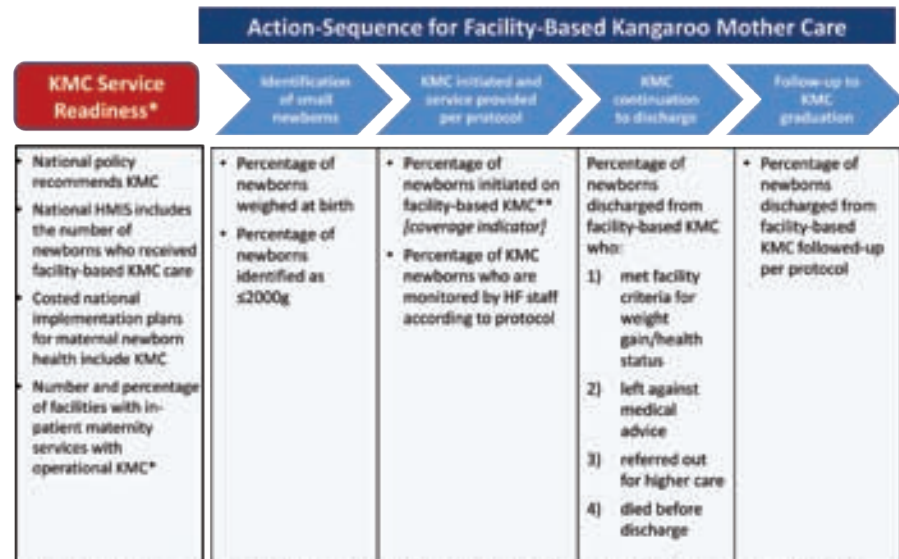
building blocks framework, and specify what minimum elements should be in place to support national-level implementation of KMC [19]. The action sequence of service delivery outlines four main steps necessary for provision of KMC at health facilities: identification of small babies; KMC initiation per protocol; KMC continuation to discharge; and follow-up to KMC graduation. Essential actions for health service providers and for caregivers and families are outlined in broad terms for each step in the action sequence.

### KMC core indicators

The 10 KMC core indicators are summarized in **Figure 3** according to the framework and defined in **Table 1**. Table S2 in **Online Supplementary Document** provides further information on the limitations and additional data collection considerations for each indicator.

### Service readiness indicators

Consensus was reached on four indicators of service readiness, namely national level policy for KMC, availability of KMC indicators in HMIS, costed operational plans for KMC and availability of KMC services meeting national minimum standards at facilities with inpatient maternity services. The service readiness indicators are qualitative milestones or benchmark indicators, which build on existing indicators and data collection efforts at national/global level, such as the Countdown to 2015 policy indicators. Primary data sources are Ministry of Health and implementing partners and the data can be aggregated at global level to track progress across countries. Availability of data are mixed. The indicator on national level policy for KMC was added in 2014 to the set of policy indicators tracked in Countdown to 2015 and data are available for two years (2014 and 2015), but there are no clear plans for tracking post 2015 [14]. The other three indicators are not currently tracked and would need to be collected in coordination with countries. Limited data on the availability of KMC services at health facilities are currently captured through several large facility assessment tools, including the Demographic and Health Survey Program's (ICF International) Service Provision Assessment (SPA) and the WHO's Service Availability and Readiness Assessment (SARA) and the revised Emergency Obstetrics and Newborn Care (EmONC) assessment tools [20]. However, these assessments are conducted infrequently, and countries investing in scaling up KMC services may need to establish other more frequent mechanisms to track KMC service availability, such as supervision or periodic audits. Further work is also needed to define 'operational' KMC, but at minimum it should specify availability of trained staff, space, and supplies. Defining these minimum components was also considered critical to the ENAP metrics measurement improvement plan, which recommended tracking the availability of KMC services as a process indicator [14].



NOTES: Service readiness indicators are meant primarily for national level tracking; Service delivery indicators (under the action sequence) are meant primarily for facility and district level use to monitor implementation and identify areas for improvement (selected indicators can be aggregated for national reporting); \* Proposed ENAP metrics process indicator; \*\* proposed ENAP metrics coverage indicator and recommended for inclusion in national HMIS; both ENAP metrics indicators prioritized for further testing

**Figure 3.** Summary of core Kangaroo Mother Care (KMC) indicators according to framework.

### Service delivery indicators

Six indicators were identified for service delivery, including weighing of babies at birth, identification of those  $\leq 2000$  g, initiation of facility-based KMC (coverage indicator), monitoring of KMC for quality, status of babies at discharge from KMC and level of follow-up according to national protocol. The service delivery indicators focus on capturing service utilization and elements of quality (both process and outcomes) and are intended for use primarily at the facility and district level for assessing KMC implementation and identifying program improvement needs. As such, the primary data sources are routine facility records and the list of recommended indicators was kept as short as possible to minimize burden on health staff.

Two indicators focus on identification of babies eligible for KMC. The percentage of newborns weighed at birth is captured by both the Demographic Health Survey (DHS) and the Multiple Indicator Cluster Survey (MICS), but only every few years. Weighing of babies at birth can also be estimated on a more routine basis using facility labour and delivery (L&D) records, although in most cases only facility births would be included and quality of recording is often poor. The DHS and MICS household surveys also estimate the percentage of low birth-weight babies, defined as birth weight  $< 2500$  g. However, as the evidence base for KMC is for babies weighing  $\leq 2000$  g at birth, the recommended indicator reinforces this cut-off and encourages routine data collection through facility L&D records. Most countries have space to record actual birth weight, but to our knowledge very few high burden countries currently collate and report on the number of live births weighing  $\leq 2000$  g. In addition to improving measurement of birth weight, investment in approaches to strengthen assessment of gestational age during antenatal care and at delivery are necessary to better target KMC interventions towards who would benefit most as the proportion of LBW babies that are pre-term will vary by setting.

The percentage of babies initiated on facility-based KMC was identified as a coverage indicator. Defining the denominator for a KMC coverage indicator proved especially challenging. The ideal denominator would be the number of babies born weighing less or equal to 2000 g. Yet as noted earlier, few if any low-income countries currently reliably capture such data for all births. Even for those births occurring in facilities, weight is not always recorded and if recorded, not always accurate and reliable. In the interim, several denominators are recommended for further testing. The preliminary mapping exercise suggested that using the expected number of live births may be the preferred denominator until measurement of birth weight improves (see **Box 1**). In this scenario, a benchmark value or range would need to be established for interpretation, similar to that used for C-section rates.

**Table 1.** KMC indicator definitions and data sources

INDICATOR	METRIC	MAIN PURPOSE	DATA SOURCE(S) AND METHODS OF COLLECTION	FREQUENCY
<b>KMC in national policy:</b> National policy recommends KMC	<b>Yes</b> = national policy recommends KMC <b>No</b> = national policy does not recommend KMC	National reporting/ monitoring	National policy documents – record review; Key informants through interview	Annually or less
<b>KMC indicator in HMIS:</b> National HMIS includes the number of newborns who received facility-based KMC care	<b>Yes</b> = national HMIS includes the number of newborns who received facility-based KMC <b>No</b> = national HMIS does not include the number of newborns who received facility-based KMC	National reporting/ monitoring	HMIS documents – record review; Key informants through interview	Annually or less
<b>Costed plan includes KMC:</b> Costed national implementation plans for maternal newborn health include KMC	<b>Yes</b> = costed plan or plans to scale up maternal, newborn and child health intervention includes KMC components <b>No</b> = no costed implementation plan OR costed implementation plan does not include KMC components	National reporting/ monitoring	Costed plans – record review; Key informants through interview	Annually or less
<b>KMC service availability:</b> Percentage of facilities with in-patient maternity services with operational KMC	<u>Numerator:</u> Number of health facilities in which KMC is operational* <u>Denominator:</u> Number of health facilities with inpatient maternity services	National reporting/ monitoring	Facility assessments and MOH records (collected through supervision or periodic audits)	Annually or less
<b>Weighed at birth:</b> Percentage of newborns weighed at birth	<u>Numerator:</u> Number of newborns weighed at birth† <u>Denominator:</u> Number of live births	Facility monitoring	Interviews with mothers + child health card review – collected through household surveys; L&D registers – collected through record review as part of facility assessment or supervision	Periodic for household surveys; routinely (monthly/quarterly) depending on need
<b>Identification of newborns ≤2000 g:</b> Percentage of live births identified as ≤2000 g	<u>Numerator:</u> Number of newborns identified as ≤2000 g <u>Denominator:</u> Number of live births	Facility monitoring	L&D registers – collected through HMIS (see notes) or through register review as part of supervision or facility assessment	Routinely (monthly/quarterly)
<b>KMC coverage‡:</b> Percentage of newborns initiated on facility-based KMC	<u>Numerator:</u> Number of newborns initiated on facility-based KMC§ <u>Denominator:</u> Expected number of live births OR expected number of LBW babies	Facility monitoring & National reporting/ monitoring	KMC registers – reported through HMIS or collected through register review as part of facility assessment; Denominator available through national and global estimates updated annually	Annually
<b>KMC monitoring:</b> Percentage of KMC newborns who are monitored by health facility staff according to protocol	<u>Numerator:</u> Number of newborns admitted to KMC who are monitored by health facility staff according to protocol (includes at minimum: assessing feeding, STS duration, weight, temperature, breathing, heart rate, urine/stools) <u>Denominator:</u> Number of newborns initiated on facility-based KMC	Facility monitoring	KMC patient charts – collected through record review as part of facility assessment/supervision visits	Quarterly or less; to be determined at country level visits
<b>Status at discharge from KMC facility:</b> Percentage of newborns discharged from KMC facility who: met facility criteria for weight gain/health status; left against medical advice; referred out; or died before discharge	<u>Numerator:</u> Number of newborns discharged from facility-based KMC who: 1) met facility criteria for weight gain, health status, feeding, thermal regulation, family competency, etc; 2) left against medical advice; 3) referred out for higher level care; 4) died before discharge <u>Denominator:</u> Number of newborns discharged from facility-based KMC	Facility monitoring	KMC registers – reported through HMIS or collected through register review as part of facility assessment	Routinely (monthly/quarterly)
<b>KMC follow-up:</b> Percentage of newborns discharged from facility-based KMC that received follow-up per protocol	<u>Numerator:</u> Number of newborns discharged from facility-based KMC that received follow-up per protocol <u>Denominator:</u> Number of newborns discharged alive who received facility-based KMC¶	Facility monitoring	KMC registers/records – reported through HMIS or collected through register review as part of facility assessment and/or) Interviews with caregivers/mothers of newborns discharged from KMC	Routinely (monthly/quarterly)

KMC – Kangaroo Mother Care; HMIS – Health Management Information System; MOH – Ministry of Health; LBW – low birth weight; L&D – Labour and Delivery; STS – skin-to-skin

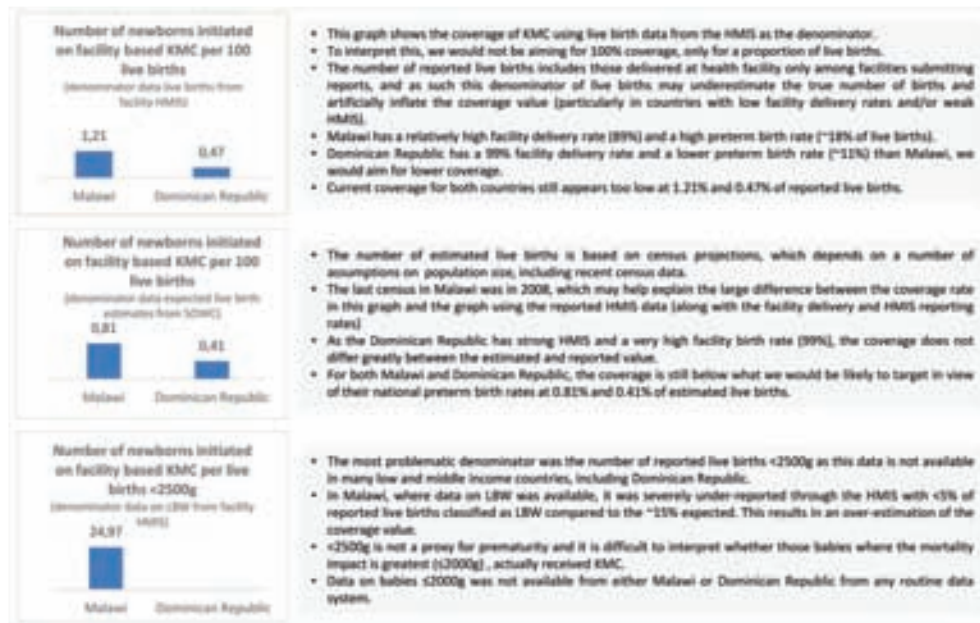
\*KMC elements already collected through Service Provision Assessments (SPA) include: staff receiving in-service training on KMC; identified space for KMC; and availability of functional infant scale. This indicator has been prioritized for further testing by the KMC Acceleration and ENAP metrics group, with particular focus on identifying and testing additional KMC elements for inclusion in future harmonized facility assessments, supervision checklists and MOH audits.

†Countries will need to define a timeframe for 'weighed at birth'. In some settings, this may include babies weighed at admission to the health care facility within a certain timeframe after delivery and babies weighed at home by a trained provider with weight documented on a maternal held record.

‡Indicator recommended as priority for inclusion in national HMIS.

§This may include facility-initiated ambulatory KMC as in Latin America (eg, Colombia); this indicator has been prioritized for further testing by the KMC Acceleration and ENAP metrics group, with particular focus on establishing the most feasible, valid, and reliable denominator and benchmarks for interpretation.

¶Countries should define their own denominator based on the national protocol for follow-up care of small and sick newborns, with an ideal denominator that captures all those infants discharged alive that were potential candidates for KMC.



**Figure 4.** Analysis of denominator options for Kangaroo Mother Care (KMC) coverage indicator in Malawi and Dominican Republic.

Three of the recommended service delivery indicators serve as proxies for quality of care processes and outcomes. While in facility, KMC babies require daily monitoring to assess and record their positioning, feeding, and weight gain and to check for signs of illness or other complications. One core indicator tracks the percentage of KMC babies who are monitored according to the national protocol by reviewing patient charts or other relevant facility records, through supervision visits or periodic assessments (monitoring adherence with recommended processes). The status of babies at the time of discharge from KMC is also an important proxy measure of quality of care, monitoring overall performance through a critical outcome indicator. Status at discharge should be captured in a KMC or postnatal register and include the following categories: met facility criteria for weight gain, health status, feeding, thermal regulation, and family competency with KMC (the ideal); died before discharge; left against medical advice (defaulters); and referred out for higher level care. These categories are similar to those used for community management of acute malnutrition (CMAM) programs for performance monitoring; however unlike CMAM, protocols differ substantially by country and there is insufficient data and experience to establish international minimum performance standards for KMC [21]. The third proxy indicator for quality relates to the level of follow-up post-discharge from facility KMC, which can be regarded as both a process and intermediate outcome. It is common for low birth weight and preterm babies to be discharged at 1500–1800 g to reduce exposure to nosocomial infections and allow space for other patients. Adherence to regular follow-up care that involves tracking growth and addressing other complications of prematurity is critical for improved outcomes of these still vulnerable preterm babies. Improving measurement of gestational age during pregnancy and/or by clinical assessment of the baby is essential for better targeting of clinical interventions and identification of infants who will require long-term and specialized care and follow-up. As discharge criteria and follow-up schedules vary by country, the indicator definition for follow-up will need to be tailored in each country accordingly. Assessing follow-up through routine sources can be complicated if babies receive follow-up care at different facility than where they received KMC, in which case periodic assessments may be required to supplement routine data.

## DISCUSSION

To the best of our knowledge, this is the first attempt to use a global consultation process to identify a prioritized set of core indicators to track country progress towards scaling up KMC. Both ENAP and the KMC Acceleration Partnership have set ambitious goals for reducing newborn morbidity and mortality through improved coverage of high impact interventions, including KMC [6,11]. Better quality data and measurement of KMC will be critical in accelerating progress of implementation and supporting scale up of the intervention. As has been seen with child health programs (eg, vaccinations) good quality, compa-

able data allows for informed planning, decision-making and targeting of programs. As direct complications of prematurity are now the leading cause of child death, comparable data are critical to foster global visibility, policy attention and accountability structures within the Sustainable Development Goals for child health. This requires a consistent approach to measurement of KMC with standardized indicators and data collection methodologies that can be captured in sustainable, country-driven health information systems. We employed a consensus-based process to develop a measurement framework and identify a set of 10 core indicators for measuring progress of KMC implementation. The resulting framework can be used to help program managers at the country level plan and set milestones that will be comparable between different settings. At a facility level, program implementers can use the service delivery indicators to identify areas for quality improvement.

The indicator selection and refinement process had several strengths. We engaged a broad range of KMC and measurement experts representing global and country level perspectives and diverse technical and methodological expertise. The candidate indicators were selected through a literature review, including peer reviewed and grey literature (surveys and program documents), and canvassing of KMC researchers and implementers to ground the work in experience with existing measures. We conducted a preliminary mapping exercise to look at availability of data for some of the most critical indicators and we consulted with country-based KMC implementers to assess the feasibility and relevance of the proposed indicators. To avoid overburdening health systems and frontline workers with unnecessary data collection requirements, we intentionally kept the list to a minimum set and focused on indicators with potential to be collected within existing, sustainable systems. Finally, the overlap between members of the ENAP metrics stream and the KMC Acceleration Partnership facilitated close collaboration and alignment of the recommended process and coverage indicators. This collaboration also allowed for wide consultation and shared learning with other groups facing similar challenges to measure interventions for newborns requiring extra care (eg, neonatal resuscitation, treatment of neonatal infections) [14].

The development of a measurement framework specific to KMC played a critical role in guiding the process of selecting and refining the indicators. Inclusion of the WHO building blocks helped ensure a health systems approach and the action sequence identified the major steps that need to take place to deliver high quality KMC services. Use of the framework allowed us to ensure indicators were evenly spread along the continuum of care from service readiness to service delivery. The framework also provides a useful reference point for program implementers, evaluators and researchers to identify additional indicators on aspects of readiness and service delivery. The intent was not to create a rigid framework, but to prioritize indicators that are relevant to implementation across a wide range of settings. The expectation is that individual programs will identify additional indicators that are program specific and adapt the framework to fit their context and data collection capacity.

As mentioned in the results, data availability is limited for most of the recommended KMC indicators. For countries in early stages of introducing KMC, the focus should be on tracking progress against the service readiness indicators. Once KMC is integrated within packages of care for preterm babies, countries can design and test data collection systems to capture more of the service delivery indicators. For countries with more established KMC services, efforts should be made to review their existing health information systems to determine the best way to integrate the recommended indicators such that select indicators are reported up to the national level. While nearly all service delivery indicators could be aggregated and reported through a national HMIS, given system constraints in most settings, priority should be given to capturing KMC initiation for tracking coverage. Countries will need to tailor some of the indicator definitions, particularly for KMC service availability, KMC monitoring, and KMC follow-up, to align with national protocols and clinical guidelines.

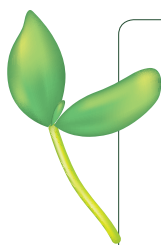
The core list of KMC indicators should be considered in light of the limitations of the process and of the indicators themselves. First, due to time constraints we were unable to undertake a formal systematic review of the literature and may have missed some relevant information. However, we did reach out to renowned KMC experts to share their experiences and materials and the majority of candidate indicators were extracted from these grey sources. Second, while we convened a series of consultations with a wide range of experts including representatives directly involved in country implementation, we were not able to sufficiently involve the principal end-users of the service delivery indicators – namely managers and service providers at district and facility levels. Third, given the aim was to develop a focused list of indicators suitable for routine systems, several aspects central to quality implementation of KMC are not reflected in the set of core indicators. Approaches to capture aspects such as timely initiation of KMC, ex-

tent of skin-to-skin and feeding practices, referral completion and health outcomes were discussed in-depth and considered only feasible within the context of research settings or special studies for the time being. In 2016, the WHO released a set of standards for improving quality of maternal and newborn care in health facilities and recommended two indicators for facilities to use to evaluate quality of KMC care; these draw attention to and have the potential to reinforce facility-level quality improvement efforts. However, measuring these quality of care indicators would require detailed information captured through daily patient charts and may not be feasible for routine monitoring and national aggregation in most settings [22].

Future work will include developing guidance for the indicators such as detailed reference sheets outlining how to collect and use the data effectively and supporting country-level partners to adapt and use the indicators. Both the KAP and ENAP metrics offer platforms for disseminating such materials to a wide audience and to collate and share additional resources and experiences gathered through collecting the indicators. The KAP regional communities of practice in Africa and Asia will convene meetings in 2016 and 2017 and provide an important opportunity to engage country-level partners to further refine the indicators. A critical next step is initiating special studies to test and validate the recommended KMC coverage indicators as outlined in the ENAP measurement improvement roadmap (see **Box 1**) [14]. This will be embedded in work to test all of the core ENAP coverage indicators for newborns with complications requiring extra care (antenatal corticosteroids, neonatal resuscitation and treatment of neonatal infection) that face similar measurement challenges. Data collection is under way to test a range of numerator and denominator options for validity (eg, sensitivity and specificity of the indicators), feasibility of measurement and usefulness through country hubs in Bangladesh and Tanzania. Another important area of future work is to establish a coordinated mechanism for global tracking of a sub-set of the core indicators to assess progress towards the KAP and ENAP goals. This will require harmonized investments in strengthening country health information systems, prioritizing capture of data to generate coverage estimates following validation efforts, as well as a system for global reporting.

## CONCLUSIONS

As KMC accelerates globally, a standardized approach to measuring implementation and progress towards effective coverage is needed. The indicators presented in this paper, identified with input from a wide range of global and country-level KMC and measurement experts, can aid efforts to strengthen monitoring systems and facilitate global tracking of KMC implementation. As data collection systems advance, we encourage program managers and evaluators to document their experiences using this framework to inform further progress and indicator refinement with the overall aim of working towards sustainable, country-led data systems.



**Acknowledgements:** We acknowledge the contributions of the following individuals who participated in one or more of the meetings to conceptualize, review, and provide inputs into the framework and indicator development process: Stella Abwao, Asri Adisamita, Rajiv Bahl, Jennifer Callaghan, Grace Chan, Nathalie Charpak, Ruth Davidge, Kim Dickson, Mary Drake, Cyril Engmann, Sabine Gabrysch, John Grove, Eni Gustina, Zelee Hill, Erin Hunter, Steve Hodgins, Lily Kak, Neena Khadka, Joy Lawn, Jim Litch, Rachel Makunde, Liz Mason, Suleiman Mohammed, Winnie Mwebesa, Harriot Nambuya, Silvia Nurdin, Magdalena Sherpa, Lara Vaz, Brendan Wackenteuer, Peter Waiswa and Evelyn Zimba.

**Disclaimer:** The opinions expressed are those of the authors and do not necessarily reflect the views of Save the Children.

**Funding:** Funding for this work was provided by Save the Children – Saving Newborn Lives.

**Authorship declaration:** TG, SM, BV and GW formed the core team leading the development of the measurement framework and indicators and consolidating participant inputs. All co-authors participated in one or more of the consultations. TG and SM wrote the manuscript and all authors reviewed the manuscript drafts and provided inputs.

**Competing interests:** The authors have completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author) and declare no conflict of interest.

- 1 Blencowe H, Cousens S, Oestergaard M, Chou D, Moller AB, Narwal R, et al. National, regional and worldwide estimates of preterm birth. *Lancet*. 2012;379:2162-72. Medline:22682464 doi:10.1016/S0140-6736(12)60820-4
- 2 Blencowe H, Cousens S, Chou D, Oestergaard M, Say L, Moller AB, et al. Born too soon: The global epidemiology of 15 million preterm births. *Reprod Health*. 2013;10 Suppl 1:S2. Medline:24625129 doi:10.1186/1742-4755-10-S1-S2
- 3 UN Inter-agency Group for Child Mortality Estimation. Levels and trends in child mortality: Report 2015. Available: [http://www.unicef.org/media/files/IGME\\_Report\\_Final2.pdf](http://www.unicef.org/media/files/IGME_Report_Final2.pdf). Accessed: 10 September 2015.
- 4 Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012;380:2197-223. Medline:23245608 doi:10.1016/S0140-6736(12)61689-4
- 5 Vesel L, Bergh AM, Kerber KJ, Valsangkar B, Mazia G, Moxon SG, et al. Kangaroo mother care: a multi-country analysis of health system bottlenecks and potential solutions. *BMC Pregnancy Childbirth*. 2015;15 Suppl 2:S5. Medline:26391115 doi:10.1186/1471-2393-15-S2-S5
- 6 WHO. UNICEF: Every Newborn: An action plan to end preventable newborn deaths Geneva: World Health Organisation 2014. Available: [http://www.everynewborn.org/Documents/Every\\_Newborn\\_Action\\_Plan-ENGLISH\\_updated\\_July2014.pdf](http://www.everynewborn.org/Documents/Every_Newborn_Action_Plan-ENGLISH_updated_July2014.pdf). Accessed: 10 September 2015.
- 7 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
- 8 Moxon SG, Lawn JE, Dickson KE, Simen-Kapeu A, Gupta G, Deorari A, et al. Inpatient care of small and sick newborns: a multi-country analysis of health system bottlenecks and potential solutions. *BMC Pregnancy Childbirth*. 2015;15 Suppl 2:S7. Medline:26391335 doi:10.1186/1471-2393-15-S2-S7
- 9 World Health Organization. WHO recommendations on interventions to improve preterm birth outcomes. Geneva: World Health Organization, 2015. Available: [http://www.who.int/reproductivehealth/publications/maternal\\_perinatal\\_health/preterm-birth-guideline](http://www.who.int/reproductivehealth/publications/maternal_perinatal_health/preterm-birth-guideline). Accessed: 10 September 2015.
- 10 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
- 11 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
- 12 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
- 13 WHO. UNICEF: Countdown to 2015: Decade of tracking progress for maternal, newborn and child survival. The 2015 report. World Health Organization: Geneva; 2015.
- 14 Moxon SG, Ruysen H, Kerber KJ, Amouzou A, Fournier 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
- 15 Brown D. 2009. Good practice guidelines for indicator development and reporting. A contributed paper Third World Forum on 'Statistics, Knowledge and Policy' Charting Progress, Building Visions, Improving Life. October 2009, Korea. Available: <http://www.oecd.org/site/progresskorea/43586563.pdf>. Accessed: 15 October 2016.
- 16 Indicators Technical Working Group of the UNAIDS Monitoring and Evaluation Reference Group. 2010. Indicator standards: Operational guidelines for selecting indicators for the HIV response. Joint United Nations Program on HIV/AIDS. Available: [http://www.unaids.org/en/media/unaids/contentassets/documents/document/2010/4\\_3\\_MERG\\_Indicator\\_Standards.pdf](http://www.unaids.org/en/media/unaids/contentassets/documents/document/2010/4_3_MERG_Indicator_Standards.pdf). Accessed: 15 October 2016.
- 17 Measure Evaluation. Family Planning and Reproductive Health Indicators Database. Selection of indicators. Available: [http://www.cpc.unc.edu/measure/prh/rh\\_indicators/overview/rationale2](http://www.cpc.unc.edu/measure/prh/rh_indicators/overview/rationale2). Accessed: 15 October 2016.
- 18 Save the Children. Kangaroo Mother Care Indicators Working Group Meeting. Meeting Summary. October 6-7, 2014. Washington DC.
- 19 World Health Organization. 2007. Everybody's Business: Strengthening health systems to improve health outcomes: WHO's Framework for Action. Available: [http://www.who.int/healthsystems/strategy/everybodys\\_business.pdf](http://www.who.int/healthsystems/strategy/everybodys_business.pdf). Accessed: 10 September 2015.
- 20 Moxon S, Guenther T, Gabrysch S, Enweronu-Laryea C, Ram P, Niermeyer S, et al. Measuring service readiness for inpatient care of small and sick newborns: a review of health facility assessment tools. Manuscript in preparation.
- 21 Save the Children UK and Humanitarian Innovation Fund. 2015. Standardised indicators and categories for better CMAM reporting. Available: <http://cmamreport.com/sites/all/themes/stc/cmam-assets/STANDARDISED%20CATEGORIES%20AND%20INDICATORS%20FOR%20BETTER%20CMAM%20REPORTING%20FINAL%20Apr%202015.pdf>. Accessed: 10 September 2015.
- 22 World Health Organization. 2016. Standards for improving quality of maternal and newborn care in health facilities. Available: [http://www.who.int/maternal\\_child\\_adolescent/documents/improving-maternal-newborn-care-quality/en/](http://www.who.int/maternal_child_adolescent/documents/improving-maternal-newborn-care-quality/en/). Accessed: 15 October 2016.



# Readiness of hospitals to provide Kangaroo Mother Care (KMC) and documentation of KMC service delivery: Analysis of Malawi 2014 Emergency Obstetric and Newborn Care (EmONC) survey data

Kondwani Chavula<sup>1</sup>, Dyson Likomwa<sup>1</sup>, Bina Valsangkar<sup>2</sup>, Richard Luhanga<sup>1</sup>, Lydia Chimtembo<sup>1</sup>, Queen Dube<sup>3</sup>, Wasihun Andualem Gobezie<sup>4</sup>, Tanya Guenther<sup>2</sup>

<sup>1</sup> Save the Children International, Lilongwe, Malawi

<sup>2</sup> Saving Newborn Lives, Save the Children, Washington, D.C., USA

<sup>3</sup> University of Malawi, College of Medicine & Ministry of Health, Blantyre, Malawi

<sup>4</sup> Averting maternal death and disability (AMDD), Ethiopia

**Background** Malawi introduced Kangaroo Mother Care (KMC) in 1999 as part of its efforts to address newborn morbidity and mortality and has continued to expand KMC services across the country. Yet, data on availability of KMC services and routine service provision are limited.

**Methods** Data from the 2014 Emergency Obstetric Newborn Care (EmONC) survey, which was a census of all 87 hospitals in Malawi, were analyzed. The WHO service availability and readiness domains were used to generate indicators for KMC service readiness and an additional domain for documentation of KMC services was included. Levels of KMC service delivery were quantified using data extracted from a 12-month register review and a KMC initiation rate was calculated for each facility by dividing the reported number of babies initiated on KMC by the number of live births at facility. We defined three levels of KMC readiness and two levels of KMC operational status.

**Results** 79% of hospitals (69/87) reported providing inpatient KMC services. More than half of the hospitals (62%; 54/87) met the most basic definition of readiness (staff, space for KMC and functional weighing scale) and 35% (30/87) met an expanded definition of readiness (guidelines, staff, space, scale and register in use). Only 15% (13/87) of hospitals had all KMC tracer items. Less than half of the hospitals (43%; 37/87) met criteria for KMC operational status at minimum levels ( $\geq 1/100$  live births), and just 16% (14/87) met criteria for KMC operational status at routine levels ( $\geq 5/100$  live births).

**Conclusions** Our study found large differences between reported levels of KMC services and documented levels of KMC readiness and service provision among hospitals in Malawi. It is recommended that facility assessments of services such as KMC include record reviews to better estimate service availability and delivery. Further efforts to strengthen the capacity of Malawian hospitals to deliver KMC are needed.

## Correspondence to:

Kondwani Chavula  
Save the Children Malawi  
Ngerengere House, Off Mchinji Road  
P. O. Box 30374  
Lilongwe, Malawi  
Kondwani.Chavula@savethechildren.org

Preterm birth is one of the leading causes of newborn morbidity and mortality globally [1–3]. Malawi has one of the highest rates of preterm births in the world, with an estimated 18% of all live births occurring before 37 completed weeks of gestation [2]. Kangaroo Mother Care (KMC) is strongly recommended by the World Health Organization (WHO) for the routine care of stable newborns weighing  $\leq 2000$  g as soon as they are clinically stable as an evidence–

based intervention to improve preterm birth outcomes [4]. Kangaroo Mother Care is defined by WHO as early, continuous and prolonged skin-to-skin contact between the mother (or other caregiver) and the baby, and exclusive breastfeeding (ideally) or feeding with expressed breastmilk [4].

Malawi was an early adopter of KMC, introducing the intervention on a pilot basis in 1999 as part of its efforts to address newborn morbidity and mortality [5]. In 2005, KMC was integrated into national policy as routine care of preterm and low birth weight (LBW) babies. During the same period there was adoption of the Malawi National Guidelines on KMC [6] and incorporation of KMC into the Ministry of Health (MoH) workplan for 2005/6. The KMC guidelines were revised in 2009 to incorporate guidelines for ambulatory and community KMC [7] and KMC was integrated into the Sexual and Reproductive Health and Rights programs [8]. Malawi continued the expansion of KMC services across the country and by 2011, KMC was reportedly established in all central- and district-level hospitals as well as several first-level health facilities [5].

In July 2015, Malawi launched its Every Newborn Action Plan (ENAP), which aims to bring partners together to accelerate progress towards ending preventable newborns deaths. The major goal of the Malawi ENAP is to achieve equitable and high-level coverage of quality essential interventions and commodities for maternal and newborn health and ultimately halving the NMR to 15 per 1000 live births by 2035 [9]. High-impact, cost-effective interventions for newborn health, like breastfeeding support and KMC, form one component of integrated health services for newborn health. Within its ENAP plan, Malawi has established a target that 75 percent of eligible preterm and low birth weight newborns should be managed with facility-based KMC by 2020 and 90 percent by 2035 [9].

Despite KMC being national policy in Malawi for the last decade, data on availability and use of KMC are limited. A 2012 evaluation of progress in KMC implementation in Malawi found that only 36% of the facilities assessed had integrated KMC into routine practice and none demonstrated sustainable practice [5]. Lack of documentation and poor record-keeping was found to be widespread and limited the ability of the evaluation to assess other aspects, such as the extent and quality of KMC practice [5]. The 2014 Emergency Obstetrics and Newborn Care (EmONC) survey provides a unique opportunity to address this information gap. The purpose of this paper is to assess the readiness of hospitals in Malawi to provide facility-based KMC and documentation of KMC service delivery.

## METHODOLOGY

### Study setting

Malawi is a small, land-locked country located in Southern Africa with an estimated population of 15.8 million [10]. Administratively, Malawi is organized into five zones (North, Central East, Central West, South East and South West) and 29 districts. Formal health care services are primarily provided by two main agencies: the government, through the Ministry of Health (MOH), operates about 60% of health facilities and the Christian Health Association of Malawi (CHAM) operates an estimated 39%. There is a small contribution from the private-for-profit health sector. Health services are provided at three levels: primary, secondary and tertiary. At primary level, services are delivered through rural hospitals, health centres, health posts, outreach clinics and also through community health initiatives. District and CHAM hospitals provide secondary level health care services to back up the activities of the primary level while central hospitals provide tertiary level and specialized services. At the time of the study, maternal and newborn health services for Malawi's 29 districts were provided through 87 hospitals and 468 health centres.

### Data source

In 2014, the Ministry of Health in Malawi conducted a nationwide assessment of EmONC services [11]. The sample included 365 public and private health facilities, covering all 87 hospitals and a 60% random sample of the 464 health centres with maternity services. Health facilities that did not offer maternal and newborn health (MNH) services were not included in the sampling frame. Convenience sampling was used to select providers and cases for review within each selected facility.

Data were collected using a structured questionnaire comprised of 10 modules, adapted from the generic modules developed by Averting Maternal Death and Disability (AMDD) [12]. Save the Children worked with AMDD and other stakeholders in Malawi to include additional questions related to KMC for six of the modules (Module 1: Identification of facility and infrastructure; Module 2: Human Resources; Mod-

ule 3: Essential drugs, equipment and supplies; Module 4: Facility case summary; Module 5: EmONC signal functions; and Module 7: Provider knowledge and competency for maternal newborn care). The modules with the additional KMC questions are available upon request from the authors.

Data were collected from September 23 – October 17, 2014, by 20 teams of three members, all of whom had a clinical background (nursing, midwifery or clinical medicine). Data collectors received five days of training covering the survey tools, research ethics and interview techniques and including field visits and role plays for practice. Quality assurance of data collection was conducted by a supervisor assigned to each team supplemented with a core survey support team comprised of representatives from the MOH, AMDD, Save the Children International, University of Malawi College of Medicine, Medical and Nurses and Midwives Council of Malawi. Double data entry for EmONC data was conducted in CSPro 5.0 and cleaned data files were exported to Stata 12.1 for analysis.

## Analysis of KMC readiness and operational status

Our analysis focused on the 87 hospitals, all of which provide inpatient maternity services and are expected to include facility-based KMC services according to the MoH national guidelines. We used the WHO Service Availability and Readiness Assessment (SARA) domains [13] to identify tracer items for KMC service readiness (staffing & guidelines, equipment & infrastructure, diagnostics, and medicines & commodities) and added a domain for documentation of KMC services provided (Table 1). We used standard international definitions of KMC and informal consultations with clinicians to select a list of tracer items that would be needed to implement KMC per the Malawi 2009 guidelines. Levels of KMC service delivery were quantified using data extracted from a 12-month register review (September 2013 – August 2014) and a KMC initiation rate was calculated for each facility by dividing the reported number of babies initiated on KMC by the number of live births at each facility. We defined three levels of KMC readiness (basic, expanded and full) and two levels of KMC operational status (basic readiness plus documentation of KMC services provided); refer to Table 2 for definitions. Three tracer items were con-

**Table 1.** Kangaroo Mother Care (KMC) service readiness items captured in 2014 Emergency Obstetric and Newborn Care (EmONC) survey

DOMAIN	TRACER ITEMS
Staffing & guidelines	Guidelines/protocols for KMC
	Staff providing KMC (any availability)
	Staff providing KMC available 24/7
Equipment & infrastructure	Defined space for KMC (separate room or in postnatal area)
	Designated beds for KMC (one or more)
Diagnostics	Functional infant weighing scale in delivery and/or postnatal ward
Medicines and commodities	Caps/hats for newborns in delivery area
	Linens/blankets for newborns in postnatal area
Documentation	KMC register in use
	KMC register complete and up-to-date

**Table 2.** Definitions of indicators for Kangaroo Mother care (KMC) service readiness and KMC operational status

INDICATOR	DEFINITION
<b>KMC service readiness:</b>	<b>Percentage of hospitals reporting inpatient KMC that have the following:</b>
<b>Basic</b>	1) defined space for KMC 2) at least one staff providing KMC 3) functional infant weighing scale
<b>Expanded</b>	Basic (1–3), plus: 4) KMC guidelines/protocols available 5) KMC register available and in use
<b>Full</b>	Expanded (1–5), plus: 6) caps/hats for newborn 7) linens/blankets for newborns
<b>KMC service operational status:</b>	<b>Percentage of hospitals reporting inpatient KMC who have the basic KMC elements (1–3), and documentation of:</b>
<b>Minimum</b>	≥1 KMC case initiated/100 live births in last 12 months
<b>Routine</b>	≥5 KMC case initiated/100 live births in last 12 months

sidered essential for provision of basic KMC services: defined space for KMC; at least one staff member reported to provide KMC services and a functional infant weighing scale. We classified the KMC initiation rate into two levels: minimum defined as one or more newborns initiated on KMC per 100 reported live births and routine defined as five or more newborns initiated on KMC per 100 reported live births. Data on the expected number of babies born weighing 2000 g or less and eligible for KMC are limited in Malawi; one recent study reported that 43% of all babies born low birth weight (<2500 g) were <2000 g, which given Malawi's estimated LBW rate of 13% suggests that around 6% of all live births would be eligible for KMC assuming the ≤2000 g cut-off [14,15]. Results were disaggregated by type of hospital (central, district, community and other).

## Ethical considerations

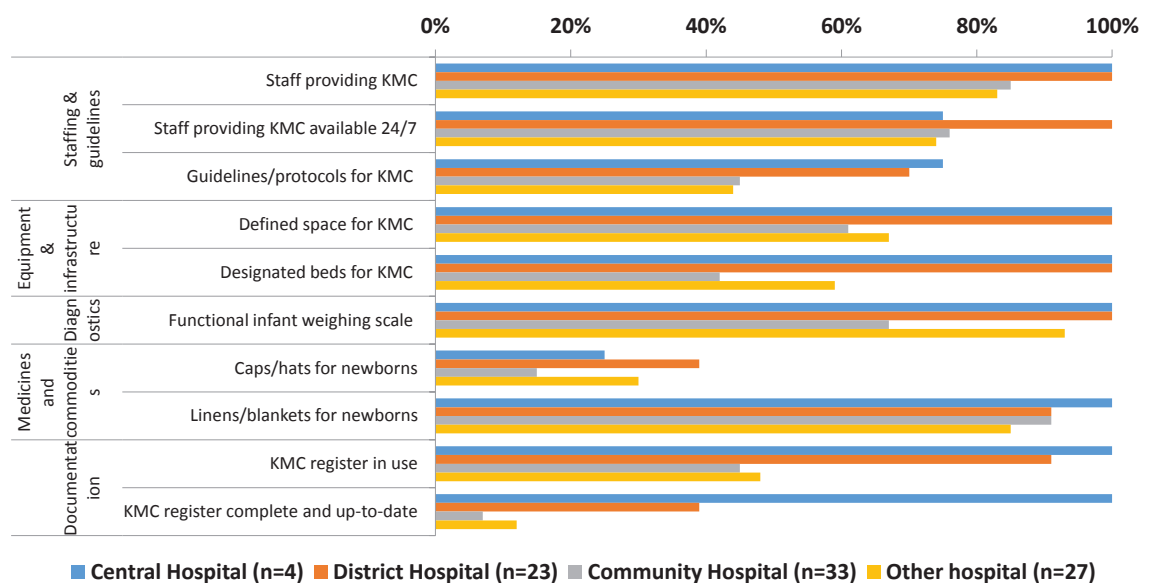
Ethical approval for the study was granted by the National Health Sciences Research Committee (NHSRC) of Malawi. The survey was led by the MOH, with technical and financial support from AMDD, Save the Children International, WHO, USAID, UNFPA, and UNICEF. Permission to conduct data collection at the facility was granted by the in-charge at each facility and individual oral consent was obtained from all individuals interviewed.

## RESULTS

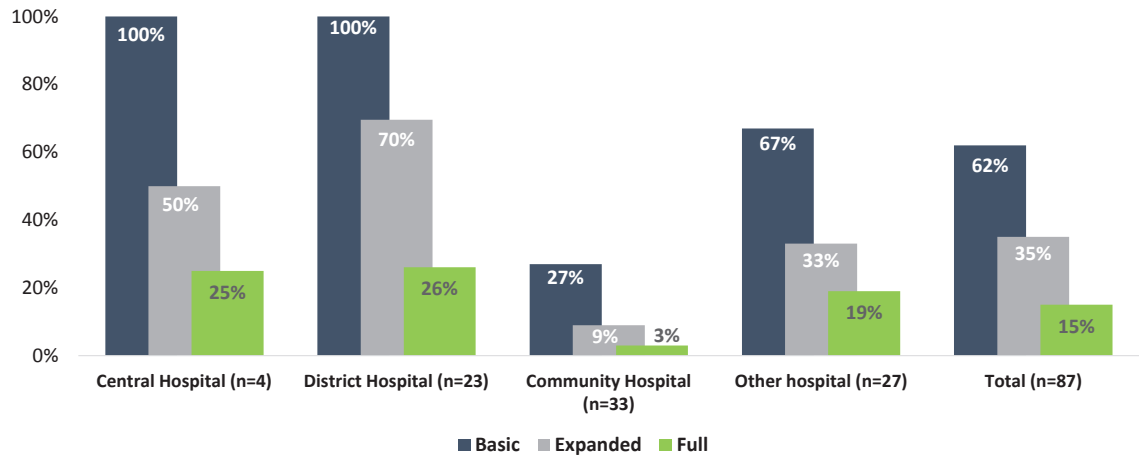
Data were available for all 87 hospitals, of which 4 were central hospitals, 23 were district hospitals, 33 were community hospitals and 27 were categorized as "other", which mainly comprised private for-profit hospitals and hospitals operated by Christian Health Association of Malawi (CHAM).

Most hospitals (79%; 69/87) reported providing inpatient KMC services (range 67% of community hospitals to 100% of central and district hospitals). **Figure 1** shows the availability of KMC tracer items by hospital type. All central and district hospitals had staff for KMC, a defined space and a functional infant weighing scale, compared to two-thirds of other hospitals and less than one-third of community hospitals. Availability of KMC guidelines, caps and hats for newborns were consistently low, even at central and district level hospitals. KMC registers were missing in more than half of community and other hospitals, and few facilities outside of the four central hospitals had up-to-date and complete KMC registers.

Sixty-two percent of hospitals (54/87) met the basic definition of readiness (staff, space for KMC and functional infant weighing scale) and 35% (30/87) met the expanded definition of readiness (guidelines, staff, space, scale and register in use) (**Figure 2**). Thirteen hospitals (15%) had all KMC tracer items. Community and other hospitals had the lowest levels of readiness.



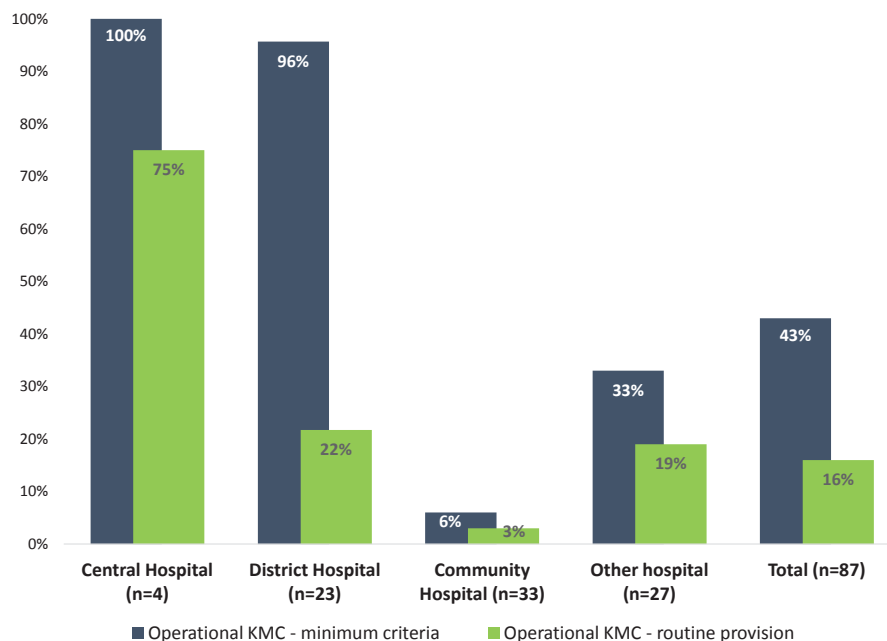
**Figure 1.** Availability of Kangaroo Mother Care (KMC) tracer items by domain and type of hospital, Malawi Emergency Obstetric and Newborn Care (EmONC) survey 2014.



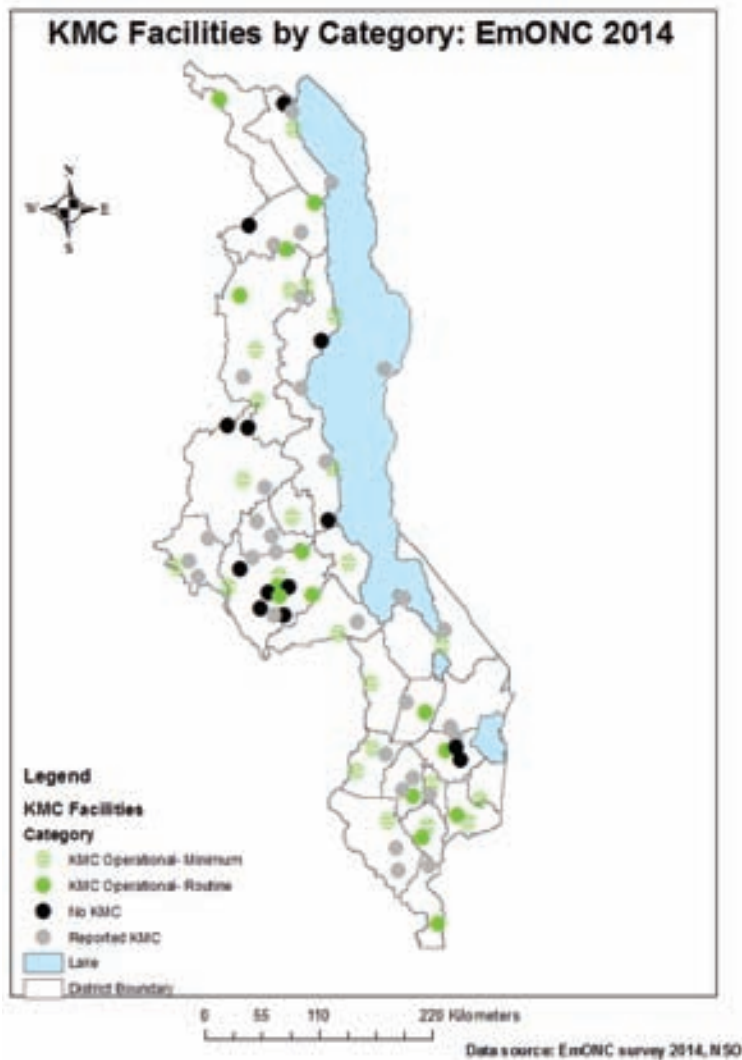
**Figure 2.** Kangaroo Mother Care (KMC) readiness among hospitals by hospital type, Malawi Emergency Obstetric and Newborn Care (EmONC) survey 2014.

The 12-month register review yielded a total of 8330 cases initiated on KMC and 211 240 live births across all 87 hospitals, for an overall KMC initiation rate of 3.9/100 live births. More than 80% (84%; 73/87) of hospitals reported providing KMC in the last three months, but just 61% (53/87) had documented cases of KMC services in the 12 months before the survey. Among facilities reporting any KMC cases, the KMC initiation rates ranged from 0.6 cases/100 live births to 17.4 cases/100 live births. Levels of KMC initiation were highest at the central hospitals; of the four central hospitals, three had KMC initiation rates  $\geq 5/100$  live births, while one had low levels of KMC initiation (1.8/100 live births). While most (96%) of the district hospitals had KMC initiation rates  $\geq 1/100$  live births, 22% had KMC initiation rates  $\geq 5/100$  live births. One third of other hospitals and 18% of community hospitals had KMC initiation rates of  $\geq 1/100$  live births. In total, 15 of the 87 hospitals recorded KMC initiation rates  $\geq 5$  cases/100 live births.

Less than half of the hospitals (43%; 37/87) met criteria for KMC operational status at minimum levels (Figure 3). All central and nearly all district hospitals (96%) met criteria for minimum operational KMC,



**Figure 3.** Kangaroo Mother Care (KMC) operational status by hospital type, Malawi Emergency Obstetric and Newborn Care (EmONC) survey 2014.



**Figure 4.** Map of Malawi showing distribution of hospitals by operational status of Kangaroo Mother Care (KMC) services, Malawi Emergency Obstetric and Newborn Care (EmONC) survey 2014.

compared to 33% of other hospitals and 6% of community hospitals (Figure 4). Fourteen of Malawi's 87 hospitals (16%) met criteria for KMC operational status at routine levels ( $\geq 5/100$  live births) (Figure 4).

## DISCUSSION

Malawi has been systematically scaling up KMC services since the intervention was introduced in 1999 and has set ambitious targets for coverage of KMC as part of its ENAP plan, aiming for 75% of eligible newborns to receive KMC by 2020. However, our analysis of the 2014 EmONC needs assessment in Malawi, the first such survey to capture detailed information on KMC services at national level, suggest that much more needs to be done if Malawi is to reach its goals. Readiness of hospitals to provide KMC was just a fraction of reported service availability. While nearly 80% of hospitals reported providing KMC services, less than two-thirds of hospitals had the minimum tracer items and only one in six had all tracer items. Our study also found poor documentation of KMC services and low levels of KMC initiation, apart from a few hospitals with well-established KMC services. Overall, just 14 of Malawi's 87 hospitals met the criteria for basic readiness and demonstrated providing KMC services routinely (at least 5 cases initiated on KMC per 100 live births).

Readiness to provide KMC services was limited primarily by lack of guidelines, caps/hats for newborns and service documentation. Having national guidelines in place at health facilities and health workers trained to use them in addition to emphasis on skills strengthening through mentorship sessions can help ensure standardization of service provision. At present, provision of caps/hats is not standard practice as mothers are expected to bring their own to the facility. Availability of caps/hats, which support thermal

care, is especially important for women experiencing preterm birth who may not be able to provide their own. Our study found that KMC documentation continues to be a challenge, particularly for district and other hospital types. At the time of the study, a register and monthly report form to track KMC services was developed, but the tools were not nationally endorsed and dissemination was ad hoc. Consequently, routine service data for KMC were limited and incomplete.

Readiness and documentation of service provision were lowest among community level hospitals. Many community hospitals lacked designated beds for KMC and basic equipment such as a functional infant scale and fewer than one in five initiated at least one case per 100 live births on KMC. In some districts, inpatient KMC services have not scaled up to community level hospitals, which often lack the infrastructure and human resource capacity to manage all units as a hospital. The common practice has been to initiate babies on KMC and refer them to facilities with inpatient KMC or to community for ambulatory KMC, leading to documentation challenges as these referrals are often not recorded.

Our study found stark differences between reported availability of KMC, readiness to provide KMC, and documented KMC service provision. Facility reports of service availability overestimated the level of KMC services, especially when the expanded or full set of tracer items were applied. The recommended global ENAP process indicator for KMC is the proportion of facilities in which a space is identified for KMC and where staff have received training in KMC in the last two years [16]. The ENAP process indicator is similar to our definition of basic readiness (staffing, defined space, and a scale), which was met by most facilities. While the EmONC study tool gathered information about staff availability to provide KMC and not KMC training directly, our results suggest that reporting on the ENAP process indicator would overestimate the availability of KMC services in Malawi. The periodic capture and use of several additional tracer items will provide a better picture of facility readiness to provide KMC services. Work is under way to develop a standard list of tracer items by KMC experts as part of the ENAP indicator development and validation process [16]. We also assessed whether facilities had 'operational' KMC, by combining basic criteria of readiness with levels of documented KMC service provision, and found that less than half of hospitals met the most minimum level of operational status, largely due to low levels of documented KMC initiation. This suggests that capturing readiness alone, as measured by availability of tracer items, is also prone to exaggerate service availability. Measures of service delivery should be captured alongside readiness where possible to obtain a clearer picture of how operational KMC is in a given facility.

While our findings suggest that we can improve the assessment of KMC service availability through better measurement of key inputs (readiness tracer items) and service delivery (operational status), understanding the strength and quality of KMC implementation at the facility and patient level will also be critical to achieving impact. As national surveys are not necessarily appropriate for gathering information on quality of care, supplementary studies and quality initiatives will be necessary for a complete picture of KMC service provision. Reporting on KMC availability, readiness, and operational status are necessary, but not sufficient indicators of KMC provision. Indeed, the presence of staff, supplies, and space for KMC is a prerequisite for quality implementation of KMC; but assessment of the quality of key components of KMC—skin-to-skin care and exclusive breastfeeding—is also needed to achieve meaningful process evaluation and scale-up of this life-saving intervention.

Building on momentum from the launch of ENAP, Malawi is investing in efforts to strengthen quality of newborn care services for small and sick babies, including KMC, and to improve documentation and reporting. The MoH is collaborating with partners, including Save the Children, MaiKhanda and others, to create an institutionalized mechanism for quality improvement of services for small and sick newborns through strengthening system building blocks such as leadership, financing, staffing, essential drugs and supplies, information systems, and ownership and partnership. The initiative aims at integrating functional small and sick newborn units, capacity building through mentorship and coaching, documentation and sharing of learning in all central and district hospitals in Malawi. Efforts such as these, in addition to Malawi's participation in the KMC Acceleration Partnership Community of Practice, provide an opportunity to improve quality and strength of implementation of KMC at the facility and patient level.

Since the EmONC study was completed, the Malawi Reproductive Health Directorate and Central Monitoring and Evaluation Department (CMED), have taken important steps to address the poor documentation and reporting of KMC [17]. In October 2015, the MoH began rolling out a national routine reporting system for KMC, which includes a simplified, user-friendly KMC register and reporting tool designed to generate a set of core indicators for tracking KMC implementation and making clinical and management decisions to improve the quality of KMC services. Data are entered at the district level into the DHIS2 (Malawi's health information system platform) and the core indicators are calculated automatically. On-

going efforts are needed to strengthen the timeliness, completeness and quality of the data and encourage regular use at facility, district and national levels.

### Limitations

This study has some important limitations. We looked at the availability of selected tracer items for KMC services; other items required to provide quality care for small babies, such as nasal gastric feeding tubes, cups and spoons for feeding, and patient monitoring charts, were not captured. At the time the Malawi EmONC survey tools were being developed, consensus regarding what tracer items should be captured for KMC was not available. The preparation of such a standardized list, as currently in process through the ENAP metrics working group, will improve such assessments in future. We relied on KMC registers to look for evidence of KMC service delivery. However, the availability and completeness of register data are often low, as was seen in this assessment. Some facilities may have been providing KMC services without using the registers, which could underestimate the level of KMC services being provided. Further, we were unable to assess the quality of the register data, and it is unclear how data quality issues would affect the results. We attempted to assess 'operational' KMC, combining measures of basic readiness with documentation of service delivery. However, EmONC surveys, like most facility assessments, rarely include an observational component and are unable to determine important aspects of the quality of care provided, and as such our measures of operational KMC do not take this into account.

### CONCLUSIONS

We found large differences between reported levels of KMC services and documented levels of KMC readiness and service provision among hospitals in Malawi. While many hospitals met the basic criteria for KMC readiness, few had most or all tracer items. Levels of documented KMC initiation were much lower than needed to achieve high coverage of KMC for preterm and LBW babies in Malawi.

We recommend that, when feasible, facility assessments of services, such as KMC, include record reviews to better estimate service availability and delivery. Further efforts to strengthen capacity of Malawian hospitals to deliver KMC are needed, particularly for district, community and other hospitals. Such efforts should include routine reviews of KMC data by facility for gaps and ensuring basic items are available to hospitals providing inpatient KMC. Regular assessment of levels of KMC service delivery through the existing DHIS2 are required to identify under-performing facilities and provide further support and supervision.



**Acknowledgements:** We thank Mary Kinney for her helpful inputs to the paper.

**Disclaimer:** The opinions expressed are those of the authors and do not necessarily reflect the views of Save the Children.

**Funding:** Funding for this work was provided by Save the Children – Saving Newborn Lives.

**Authorship Contributions:** KC and TG conceived of the study, conducted data analysis, and were primary drafters of the manuscript. DL, BV, RL, LC, QD, and AW contributed to strengthening and finalizing the manuscript.

**Competing interests:** The authors have completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author) and declare no conflict of interest.

### REFERENCES

- 1 Blencowe H, Cousens S, Oestergaard M, Chou D, Moller AB, Narwal R, et al. National, regional and worldwide estimates of preterm birth. *Lancet*. 2012;379:2162-72. Medline:22682464 doi:10.1016/S0140-6736(12)60820-4
- 2 World Health Organization. Born Too Soon: The Global Action Report on Preterm Birth. World Health Organization. Geneva: World Health Organization Geneva, 2012. Available: [http://www.who.int/pmnch/media/news/2012/preterm\\_birth\\_report/en/](http://www.who.int/pmnch/media/news/2012/preterm_birth_report/en/). Accessed: 28 March 2017.
- 3 UNICEF UN Inter-agency Group for Child Mortality Estimation. Levels and Trends in Child Mortality. New York: UNICEF, 2015. Available: [http://www.unicef.org/media/files/IGME\\_Report\\_Final2.pdf](http://www.unicef.org/media/files/IGME_Report_Final2.pdf). Accessed: 28 March 2017.
- 4 World Health Organization. WHO recommendations on interventions to improve preterm birth outcomes. Geneva: World Health Organization, 2015. Available: [http://www.who.int/reproductivehealth/publications/maternal\\_perinatal\\_health/preterm-birth-guideline](http://www.who.int/reproductivehealth/publications/maternal_perinatal_health/preterm-birth-guideline). Accessed: 15 March 2017.

- 5 Bergh AM, Banda L, Lipato T, Ngwira G, Luhanga R, Ligowe R. Evaluation of Kangaroo Mother Care services in Malawi. MCHIP, 2012. Available: <http://www.mchip.net/sites/default/files/Malawi%20KMC%20Report.PDF>. Accessed: 22 February 2017.
- 6 Ministry of Health. Malawi. Malawi National Guidelines for Kangaroo Mother Care. Lilongwe: Ministry of Health, 2005. Available: <http://www.healthynewbornnetwork.org/hnn-content/uploads/Malawi-KMC-National-Guidelines-Final-Rev-March-2009.doc>. Accessed: 22 February 2017.
- 7 Ministry of Health. Malawi. Malawi National Guidelines for Kangaroo Mother Care. Lilongwe: Ministry of Health, 2009. Available: <http://www.healthynewbornnetwork.org/resource/malawi-national-kmc-guidelines-2009/>. Accessed: 22 February 2017.
- 8 Ministry of Health. Malawi. Sexual and Reproductive Health and Rights Policy. Lilongwe: Ministry of Health, 2015. Available: <http://malawi.unfpa.org/publications/sexual-and-reproductive-health-policy>. Accessed: 22 February 2017.
- 9 Ministry of Health. Malawi. Every Newborn Action Plan: An Action Plan to End Preventable Deaths in Malawi. Lilongwe: Ministry of Health, 2015. Available: [http://www.who.int/pmnch/media/events/2015/malawi\\_enap.pdf](http://www.who.int/pmnch/media/events/2015/malawi_enap.pdf). Accessed: 22 February 2017.
- 10 National Statistics Office. Malawi. Malawi Population Projections. Zomba: National Statistics Office, 2008. Available: [www.nsomalawi.mw/publications/134-population-projections-for-malawi.html](http://www.nsomalawi.mw/publications/134-population-projections-for-malawi.html). Accessed: 22 February 2017.
- 11 Ministry Of Health. Malawi. Malawi Emergency Obstetric and Newborn Care Needs Assessment, 2014. Lilongwe: Ministry of Health, 2015. Available: [http://www.healthynewbornnetwork.org/hnn-content/uploads/Malawi-EmONC-Report-June-2015\\_FINAL.pdf](http://www.healthynewbornnetwork.org/hnn-content/uploads/Malawi-EmONC-Report-June-2015_FINAL.pdf). Accessed: 22 February 2017.
- 12 Averting Maternal Deaths and Disability. EmOC Needs Assessment Toolkit. New York: Columbia University Mailman School of Public Health, 2014. Available: <https://www.mailman.columbia.edu/research/averting-maternal-death-and-disability-amdd/toolkit>. Accessed: 8 August 2016.
- 13 World Health Organisation. Service Availability and Readiness Assessment (SARA): An annual monitoring system for service delivery. Geneva: World Health Organization, 2015. Available: [http://apps.who.int/iris/bitstream/10665/149025/1/WHO\\_HIS\\_HSI\\_2014.5\\_eng.pdf?ua=1](http://apps.who.int/iris/bitstream/10665/149025/1/WHO_HIS_HSI_2014.5_eng.pdf?ua=1). Accessed: 22 February 2017.
- 14 Ahlsén AK, Spong E, Kafumba N, Kamwendo F, Wolff K. Born too small: who survives in the public hospitals in Lilongwe, Malawi? Arch Dis Child Fetal Neonatal Ed. 2015;100:F150-4. Medline:25516259 doi:10.1136/archdis-child-2013-305877
- 15 National Statistical Office. Malawi MDG Endline Survey 2014. Zomba: National Statistical Office, 2015. Available: [http://www.nsomalawi.mw/images/stories/data\\_on\\_line/demography/MDG%20Endline/MES%202014%20Report.pdf](http://www.nsomalawi.mw/images/stories/data_on_line/demography/MDG%20Endline/MES%202014%20Report.pdf). Accessed: 28 March 2017.
- 16 Moxon SG, Ruysen H, Kerber KJ, Amouzou A, Fournier 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
- 17 Save the Children Malawi. Development of a National Routine Reporting System for Kangaroo Mother Care (KMC) Services in Malawian Health Facilities. Lilongwe: Save the Children, 2015. Available: <http://www.healthynewbornnetwork.org/hnn-content/uploads/KMC-Register-Brief-and-Forms-Final-2015.10.09-web.pdf>. Accessed: 22 February 2017.

# Developmental assistance for child and adolescent mental health in low- and middle-income countries (2007–2014): Annual trends and allocation by sector, project type, donors and recipients

Jasmine Turner<sup>1</sup>,  
Hugo Pigott<sup>2</sup>,  
Mark Tomlinson<sup>3</sup>,  
Mark JD Jordans<sup>1,4</sup>

<sup>1</sup> Department of Research and Development, War Child Holland, the Netherlands

<sup>2</sup> Independent consultant, Accra, Ghana

<sup>3</sup> Department of Psychology, Stellenbosch University, Stellenbosch, South Africa

<sup>4</sup> Centre for Global Mental Health, King's College London, UK

**Background** Globally, mental disorders are the leading cause of disability among children and adolescents. To date, there has been no estimate of developmental assistance supporting mental health projects that target children and adolescents (DAMH–CA). This study aimed to identify, describe and analyse DAMH–CA with respect to annual trends (2007–2014), sector, project type, recipient regions, and top donor and recipient countries, and estimate annual DAMH–CA per child/adolescent by region.

**Methods** Developmental assistance for all projects focused on children and adolescent mental health between 2007 and 2014 was identified on the Organisation for Economic Co-operation and Development's (OECD) Creditor Reporting System, and analysed by target population, sector, project type, donors, and recipients. The study did not include governmental or private organisation funds, nor funding for projects that targeted the community or those that included mental health but not as a primary objective.

**Results** Between 2007 and 2014, 704 projects were identified, constituting US\$ 88.35 million in DAMH–CA, with an average of 16.9% of annual development assistance for mental health. Three quarters of DAMH–CA was used to fund projects in the humanitarian sector, while less than 10% was directed at mental health projects within the education, HIV/AIDS, rights, and neurology sectors. DAMH–CA was predominantly invested in psychosocial support projects (US\$ 63.24 million, 72%), while little in absolute and relative terms supported capacity building, prevention, promotion or research, with the latter receiving just US\$ 1.2 million over the eight years (1.4% of total DAMH–CA). For 2014, DAMH–CA per child/adolescent was US\$ 0.02 in Europe, less than US\$ 0.01 in Asia, Africa, and Latin America and the Caribbean, and US\$ 0 in Oceania.

**Conclusions** To mitigate the growing burden of mental and neurological disorders, increased financial aid must be invested in child and adolescent mental health, especially with respect to capacity building, research and prevention of mental disorder projects. The present findings can be used to inform policy development and guide resource allocation, as current developmental assistance is described by sector and project type, thereby facilitating the identification of specific areas of investment need.

## Correspondence to:

Mark Jordans  
IoPPN  
16 De Crespigny Park  
Camberwell  
London SE5 8AF  
UK  
mark.jordans@warchild.nl

The Global Burden of Disease Study 2013 indicates that 21.2% of years lost to disability (YLDs) and 7.1% of disability-adjusted life years (DALYs) are attributable to mental illness, although a more recent analysis suggests that 32.4% and 13.0%, respectively, are more accurate estimates due to exclusion/inclusion issues



with the original calculations [1,2]. An estimated 20% of all children and adolescents have some form of mental disorder, one quarter of which are severe [3,4]. Globally, mental illness (including substance use disorders) is the leading cause of disability in children and young people, accounting for 54.2 million YLDs (25% of disability in children and youth), 6.3 million years of life lost (YLL) due to suicide, and 61.8 million DALYs (5th leading cause of disability and 6.3% of all DALYs, when including the burden of suicide) [4]. Despite this, the vast majority of children around the world do not have access to mental health services, largely due to severe shortages of mental health professionals, lack of health worker training on children and adolescent mental health, and stigma [5–8]. Given that an estimated 50% of adult mental disorders have their onset during adolescence, the lack of care and early intervention during this period results in lost economic productivity by both the individual and their carers [8–10]. The World Health Organisation (WHO), researchers, advocates and clinicians repeatedly recommend prioritising increased prevention of mental disorders, and the adoption of a life-course approach to address the mental health burden [11–13].

Significant strides have been made in recent years to increase the visibility and funding of mental health. The WHO has published a comprehensive mental health action plan (2013–2020) and, with the World Bank, mobilised a global alliance to scale up implementation and prioritise mental health [14–16]. Additionally, mental health and substance use disorders have been included in two health targets of the Sustainable Development Goals [17]. Despite this progress, there is a need for increased funding to support the growth of mental health services. In 2007, the Lancet Global Mental Health Group recommended an investment of US\$ 2 per person per year to scale up basic mental health care packages in LICs, and US\$ 3–4 in LMICs which, when taking inflation into account, is equivalent to US\$ 2.34 and US\$ 3.51–4.68 in 2017 US dollars [18,19]. In practice, average governmental spending on mental health ranges from 0.5% (equivalent to US\$ 0.20 per capita) of the annual health budget in low-income countries (LICs), 1.9% in low-middle-income countries (LMICs), and 2.4% and 5.1% in upper-middle- and higher-income countries (UMCs and HICs, respectively) [20].

A recent analysis of developmental assistance for mental health (DAMH) between 2007 and 2013 reveals a similar picture, with assistance more than trebling in actual terms from US\$ 53.67 million in 2007 to US\$ 196.62 million in 2013, yet remaining less than 1% of total developmental assistance for health [21]. The annual DAMH average of US\$ 133.57 million is over 50 times less than funding for HIV, despite mental, neurological and substance use disorders causing more than double the global burden of HIV/AIDS [22]. DAMH per capita was calculated to be US\$ 0.05 which, when added the average governmental spending, still leaves each individual in LICs US\$ 2.09 short per year (accounting for inflation) of the minimum investment needed to scale up basic mental health care packages [21].

To the best of our knowledge, to date there is no estimate of DAMH specifically targeted at children and adolescents (DAMH-CA). The present study aims to estimate DAMH-CA between 2007 and 2014, and to describe, analyse and comment on the distribution by year, sector, project type, top recipients and donors, recipient regions, as well as estimate annual DAMH-CA per child/adolescent by region.

## METHODS

An adaptation of previous studies' methods using the Creditor Reporting System (CRS) of the Organisation for Economic Co-operation and Development's (OECD) Development Assistance Committee (DAC) was used to identify DAMH-CA [21,23–25].

### Data source

The CRS is a comprehensive, open-access database detailing aid activities reported by DAC and non-DAC countries, multilateral organisations, and private donors [26]. It provides project-by-project information on donor disbursements (recorded by funding year), and is considered the most reliable source on development assistance projects. CRS spreadsheets detailing all aid projects between 2007 and 2014 were downloaded and searched within for children and adolescent mental health projects (project identification methodology detailed below). Data collected was conducted between November 2015 and May 2016.

### Identification of children and adolescent mental health projects

DAMH-CA is defined here as financial aid for mental health projects targeted at children and/or adolescents. Projects were identified using keywords searches in the project title, short description and long description. The keywords were based on language used for diagnosing and describing mental illness and psychosocial well-being, developed in English and translated into Italian, French, Portuguese, Spanish,

German and Dutch (see Table S1 in the **Online Supplementary Document**). When a project included a keyword, the title and descriptions were read to determine whether it should be included, and to ascertain and record the target population (community, adults, or children and adolescents). Most of these keywords were the same as those used by previous studies [21] with some additions and/or modifications, eg, “therap-“ to include therapy, therapist, therapeutic. “Wellbeing” was another addition and did present challenges due to the broad concept it encompasses; the project was included if mental well-being was specified, otherwise it was coded as a ‘multicomponent’ project (ie, a project which included mental health within its remit but not as a primary objective).

In line with previous methodologies [21], projects were searched for in the education, health, government and civil services, other social infrastructure and services, and humanitarian aid sectors. Due to the heterogeneity of multicomponent projects in terms of size and aims, it was not possible to identify the proportion of funding directed at mental health therefore, unlike previous studies [21], developmental assistance for multicomponent projects was not included as DAMH–CA.

### **Categorisation by sector, project type and population group**

Once data collection was complete, a random selection of projects was open-coded for sector and project type by two researchers [JT, HP] (n=60). The codes were discussed and finalised, and then the remaining projects were manually coded (see Table S2 for definitions of sector and project type, and Text S1 for detailed description of coding process, both in **Online Supplementary Document**).

The data were then tracked for trends in DAMH–CA by year, sector, programme type, donor, and recipient. Programmes targeting the general community were not included to prevent overestimation of DAMH–CA, as it became apparent from project descriptions that the overarching focus of community programmes was on adults, and that children and adolescents were generally secondary beneficiaries at best.

The CRS gives three options for the recording of financial data; *actual disbursement*, *money received*, and *commitment*. It was decided to use *actual disbursement* data, as *money received* was rarely recorded and there were often discrepancies between *actual disbursement* and *commitment*. When *actual disbursement* data were not available, *money received* was used instead, and for the remaining projects (n=73) *commitments*, multiplied by the *actual disbursement:commitment* ratio (1:0.816), were used. DAMH–CA is presented in 2013 constant dollars, ie, adjusted for inflation and using the exchange rates of 2013 [27].

### **Recipient and donor data, annual DAMH–CA per child/adolescent, and cumulative DAMH–CA by recipient country**

Recipient and donor information for most projects was available, therefore the top ten cumulative (ie, aggregated 2007–2014 DAMH–CA) recipients and donors between 2007 and 2014 were identified. DAMH–CA for unspecified recipients (n=14) amounted to US\$ 4.61 million. It was possible to calculate annual DAMH–CA per child/adolescent by region, using UN Department of Economic and Social Affairs population data (see Text S2 in **Online Supplementary Document** for detailed methodology). The percentage of cumulative DAMH–CA per recipient country was mapped, although the West Bank and Gaza were excluded as they received a disproportionate percentage (28.0%) of DAMH–CA, which would have rendered the shading of the rest of the map uninformative. DAMH–CA for unspecified recipient counties was not included in the mapping (US\$ 5.85 million, 6.6%), and DAMH–CA for unspecified regions (US\$ 4.61 million, 5.2%) was not included for the per child/adolescent calculations.

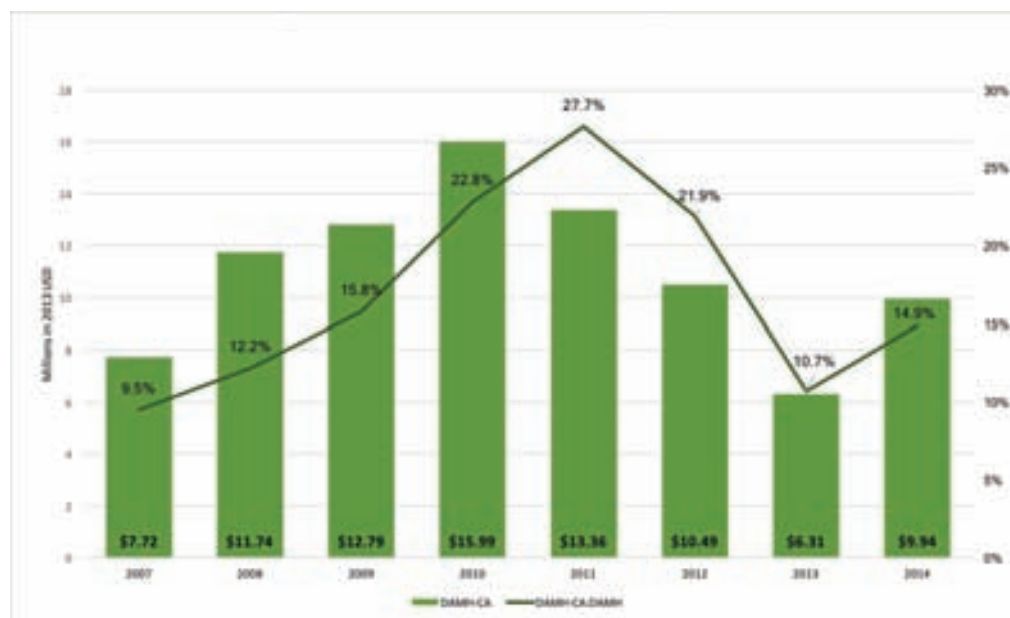
## **RESULTS**

### **DAMH by target population**

DAMH between 2007 and 2014 was US\$ 550.79 million, of which DAMH–CA was US\$ 88.35 million (16.0%). DAMH targeted at adults was US\$ 17.74 million (3.2%), and community DAMH was US\$ 444.70 million (80.7%). 704 mental health projects targeting children and adolescents were identified, making the average spending per project almost US\$ 125 500.

### **Annual trends of DAMH–CA**

Annual DAMH–CA between 2007 and 2014 was an average of US\$ 11.04 million per year, with no clear annual trend (**Figure 1**). DAMH–CA ranged from a peak in 2010 of US\$ 15.99 million to US\$ 6.31 mil-



**Figure 1.** Annual trends of developmental assistance supporting mental health projects that target children and adolescents (DAMH-CA) and as a proportion of DAMH, 2007–2014 (US\$ millions).

lion in 2013. DAMH-CA constituted an annual average of 16.9% of DAMH, ranging from 9.5% in 2007 to 27.7% in 2011. Developmental assistance for multicomponent programmes totalled US\$ 129.30 million (annual average = US\$ 16.16 million). Annually, it varied considerably, from US\$ 75.26 million in 2007 to US\$ 3.22 million in 2011 (see Figure S1 in **Online Supplementary Document**). When considered as a proportion of DAMH, there is possibly a shift from lower to higher annual investment in multicomponent programmes in more recent years, although this would need to be tracked over more years to confirm.

### DAMH-CA by sector and project type

The humanitarian sector received the highest cumulative proportion of DAMH-CA (US\$ 41.72 million, 47.2%; **Table 1**). This was followed by the health and substance use sectors (US\$ 25.86 million [29.3%] and US\$ 10.99 million [12.4%], respectively). Projects addressing autism received US\$ 3.91 million over the eight years, equivalent to 4.4% of total DAMH-CA. Other sectors (education, HIV/AIDS, rights, and neurological disorders) received a cumulative total of US\$ 5.87 million (6.6% of total DAMH-CA). Many projects supported children and adolescents affected by, or at risk of, specific diseases/disorders, namely substance use, autism, HIV/AIDS and neurological disorders. To provide richer data on the types of projects funded through DAMH-CA, those addressing the mental health impact of these diseases and disorders on children and adolescents have been differentiated from general health projects.

**Table 1.** Distribution of cumulative (2007–2014) DAMH-CA by sector and project type with percentages of sector spending for each project type, US\$ millions\*

	CAPACITY BUILDING	PREVENTION	PROMOTION	PSYCHOSOCIAL SUPPORT	RESEARCH	SECTOR TOTALS (%)
Humanitarian	0.69 (1.7%)	1.60 (3.8%)	1.76 (4.2%)	37.67 (90.3%)	N/I	<b>41.72 (47.2%)</b>
Health	3.01 (11.6%)	0.76 (2.9%)	2.18 (8.4%)	19.91 (77.0%)	N/I	<b>25.86 (29.3%)</b>
Substance use	0.45 (4.1%)	7.17 (65.2%)	0.40 (3.6%)	2.53 (23.0%)	0.44 (4.0%)	<b>10.99 (12.4%)</b>
Autism	1.74 (44.6%)	N/I	0.18 (4.6%)	1.34 (34.3%)	0.65 (16.6%)	<b>3.91 (4.4%)</b>
Education	2.10 (61.9%)	N/I	0.65 (19.3%)	0.57 (16.8%)	0.07 (2.0%)	<b>3.39 (3.8%)</b>
HIV/AIDS	0.20 (16.2%)	0.01 (1.0%)	N/I	1.04 (82.8%)	N/I	<b>1.26 (1.4%)</b>
Rights	0.05 (7.5%)	N/I	0.64 (92.5%)	N/I	N/I	<b>0.69 (0.8%)</b>
Neuro	0.32 (60.3%)	N/I	N/I	0.18 (33.3%)	0.03 (6.4%)	<b>0.53 (0.6%)</b>
Project type totals (%)	<b>8.57 (9.7%)</b>	<b>9.54 (10.8%)</b>	<b>5.81 (6.6%)</b>	<b>63.24 (71.6%)</b>	<b>1.20 (1.4%)</b>	<b>88.35 (100.0%)</b>

DAMH-CA – developmental assistance supporting mental health projects that target children and adolescents, N/I – no investment

\*Row percentages show the percentage of DAMH-CA for project types within the different sectors. Percentages in the ‘Sector totals’ column show the percentage that each sector constitutes of the total DAMH-CA.

Over 90% of DAMH–CA in the humanitarian sector was directed at the provision of psychosocial support, where the majority of projects funded counselling and well-being promotion projects. Very little was invested in building the capacity of communities and aid workers to sustain the psychosocial response in the long-term, and none was invested in research. Similarly, the majority of DAMH–CA invested in the general health sector was aimed at the provision of psychosocial support, of which 68% of the projects were funded by UNICEF, predominantly as child friendly spaces. Other project types included training caregivers and counsellors in psychosocial support, well-being promotion projects, and the construction of youth mental health centres. Again, there was no investment in research. Substance use was the only sector which received DAMH–CA for all project types, and which was aimed predominantly at prevention of drug abuse, particularly in South America. Other projects included psychoeducation and improvement of counselling services, and many focused on street children. Much of the DAMH–CA for autism projects funded those supporting and training families to take care of their children's needs, establishing specialised centres and services, and providing therapy. Over half of the projects in the education sector had a capacity-building focus; the development of mental health services in schools, and construction of specialised centres. Others involved well-being promotion programmes and those supporting cognitive development. There were no projects focusing on prevention of mental disorders within education. Over 80% of DAMH–CA targeting children and adolescents affected directly or indirectly by HIV/AIDS funded the provision of psychosocial support, and most of the remaining assistance was spent on projects that trained counsellors or increased access to counsellors. Very little money – less than 1% of total DAMH–CA – was invested in the promotion of rights of people with mental health problems, but those that were funded frequently supported young people to be advocates for mental health and mental health care. Lastly, most funding within neuro supported the building of observatories, establishment of services, and training of staff for improved care for children and adolescents with neurological and intellectual disabilities.

Capacity building projects frequently focused on building centres and facilities, training professionals who are in contact with children and adolescents (from general health staff to teachers to counsellors) as well as families and/or caregivers, and especially targeted children with autism and mentally handicapped children. Prevention projects predominantly aimed to prevent drug and alcohol abuse, especially targeting vulnerable children and youth, such as street children and orphans. Promotion projects varied considerably, and often aimed to promote well-being, and social and cognitive development. Projects providing psychosocial support received over 70% of total DAMH–CA, and were offered through community–, school– and centre–based initiatives.

The project descriptions were quite vague, making it difficult to glean deeper information about the forms of support. Lastly, many of the research studies were epidemiological, and focused on children with autism and special needs.

### Annual trends of DAMH–CA by sector

Absolute DAMH–CA in the humanitarian sector decreased 73.8% between 2011 and 2014 (Figure 2 for annual distribution of DAMH–CA by sector). Since 2012, the health sector received the largest proportion of DAMH–CA, ranging between 37.1% and 53.2%, with an annual average of US\$ 3.23 million. DAMH–CA for substance use disorders dropped from 2011, with a 2011–2014 average of just US\$ 0.27 million per year, equating to an almost 90% decrease in average spending compared with 2007–2010 (average = US\$ 2.48 million).



**Figure 2.** Annual distribution of developmental assistance supporting mental health projects that target children and adolescents (DAMH–CA) by sector, 2007–2014 (US\$ millions).

DAMH-CA for autism fluctuated between 2007 and 2014, with an investment of US\$ 1.06 million in 2014 (11.0% of DAMH-CA for 2014). Combined annual DAMH-CA for HIV/AIDS, education, rights, and neurological disorders was extremely low in absolute terms, never reaching US\$ 1 million (see Table S3 in **Online Supplementary Document** for breakdown of annual spending).

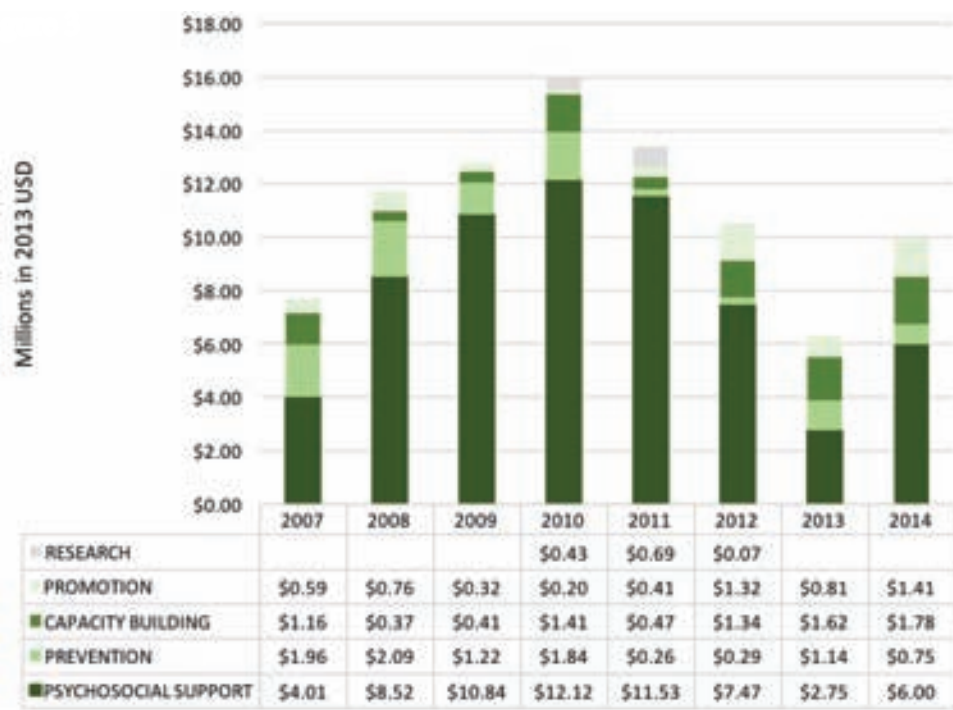
### Annual DAMH-CA by project type

DAMH-CA for projects providing psychosocial support consistently received the largest annual proportions, although decreased in 2014 by more than half compared to its peak in 2010 (**Figure 3**). DAMH-CA for prevention projects more than halved from an average of US\$ 1.78 million between 2007 and 2010 to US\$ 0.61 million between 2011 and 2014. DAMH-CA for capacity building steadily increased since 2011, although remained low in absolute terms between 2007 and 2014 with an average of US\$ 1.7 million per year. Promotion projects consisted of just 6.6% (US\$ 5.81 million) of DAMH-CA between 2007 and 2014. Funding supporting research on child and adolescent mental health consisted of just 1.4% of DAMH-CA between 2007 and 2014, and received no funds in 2007–2009, 2013, and 2014.

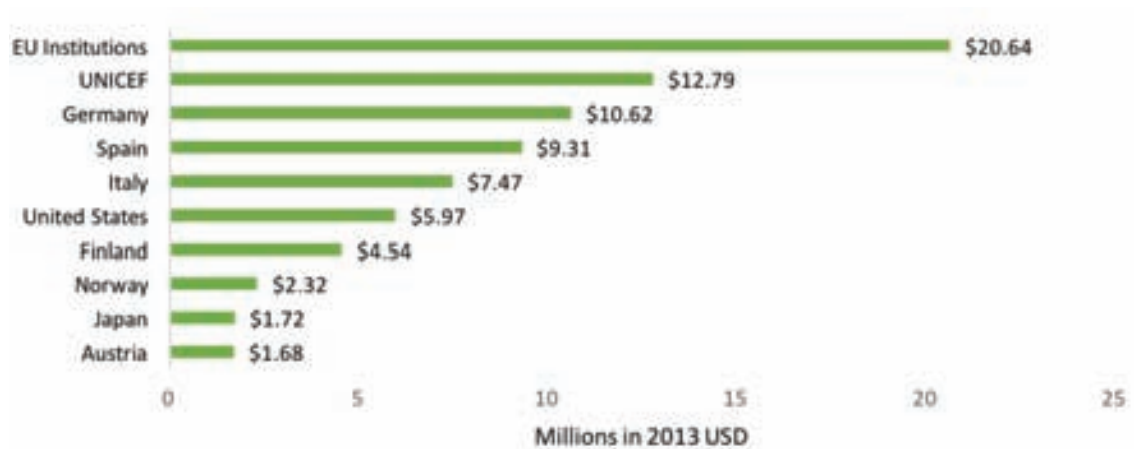
### Top ten donors and recipients of DAMH-CA, and DAMH-CA per child/adolescent

Most DAMH-CA was multilateral, ie, aid from international institutions with governmental membership. EU institutions and UNICEF together invested over US\$ 33 million – almost 40% – of DAMH-CA between 2007 and 2014 (**Figure 4**). The majority (62.1%) of funding from EU institutions was for humanitarian projects, mainly implemented in the West Bank and Gaza. Many projects by UNICEF supported child friendly spaces, implemented all over the world. Germany provided DAMH-CA for almost one quarter (22.8%) of total DAMH-CA targeting autism, and one fifth of the total DAMH-CA for capacity building. DAMH-CA from Spain constituted almost 60% of total funding for substance use projects, predominantly implemented in South America. DAMH-CA from Italy had no consistent pattern for recipient, project type, or sector. Most aid (77.8%) from the USA was for humanitarian projects in Africa and the Middle East, although several of their projects did not specify the recipient. Over one third (36.4%) of DAMH-CA from Finland, Norway, Japan and Austria was directed Sub-Saharan Africa, predominantly funding psychosocial support projects within the health sector.

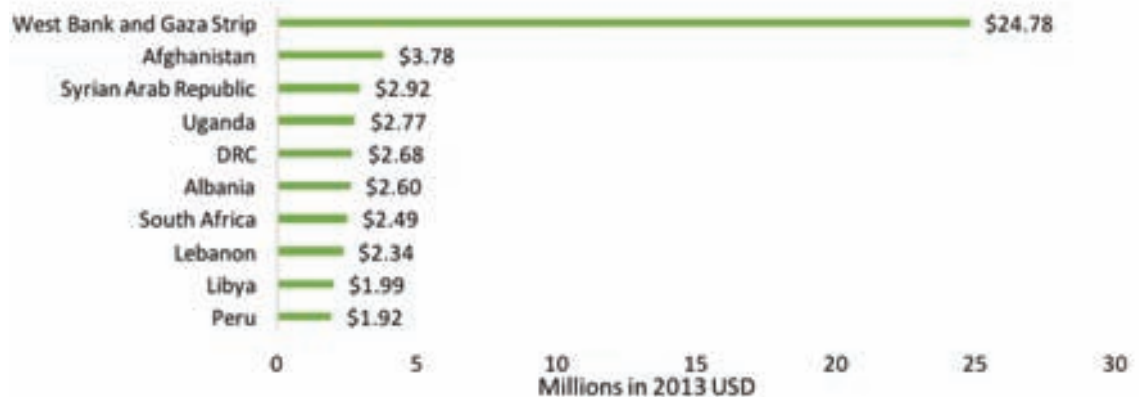
The West Bank and Gaza received almost one third of total DAMH-CA (US\$ 24.78 million; 28.0%; **Figure 5**), over half of which (US\$ 14.4 million) came from EU institutions. All DAMH-CA for Afghanistan was from UNICEF and Germany, and funded psychosocial support and child friendly spaces. Eight out



**Figure 3.** Annual distribution of developmental assistance supporting mental health projects that target children and adolescents (DAMH-CA) by project type, 2007–2014 (US\$ millions).



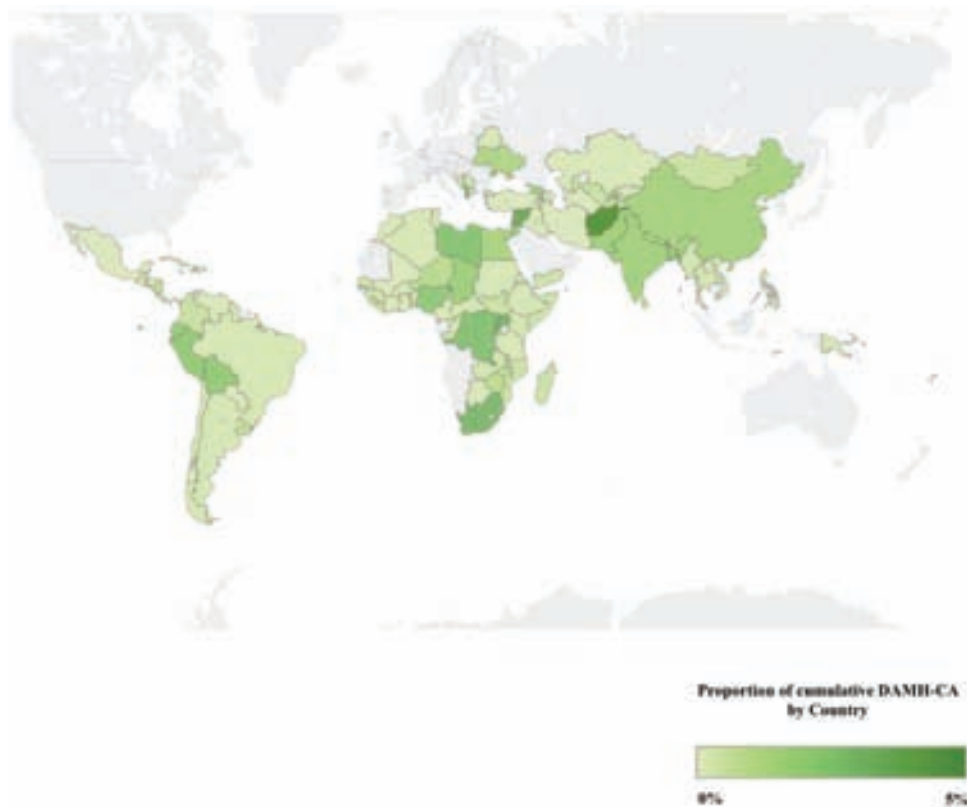
**Figure 4.** Top ten cumulative donors of developmental assistance supporting mental health projects that target children and adolescents (DAMH-CA), 2007–2014.



**Figure 5.** Top ten cumulative recipients of developmental assistance supporting mental health projects that target children and adolescents (DAMH-CA), 2007–2014.

of the 17 projects implemented in Syria were specifically aimed at Iraqi refugees, while the remaining projects encompassed psychosocial support and humanitarian action. Projects in Uganda predominantly provided psychosocial support for vulnerable children, and those in the DRC addressed both the psychosocial impact of conflict on children and capacity building. Most projects in both Albania and Peru were aimed at substance abuse prevention. In South Africa, many projects targeted children affected by HIV/AIDS and other chronic illnesses. Projects in Lebanon often concerned the psychosocial support of refugee children from neighbouring countries. DAMH-CA to Libya consisted of a single ‘emergency response’ disbursement from Italy in 2011. The proportional distribution of cumulative DAMH-CA by country is indicated in Figure 6.

Cumulative DAMH-CA (2007–2014) by region was as follows: Asia received US\$ 46.24 million (52.3%), Africa received US\$ 22.52 million (25.5%), Latin America and the Caribbean (LACAR) received US\$ 9.83 million (11.1%), Europe received US\$ 4.91 million (5.6%), and Oceania received US\$ 0.24 million (0.3%). DAMH-CA for unspecified recipient regions was US\$ 4.61 (5.2%). In 2014, the DAMH-CA for one region only (Europe) equated to more than one dollar cent per child/adolescent, three regions (Asia, Africa, Latin America & the Caribbean) received less than one cent per child/adolescent, and Oceania receive no DAMH-CA (as was the case for half of the years between 2007–2014; Table 2).



**Figure 6.** The distribution of developmental assistance projects that target children and adolescents (DAMH-CA) (2007–2014) by recipient country.

**Table 2.** Annual developmental assistance for DAMH-CA per child/adolescent by region, 2007–2014\*

	2007	2008	2009	2010	2011	2012	2013	2014
Africa	<0.01	<0.01	<0.01	<0.01	0.01	<0.01	<0.01	<0.01
Asia	0.02	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Europe	0.02	0.01	<0.01	0.04	<0.01	<0.01	<0.01	0.02
LACAR	<0.01	<0.01	<0.01	0.02	<0.01	<0.01	<0.01	<0.01
Oceania	0.02	N/I	0.03	0.01	N/I	<0.01	N/I	N/I

DAMH-CA – developmental assistance supporting mental health projects that target children and adolescents, LACAR – Latin America and the Caribbean, N/I – no investment

\*Excluded was DAMH-CA for unspecified regions (US\$ 4.61; 5.2%).

## DISCUSSION

Analysis of the OECD's CRS data revealed that DAMH-CA between 2007 and 2014 is low in both absolute terms and relative to DAMH, amounting to a total of US\$ 88.35 million (16% of DAMH 2007–2014, annual average = US\$ 11.04 million). DAMH-CA steadily increased between 2007 and 2010, progressively decreased from 2010 until 2013, and then increased from the lowest annual total of US\$ 6.31 in 2013 to US\$ 9.94 in 2014. This pattern could represent a cyclic funding trend, although lack of data on project durations prevent definitive conclusions. Another possible explanation is that funding for children and adolescent mental health is being increasingly channelled into multicomponent development projects which include mental health within their remit; for 2013 and 2014, multicomponent projects constituted 50–70% of total developmental assistance for multicomponent projects combined with DAMH-CA. While integration of mental health services is encouraging, developmental assistance focusing on mental health specifically is still fundamental to build capacity and expertise amongst health practitioners and lay professionals.

704 mental health projects targeting children and adolescents were identified, with an average of US\$ 125 500 per project. In 2014, just one region received over US\$ 0.01 in DAMH-CA per child/adolescent. Comparison of 2007 and 2014 DAMH-CA per child/adolescent revealed a decrease in two regions and no increase in others, likely due to population increase as well as insufficient funding. These figures in-

dicating that we are still far from the goal of universal mental health care and the minimum investment needed for the scale up of basic mental health care packages [18]. Given that children and young people constitute 44% of the global population, 91% of whom live in LMICs [28], developmental assistance is currently inadequate to support the mental health needs of this demographic. Considering the future 'human capital' of this age group and the burden associated with mental illness – 54.2 million YLDs, 6.3 million YLL, 61.8 million DALYs – low investment in mental well-being constitutes a missed opportunity to mitigate the increasing burden of mental and neurological disorders [4], not to mention a restriction on said generation's advancement and ability to flourish.

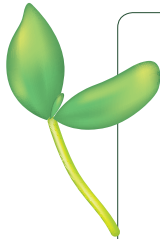
Over 75% of total DAMH-CA was invested in projects implemented within the humanitarian and health sectors, predominantly for the provision of psychosocial support but relatively little in the way of promotion, prevention, or capacity building to sustain a long-term response. A public health approach involving research, prevention, promotion, capacity building and psychosocial support is strongly recommended for mental health, however substance use was the only sector in which there was investment in all project types [29,30]. Autism, as a single disorder, received almost 5% of total DAMH-CA, and encouragingly almost half of this aimed to increase capacity through training professional and lay workers, establishing specialised institutions and services, and supporting parents to respond to their children's needs. Extremely little funding supports neurological disorders; less than 1% of total DAMH-CA.

Very little DAMH-CA supported mental health projects within the education sector; less than US\$ 4 million over the eight years. The relationship between mental health and education is bidirectional, with poor mental health being associated with early school dropout and subsequent long-term negative consequences at the individual, familial, societal and even national level while, in addition to health benefits, psychological well-being has been associated with improved educational outcomes, social relationships and productivity, and reduced absenteeism, anti-social behaviour, and crime [31]. Increased planning and financial collaboration between the health and education sectors is necessary, in particular for programmes aiming to prevent the development of mental disorders and symptoms as these are promoted as an effective and cost-effective approach, and schools are arguably the platform with the highest reach [11–13,32]. Just 11% – under US\$ 10 million – of DAMH-CA between 2007 and 2014 funded prevention projects, none of which were based in schools. The only area in which there was substantial investment in prevention programmes was against the development of drug and alcohol use disorders.

Much DAMH-CA from top donors was invested in projects within the humanitarian sector, especially directed at the West Bank and Gaza, which received almost 30% of total DAMH-CA. Many of the top recipient countries are currently or recently affected by armed conflict, or neighbouring those affected by conflict. There is a distinct lack of investment trends, suggesting a lack of direction for children and adolescent mental health care. To improve this, and increase autonomous and sustainable mental health care in LMICs, increased investment is required especially in research and capacity building. Cumulative DAMH-CA for these project types amounted to just US\$ 1.20 (1.4%) and US\$ 8.57 million (9.7%), respectively. Investment in research and capacity building is essential to equip countries with the necessary data to guide their financing of mental health services, and strengthen the design, implementation, evaluation, and scale-up of evidence-based interventions.

It must be noted that the figures presented here are an underestimate of mental health funding for children and adolescents, as they do not include governmental and private organisation funds. This is a limitation of the study, especially regarding the estimate for research funding, as much of this funding comes from such sources. However, a database (or similar) detailing research projects funded through different agencies is not available, and therefore including research projects only funded by specific funding bodies would have clouded the conceptual clarity of the study. It was also not possible to estimate the proportion of DAMH-CA for multicomponent and community-targeted projects due to their heterogeneity, although the project descriptions suggested that mental health and children and adolescents, respectively, were not the primary targets of these projects. Increased reporting transparency and detail in the CRS would give clearer insight into the distribution of the benefits of these projects. The exclusion of multicomponent programmes is the likely reason for discrepancies with previous DAMH estimates, as well as the inclusion of additional keywords geared towards the psychosocial side of mental illness, such as 'well-being' and 'therapy'. Compared to previous methodologies (eg, [23]), the method used here to identify projects targeting children and adolescents was more thorough, as each project was considered for inclusion or exclusion as opposed to the calculation of developmental assistance based on population proportions.

Given the widespread consensus that the optimal method for negotiating the increasing burden of mental health is through targeting children and adolescents as part of a life-course approach [4,13], increased DAMH-CA is crucial. Projects aiming to promote well-being and prevent mental disorders merit a substantial increase in investment, as they have been demonstrated to be both effective and cost-effective, particularly those implemented in schools and through national policy [32]. Research and capacity building especially require increased funding, to ensure appropriate and sustainable care and promotion of mental health and well-being of future generations.



**Disclaimer:** The views expressed in this article are the authors' own, and not the official position their respective institutions.

**Funding:** MT is a lead investigator with the Centre of Excellence in Human Development, University Witwatersrand, South Africa, and is supported by the National Research Foundation, South Africa.

**Authorship contributions:** JT: literature search, study design, data collection and analysis, figure design, data interpretation, and manuscript writing. HP: literature search, study design, data collection and analysis, figure design, data interpretation, and manuscript review. MT: conceived the study, and review of methods and manuscript. MJ: conceived the study, review of methods and manuscript, and overall supervision of the study. All authors have contributed to, read and approved the final manuscript.

**Conflict of interest:** The authors completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available upon request from the corresponding author), and declare no conflict of interest.

## REFERENCES

- 1 Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2015;386:743–800. Medline:26063472 doi:10.1016/S0140-6736(15)60692-4
- 2 Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. *Lancet Psychiatry*. 2016;3:171–8. Medline:26851330 doi:10.1016/S2215-0366(15)00505-2
- 3 Davidson LL, Grigorenko EL, Boivin MJ, Rapa E, Stein A. A focus on adolescence to reduce neurological, mental health and substance-use disability. *Nature*. 2015;527:S161–6. Medline:26580322 doi:10.1038/nature16030
- 4 Erskine HE, Moffitt TE, Copeland WE, Costello EJ, Ferrari AJ, Patton G, et al. A heavy burden on young minds: The global burden of mental and substance use disorders in children and youth. *Psychol Med*. 2015;45:1551–63. Medline:25534496 doi:10.1017/S0033291714002888
- 5 Tyano S, Keren M. Some reflections on the development of child and adolescent psychiatry. *World Psychiatry*. 2005;4:154–5. Medline:16633534
- 6 Saraceno B, van Ommeren M, Batniji R, Cohen A, Gureje O, Mahoney J, et al. Barriers to improvement of mental health services in low-income and middle-income countries. *Lancet*. 2007;370:1164–74. Medline:17804061 doi:10.1016/S0140-6736(07)61263-X
- 7 Patel V, Kieling C, Maulik PK, Divan G. Improving access to care for children with mental disorders: A global perspective. *Arch Dis Child*. 2013;98:323–7. Medline:23476001 doi:10.1136/archdischild-2012-302079
- 8 Morris J, Belfer M, Daniels A, Flisher A, Villé L, Lora A, et al. Treated prevalence of and mental health services received by children and adolescents in 42 low-and-middle-income countries. *J Child Psychol Psychiatry*. 2011;52:1239–46. Medline:21554305 doi:10.1111/j.1469-7610.2011.02409.x
- 9 Kessler RC, Angermeyer M, Anthony JC, DE Graaf R, Demyttenaere K, Gasquet I, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry*. 2007;6:168–76. Medline:18188442
- 10 Belfer ML. Child and adolescent mental disorders: The magnitude of the problem across the globe. *J Child Psychol Psychiatry*. 2008;49:226–36. Medline:18221350 doi:10.1111/j.1469-7610.2007.01855.x
- 11 Collins PY, Patel V, Joestl SS, March D, Insel TR, Daar AS. Grand challenges in global mental health. *Nature*. 2011;475:27–30. Medline:21734685 doi:10.1038/475027a
- 12 World Health Organization. Prevention of Mental Disorders: Effective Interventions and Policy Options. 2004. Available: [http://www.who.int/entity/mental\\_health/evidence/en/prevention\\_of\\_mental\\_disorders\\_sr.pdf](http://www.who.int/entity/mental_health/evidence/en/prevention_of_mental_disorders_sr.pdf). Accessed: 1 November 2016.
- 13 World Health Organization, Calouste Gulbenkian Foundation. Social determinants of mental health. World Health Organization: Geneva; 2014.
- 14 World Health Organization. Mental Health Action Plan 2013–2020. 2013. Available: [http://apps.who.int/iris/bitstream/10665/89966/1/9789241506021\\_eng.pdf?ua=1%5Cnhttp://apps.who.int/iris/bitstream/10665/89966/1/9789241506021\\_eng.pdf](http://apps.who.int/iris/bitstream/10665/89966/1/9789241506021_eng.pdf?ua=1%5Cnhttp://apps.who.int/iris/bitstream/10665/89966/1/9789241506021_eng.pdf). Accessed: 1 November 2016.

- 15 World Health Organization. World Bank. Investing in treatment for depression and anxiety leads to fourfold return. 2016. Available: <http://www.who.int/mediacentre/news/releases/2016/depression-anxiety-treatment/en/>. Accessed: 1 November 2016.
- 16 World Bank. Mental Health. 2016. Available: <http://www.worldbank.org/en/topic/health/brief/mental-health>. Accessed: 8 December 2016.
- 17 United Nations. Sustainable Development Goal 3. 2015. Available: <https://sustainabledevelopment.un.org/sdg3>. Accessed: 1 November 2016.
- 18 Lancet Global Mental Health Group, Chisholm D, Flisher AJ, Lund C, Patel V, Saxena S, Thornicroft G, et al. Scale up services for mental disorders: A call for action. *Lancet*. 2007;370:1241-52. Medline:17804059 doi:10.1016/S0140-6736(07)61242-2
- 19 US Inflation Calculator. Inflation Calculator. 2017. Available: <http://www.usinflationcalculator.com/>. Accessed: 21 February 2017.
- 20 World Health Organization. Investing in mental health: Evidence for Action. World Health Organization; Geneva: 2013.
- 21 Gilbert BJ, Patel V, Farmer PE, Lu C. Assessing development assistance for mental health in developing countries: 2007 – 2013. *PLoS Med*. 2015;12:e1001834. Medline:26035429 doi:10.1371/journal.pmed.1001834
- 22 Mackensie J, Kesner C. Mental health funding and the SDGs: What now and who pays? 2016. Available: <https://www.odi.org/sites/odi.org.uk/files/resource-documents/10573.pdf>. Accessed: 1 November 2016.
- 23 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
- 24 Pitt C, Greco G, Powell-Jackson T, Mills A. Countdown to 2015: assessment of official development assistance to maternal, newborn, and child health, 2003–08. *Lancet*. 2010;376:1485-96. Medline:20850869 doi:10.1016/S0140-6736(10)61302-5
- 25 Patel P, Roberts B, Guy S, Lee-Jones L, Conteh L. Tracking official development assistance for reproductive health in conflict-affected countries. *PLoS Med*. 2009;6:e1000090. Medline:19513098 doi:10.1371/journal.pmed.1000090
- 26 Organization for Economic Cooperation and Development. OECD Statistics. 2016. Available: <http://stats.oecd.org/index.aspx?DataSetCode1>. Accessed: 15 June 2016.
- 27 Organization for Economic Cooperation and Development. Technical Guide to terms and data in the Creditor Reporting System (CRS) Aid Activities database. 2016. Available: <http://www.oecd.org/dac/stats/crsguide.htm>. Accessed: 10 June 2016.
- 28 United Nations. World Population Prospects: The 2010 Revision. United Nations: New York; 2011.
- 29 Wahlbeck K. Public mental health: The time is ripe for translation of evidence into practice. *World Psychiatry*. 2015;14:36-42. Medline:25655149 doi:10.1002/wps.20178
- 30 World Health Organization. Promoting Mental Health: Concepts, Emerging Evidence, Practice: Summary Report. 2004. Available: [http://www.who.int/mental\\_health/evidence/en/promoting\\_mhh.pdf](http://www.who.int/mental_health/evidence/en/promoting_mhh.pdf). Accessed: 22 February 2016.
- 31 European Union. European Framework for Action on Mental Health and Wellbeing. EU Joint Action on Mental Health and Wellbeing. Brussels: European Union; 2016.
- 32 Petersen I, Evans-lacko S, Semrau M, Barry MM, Chisholm D, Gronholm P, et al. Promotion, prevention and protection: Interventions at the population- and community-levels for mental, neurological and substance use disorders in low- and middle-income countries. *Int J Ment Health Syst*. 2016;10:30. Medline:27069506 doi:10.1186/s13033-016-0060-z



# Uptake and predictors of early postnatal follow-up care amongst mother–baby pairs in South Africa: Results from three population–based surveys, 2010–2013

Anna Larsen<sup>1</sup>, Mireille Cheyip<sup>1</sup>,  
 Getahun Aynalem<sup>1</sup>, Thu–ha Dinh<sup>2</sup>,  
 Debra Jackson<sup>4,6</sup>, Nobubelo  
 Ngandu<sup>3</sup>, Witness Chirinda<sup>3</sup>,  
 Mary Mogashoa<sup>1</sup>, Gupreet Kindra<sup>1</sup>,  
 Carl Lombard<sup>3</sup>, Ameena Goga<sup>3,5</sup>

<sup>1</sup> Division of Global HIV/AIDS and Tuberculosis, Center for Global Health, US Centers for Disease Control and Prevention (CDC), Pretoria, South Africa

<sup>2</sup> Division of Global HIV/AIDS and Tuberculosis, Center for Global Health, US Centers for Disease Control and Prevention (CDC), Atlanta, Georgia, USA

<sup>3</sup> Health Systems Research Unit, South African Medical Research Council (HSRU, SAMRC), Pretoria, South Africa

<sup>4</sup> School of Public Health, University of the Western Cape, Cape Town, South Africa

<sup>5</sup> Department of Paediatrics, University of Pretoria, Pretoria, South Africa

<sup>6</sup> United Nations Children's Fund (UNICEF), New York, New York, USA

**Background** Achieving World Health Organization (WHO) recommendations for postnatal care (PNC) within the first few weeks of life is vital to eliminating early mother–to–child transmission of HIV (MTCT) and improving infant health. Almost half of the annual global deaths among children under five occur during the first six weeks of life. This study aims to identify uptake of three PNC visits within the first six weeks of life as recommended by WHO among South African mother–infant pairs, and factors associated with uptake.

**Methods** We analyzed data from three facility–based, nationally representative surveys (2010, 2011/12 and 2012/13) primarily designed to determine the effectiveness of the South African program to prevent MTCT. This analysis describes the proportion of infants achieving the WHO recommendation of at least 3 PNC visits. Interviews from 27 699 HIV–negative and HIV–positive mothers of infants aged 4–8 weeks receiving their six week immunization were included in analysis. Data were analyzed using STATA 13.0 and weighted for sample ascertainment and South African live births. We fitted a multivariable logistic regression model to estimate factors associated with early PNC uptake.

**Results** Over half (59.6%, 95% confidence interval (CI) = 59.0–60.3) of mother–infant pairs received the recommended three PNC visits during the first 6 weeks; uptake was 63.1% (95% CI = 61.9–64.3) amongst HIV exposed infants and 58.1% (95% CI = 57.3–58.9) amongst HIV unexposed infants. Uptake of early PNC improved significantly with each survey, but varied significantly by province. Multivariable analysis of the pooled data, controlling for survey year, demonstrated that number of antenatal visits (4+ vs <4 Adjusted odds ratio (aOR) = 1.13, 95% CI = 1.04–1.23), timing of initial antenatal visits (≤12 weeks vs >12 weeks, aOR = 1.13, 95% CI = 1.04–1.23), place of delivery (clinic vs hospital aOR = 1.5, 1.3–1.6), and infant HIV exposure (exposed vs unexposed aOR = 1.2, 95% CI = 1.1–1.2) were the key factors associated with receiving recommended PNC visits.

**Conclusions** Approximately 40% of neonates did not receive three or more postnatal care visits in the first 6 weeks of life from 2010–2013. To improve uptake of early PNC, early antenatal booking, more frequent antenatal care attendance, and attention to HIV negative women is needed.

## Correspondence to:

Ameena Goga  
 South African Medical Research Council  
 1 Soutpansberg Road  
 Pretoria 0001  
 South Africa  
 Ameena.Goga@mrc.ac.za

Global efforts to improve the health of children under one year of age have succeeded in reducing preventable infant deaths, yet infant mortality remains high in sub–Saharan Africa, especially in settings of high HIV preva-

lence [1,2]. Neonatal deaths occur predominantly from complications with preterm births, intrapartum-related issues, sepsis, and pneumonia – all requiring medical intervention within the health system if they are to be averted [3]. Most infant deaths occur during the postnatal period from birth to six weeks, making health care interactions important during this critical time [1,3,4]. If mother to child transmission of HIV (MTCT) has not occurred during pregnancy or delivery, it can occur during the early postnatal period, if adherence to maternal triple antiretroviral therapy (ART) is inadequate during breastfeeding and maternal viral load is not suppressed [5]. Interactions with the health system during the early postnatal period can aid in averting these causes of neonatal death and reducing early HIV transmission.

South Africa, as a country with slowly reducing infant mortality and static perinatal mortality, within the context of the greatest HIV epidemic in the world, suffers in the dual and associated burdens of child mortality and HIV [6]. Although infant mortality reduced from 39 per 1000 live births in 2009 to 29 per 1000 live births in 2013, neonatal mortality rates and antenatal HIV prevalence remain stable at 11 per 1000 live births since 2012 [7], and almost one third (29.7%) of pregnant women were living with HIV in 2013 [8]. This ultimately places infant populations at greater risk for HIV acquisition and death [9].

While South Africa demonstrates commitment to adopting and implementing prevention of mother to child transmission (PMTCT) guidelines [10–12], there is a paucity of data on uptake of early postnatal care (PNC). The World Health Organization (WHO) recommends that infants receive at least three PNC visits within the first six weeks of life timed at 3 days, 7–14 days, and 6 weeks to ensure positive health outcomes [4]. The WHO recommends three PNC visits over fewer PNC visits based on evidence regarding the timing and prevalence of causes of infant mortality and morbidity and is further described in the “WHO Recommendations on Postnatal Care of the Mother and Newborn” [4]. It is important to note that the optimum number and timing of PNC visits is the subject of debate, especially in resource limited settings [13]. South Africa has not yet adopted the WHO recommendation and promotes two early PNC visits within 6 days post-delivery and 6 weeks [14]. These visits support infant and child health through delivery of immunizations and dissemination of health messages to mothers such as appropriate feeding practices and recognition of danger signs. No routine home-based PNC visits were part of national policy in South Africa, at the time of this work.

National statistics regarding early PNC focus on attendance to the PNC visit timed before 6 days post-delivery. Estimates from 2014–2015 indicate that about 74.3% of mothers and infants achieved this visit, falling slightly short of the national target of 80% [14].

This paper aims to describe the current status of population-level achievement of at least three PNC facility-based visits during the first six weeks postpartum among infants achieving the 6-week immunization visit in South Africa, and to identify associated factors. This analysis identifies areas of potential intervention to improve progress toward achieving the global WHO recommendation in the interest of optimizing neonatal health.

## METHODS

### Study population

Data for this secondary data analysis were collected through the “Evaluation of the effectiveness of the national Prevention of Mother-to-Child Transmission programme on infant HIV in South Africa” surveys conducted in 2010 (June–December 2010), 2011/12 (August 2011–March 2012), and 2012/13 (October 2012–May 2013). Public sector health facilities were sampled using multi-stage, probability proportional to size methodology, and the study was powered to produce nationally-representative results of MTCT. More detailed information about the survey is available in previous publications [10,15].

For this study, mother/caregiver-infant pairs were enrolled during the infant’s first postpartum immunization visit at six weeks since national coverage of this visit is known to be very high in South Africa [11]. Infants presenting severely ill at the facility were not included in the study, and by nature of the sampling method, infants who died before 6 weeks of age were not included. A total of 30 751 mother/caregiver-infant pairs over the three survey years were eligible for the study and completed interviews.

### Data collection

Trained fieldworkers (nurses) interviewed mothers and caregivers about socio-demographic information, infant health and feeding practices, and postnatal care. Mothers were also interviewed about their HIV

status and testing practices, and care received during the antenatal and intrapartum periods. Infant dried blood spot samples (iDBS) were taken at the time of the interview (4–8 weeks postpartum) to detect the presence of maternal HIV antibodies and infant HIV infection. Responses to the question on facility-based PNC were based on maternal recall; however, fieldworkers cross-checked information recalled by the mother with information documented in the infant's Road to Health booklet. Gestational age was abstracted from the infant Road to Health booklet.

## Data analysis

This particular analysis was restricted to 27 699 mothers (not other caregivers) with available PNC information, regardless of the availability of infant HIV test results (9278 from 2010, 9542 from 2011/12, and 8879 from 2012/13). Data were weighted to adjust for sampling methods, sample ascertainment, and South African live births. We performed frequencies to assess the proportion of infants who achieved at least 3 PNC visits in a facility within the first six weeks post-delivery vs those who had fallen short of the recommended visits. The six-week immunization visit was included as a PNC visit.

We performed frequencies to describe socio-demographic characteristics of the mother–infant pairs, as well as frequencies to examine factors related to antenatal and delivery care. HIV exposure was defined by maternal self-reported HIV status during pregnancy and labor. To determine whether variables predicting uptake of early PNC were different by HIV exposure status, we tested the associations between each variable of interest and the PNC outcome variable separately among HIV exposed infants and among HIV unexposed infants. When it was established that the same associations were seen regardless of HIV exposure status, we pooled these data. Justification was also established to pool the data from all survey years when significant changes were not seen among independent variables over time.

To estimate uptake of early PNC as an odds ratio adjusted for all covariates, we fitted a multivariable logistic regression model using purposeful selection of variables [12]. Parameters were initially included in the multivariable model if they had a Wald test result that was significant at 25% in bivariate logistic regression. Prior studies indicate that traditional cut-off levels for significance, such as the commonly used 5% significance level, fail to identify important variables [16,17]. Variables were then eliminated from the model if they were no longer significant at 10% in the multivariable model and if their removal did not change any parameter by more than 15%. Variables not significant in bivariate regression were then placed in the model and retained if they became significant when added to the multivariable model or if they changed any parameter by more than 15% [12]. All analyses were performed using STATA 13.

## Ethical consideration

The protocol for the cross-sectional survey was approved by Human Subjects Division at the United States Centers for Disease Control and Prevention within the Center for Global Health and the institutional review board of the South African Medical Research Council. Mothers provided written informed consent prior to the onset of the interview.

## RESULTS

### Population characteristics

Within the population included for this analysis, one third (30.0%, 95% CI=29.4–30.6) of infants were HIV exposed based on maternal self-reported HIV status (**Table 1**). Most infants were 6 weeks old (79.9%, 95% CI=79.3–80.4) at the time of the interview, while smaller proportions were 4–5 weeks old or 7–8 weeks old. Almost all infants were black (92.8, 95% CI=92.5–93.1).

The mean age of mothers was 26.1 years (standard deviation: 6.3) and the majority of them were single (75.5%, 95% CI=74.9–76.1). Most women reported that their highest level of education attended was grades 8–12 (79.3%, 95% CI=78.7–79.8). Over half of the women were multiparous (61.1%, 95% CI=60.4–61.7). Over half of the women (55.9%, 95% CI=55.2–56.5) were able to correctly identify all modes of MTCT.

### Uptake of early PNC

About 60% (59.6%, 95% CI=59.0–60.3) of those included in the survey achieved at least 3 PNC visits in the first 6 weeks of life, with a median of 3 PNC visits per infant (**Table 2**). Approximately 46.0% (95%

**Table 1.** Characteristics of study population, overall and by year (N=27 699)

	TOTAL N = 27 699		SURVEY YEAR 2010, n = 9278	SURVEY YEAR 2011, n = 9542	SURVEY YEAR 2012, n = 8879
	Unweighted N (%)	Weighted %, (95% CI)	Weighted % (95% CI)	Weighted % (95% CI)	Weighted % (95% CI)
<b>Infant HIV exposure</b>					
<b>Mother's self-reported HIV status (proxy for HIV exposure):</b>					
HIV unexposed	19 035 (68.72)	66.80 (66.16–67.43)	67.73 (66.59–68.85)	65.90 (64.81–66.97)	66.79 (65.70–67.86)
HIV exposed	7693 (27.77)	30.00 (29.39–30.63)	29.79 (28.69–30.92)	28.88 (27.84–29.94)	30.00 (29.39–30.63)
No response (NR)	895 (3.23)	2.95 (2.74–3.17)	2.13 (1.83–2.49)	4.96 (4.51–5.45)	2.95 (2.74–3.17)
Chose not to answer	76 (0.27)	0.25 (0.20–0.32)	0.34 (0.24–0.50)	0.26 (0.18–0.39)	0.25 (0.20–0.32)
<b>Maternal characteristics:</b>					
<b>Province</b>					
Gauteng	4673 (9.24)	26.83 (26.19–27.47)	29.56 (28.37–30.79)	27.09 (26.02–28.20)	23.88 (22.88–24.90)
Eastern Cape	2560 (9.24)	9.63 (9.27–10.01)	8.48 (7.81–9.19)	9.41 (8.85–9.99)	10.99 (10.33–11.68)
Free State	2947 (10.64)	4.96 (4.78–5.16)	4.96 (4.66–5.28)	4.77 (4.45–5.10)	5.17 (4.83–5.52)
Kwa-Zulu Natal	3356 (12.12)	21.43 (20.81–22.07)	21.17 (20.63–22.82)	21.46 (20.40–22.57)	21.13 (20.05–5.53)
Limpopo	3077 (11.11)	10.81 (10.44–11.19)	9.99 (9.37–10.64)	10.64 (10.01–11.31)	11.78 (11.14–12.44)
Mpumalanga	3211 (11.59)	7.79 (7.52–8.06)	7.64 (7.21–8.09)	7.88 (7.43–8.35)	7.84 (7.33–8.37)
Northern Cape	1227 (4.43)	2.04 (1.93–2.16)	1.69 (1.51–1.89)	2.21 (2.02–2.42)	2.22 (2.01–2.44)
North West	2949 (10.65)	7.46 (7.19–7.74)	7.44 (7.01–7.89)	7.19 (6.76–7.65)	7.75 (7.24–8.30)
Western Cape	3699 (13.35)	9.05 (8.76–9.35)	8.54 (8.06–9.04)	9.35 (8.86–9.87)	9.25 (8.74–9.80)
<b>Mean age (standard deviation)</b>	26.14 (6.29)	26.11 (6.28)	25.95 (6.18)	26.09 (6.37)	26.29 (6.30)
<b>Age category:</b>					
≤19	4188 (15.12)	15.15 (14.67–15.63)	15.47 (14.63–16.35)	15.92 (15.12–16.77)	14.05 (13.26–14.87)
20–24	8395 (30.31)	30.52 (29.91–31.13)	30.45 (29.36–31.57)	30.18 (29.15–31.22)	30.91 (29.87–31.97)
25–34	11 763 (42.47)	42.29 (41.64–42.95)	42.54 (41.36–43.73)	41.42 (40.31–42.53)	42.93 (41.82–44.05)
≥35	3282 (11.85)	11.75 (11.33–12.18)	10.63 (9.92–11.38)	12.48 (11.75–13.25)	12.11 (11.40–12.86)
NR	71 (0.26)	0.30 (0.23–0.39)	0.91 (0.69–1.19)	0 (0.00)	0 (0.00)
<b>Level of education:</b>					
Grades 1–7	3996 (14.42)	13.61 (13.17–14.06)	14.54 (13.74–15.38)	12.95 (12.23–13.71)	13.37 (12.63–14.14)
Grades 8–12	21 795 (78.69)	79.27 (78.73–79.80)	78.14 (77.15–79.09)	79.85 (78.93–80.73)	79.80 (78.89–80.68)
Completed tertiary/technical/ university	1459 (5.27)	5.60 (5.30–5.92)	5.42 (4.89–5.99)	5.75 (5.24–6.31)	5.65 (5.15–6.18)
None/NA	449 (1.62)	1.52 (1.37–1.68)	1.91 (1.63–2.25)	1.46 (1.22–1.74)	1.19 (0.97–1.46)
<b>Marital status:</b>					
Single/widowed/divorced/separated	20 393 (73.62)	75.46 (74.90–76.01)	75.36 (74.33–76.35)	75.07 (74.11–76.01)	75.96 (75.01–76.88)
Married/co-habiting	7306 (26.38)	24.54 (23.99–25.10)	24.64 (23.65–25.67)	24.93 (23.99–25.89)	24.04 (23.12–24.99)
<b>Parity:</b>					
Multiparous	16 868 (60.90)	61.08 (60.43–61.72)	59.50 (58.32–60.66)	61.43 (60.33–62.52)	62.28 (61.17–63.37)
Primiparous	10 831 (39.10)	38.92 (38.28–39.57)	40.50 (39.34–41.68)	38.57 (37.48–39.67)	37.72 (36.63–38.83)
<b>Correct identification of all MTCT modes:</b>					
Yes	15 166 (54.75)	55.86 (55.20–56.52)	46.83 (45.64–48.03)	61.39 (60.30–62.47)	59.16 (58.05–60.26)
No	12 533 (45.25)	44.14 (43.48–44.80)	53.17 (51.97–54.36)	38.61 (37.53–39.70)	44.14 (43.48–44.80)
<b>Infant characteristics:</b>					
<b>Gender:</b>					
Male	13 926 (50.28)	50.34 (49.67–51.00)	50.26 (49.07–51.46)	50.11 (48.98–51.24)	50.64 (49.51–51.77)
Female	13 773 (49.72)	49.66 (49.00–50.33)	49.74 (48.54–50.93)	49.89 (48.76–51.04)	49.36 (48.23–50.49)
<b>Age:</b>					
4–5 weeks	1959 (7.07)	6.28 (5.98–6.59)	7.83 (7.26–8.45)	7.18 (6.64–7.76)	3.85 (3.44–4.29)
6 weeks	21 753 (78.53)	79.86 (79.33–80.38)	73.64 (72.60–74.65)	80.49 (79.60–81.36)	85.32 (84.51–86.09)
7–8 weeks	3987 (14.39)	13.87 (13.42–14.32)	18.53 (17.64–19.45)	12.33 (11.61–12.08)	10.84 (10.16–11.55)
<b>Population group:</b>					
Black	24 911 (89.93)	92.79 (92.49–93.07)	92.99 (92.46–93.48)	92.33 (91.81–92.82)	93.04 (92.54–93.52)
White	135 (0.49)	0.47 (0.39–0.57)	59.32 (43.11–81.57)	0.46 (0.34–0.64)	0.35 (0.25–0.51)
Coloured	2541 (9.17)	6.27 (6.01–6.53)	5.861 (5.43–6.32)	6.69 (6.26–7.16)	6.24 (5.79–6.71)
Indian	73 (0.26)	0.36 (0.28–0.46)	0.43 (0.28–0.65)	0.40 (0.27–0.61)	0.25 (0.16–0.40)
Other	39 (0.14)	0.12 (0.08–0.16)	0.13 (0.08–0.22)	0.11 (0.05–0.21)	0.11 (0.06–0.20)

Table 1. Continued

	TOTAL N = 27 699		SURVEY YEAR 2010, N = 9278	SURVEY YEAR 2011, N = 9542	SURVEY YEAR 2012, N = 8879
	Unweighted N (%)	Weighted %, (95% CI)	Weighted % (95% CI)	Weighted % (95% CI)	Weighted % (95% CI)
<b>Gestational age from Road to Health Booklet:</b>					
≥37 weeks	18843 (68.03)	69.05 (68.44–69.65)	82.03 (81.12–82.90)	62.28 (61.19–63.35)	63.11 (62.02–64.18)
<37 weeks	3646 (13.16)	12.62 (12.19–13.06)	17.97 (17.10–18.88)	10.04 (9.40–10.72)	9.97 (9.32–10.67)
NR	5210 (18.18)	18.33 (17.84–18.83)	0 (0.00)	27.68 (26.70–28.69)	26.92 (25.95–27.91)
<b>Hospitalized within first 6 weeks of life:</b>					
No	25708 (92.81)	92.88 (92.53–93.22)	91.04 (91.34–91.69)	91.67 (91.01–92.28)	95.90 (95.41–96.35)
Yes	1966 (7.10)	7.06 (6.72–7.41)	8.88 (8.23–9.58)	8.28 (7.67–8.93)	4.04 (3.60–4.53)
NR	25 (0.09)	.06 (0.04–0.09)	0.08 (0.04–0.15)	0.05 (0.03–0.11)	0.06 (0.03–0.13)
<b>ANC and delivery characteristics:</b>					
<b>Received support from community health worker:</b>					
No	8953 (32.32)	31.51 (30.90–32.13)	38.75 (37.59–39.93)	26.35 (25.40–27.32)	29.59 (28.59–30.62)
Yes	14419 (52.06)	51.39 (50.72–52.05)	61.25 (60.07–62.41)	48.30 (47.17–49.42)	44.82 (43.69–45.95)
NR	4327 (15.62)	17.1 (16.61–17.61)	0 (0.00)	25.36 (24.36–26.37)	25.59 (24.61–26.59)
<b>ANC visits:</b>					
0–3 visits	4382 (15.82)	16.43 (15.94–16.94)	19.14 (18.20–20.13)	15.85 (15.04–16.68)	14.37 (13.60–15.17)
4–5 visits	8212 (29.65)	30.66 (30.05–31.28)	28.39 (27.13–29.49)	31.94 (30.90–33.01)	31.61 (30.57–32.67)
≥6 visits	6519 (23.54)	22.77 (22.22–23.33)	19.91 (18.99–20.86)	23.06 (22.14–24.02)	25.27 (24.29–26.28)
NR	8586 (31.00)	30.13 (29.53–30.74)	32.56 (31.47–33.67)	29.15 (28.13–30.19)	28.75 (27.73–29.78)
<b>Timeliness of first ANC visit:</b>					
≤12 weeks	7695 (27.78)	26.20 (25.63–26.78)	23.46 (22.49–24.45)	25.19 (24.24–26.17)	29.92 (28.89–30.95)
13–20 weeks	10061 (36.32)	37.25 (36.61–37.90)	33.60 (32.48–34.74)	36.90 (35.82–38.00)	41.18 (40.07–42.30)
≥21 weeks	6536 (23.60)	24.55 (23.98–25.13)	28.95 (27.86–30.06)	23.44 (22.49–24.42)	21.35 (20.44–22.29)
NR	3407 (12.30)	11.99 (11.57–12.43)	13.99 (13.19–14.84)	14.46 (13.69–15.27)	7.55 (6.99–8.15)
<b>Delivery location:</b>					
Hospital	21466 (77.50)	78.14 (77.59–78.67)	77.88 (76.88–78.85)	78.63 (77.70–79.53)	77.89 (76.94–78.80)
Clinic	4724 (17.05)	16.61 (16.13–17.10)	16.39 (15.53–17.29)	15.68 (14.88–16.50)	17.76 (16.92–18.63)
Home/other	1282 (4.63)	4.48 (4.21–4.76)	5.73 (5.20–6.30)	4.28 (3.85–4.75)	3.46 (3.06–3.90)
NR	227 (0.82)	0.78 (0.68–0.89)	0 (0.00)	1.42 (1.19–1.68)	0.90 (0.71–1.14)
<b>Delivery method:</b>					
Vaginal delivery	21424 (77.35)	76.83 (76.26–77.39)	78.51 (77.51–79.48)	75.48 (74.48–76.45)	76.55 (75.57–77.50)
Caesarean section	5957 (21.51)	22.17 (21.62–22.73)	21.42 (20.46–22.42)	21.75 (20.81–22.71)	23.33 (22.38–24.31)
Do not know	38 (0.14)	0.12 (0.08–0.17)	0.07 (0.04–0.13)	0.18 (0.10–0.30)	0.12 (0.07–0.20)
NR	280 (1.01)	0.87 (0.76–1.00)	0 (0.00)	2.60 (2.27–2.97)	0 (0.00)

NR – not recorded

CI = 44.8–47.2) of respondents from the 2010 survey, 58.2% (95% CI = 57.1–59.0) from the 2011/12 survey, and 74.4% (95% CI = 73.4–75.3) of respondents from the 2012/13 survey year achieved the three WHO recommended PNC visits. The province with the highest proportion of infants receiving at least 3 PNC was Free State (85.1%, 95% CI = 83.7–86.4). The province with the lowest proportion of infants receiving at least 3 PNC was Northern Cape (2.04%, 95% CI = 1.93–2.16). Notably, over 94% (94.36, 95% CI = 92.48–96.30) of infants at the 6-week immunization visit had achieved at least 2 early PNC visits (the current guideline in South Africa).

### Factors associated with early PNC uptake

A higher proportion of infants who were HIV exposed achieved 3 PNC (63.1%, 95% CI = 61.9–64.3), compared to HIV unexposed infants (58.1%, 95% CI = 57.3–58.9) (Table 3). This relationship persisted (aOR = 1.2, 95% CI = 1.1–1.2) when adjusted for all other covariates (Table 4).

Province of residence was highly associated with uptake of WHO recommended early PNC. Infants residing in all provinces (except Eastern Cape) were at least 30% more likely to receive 3 or more early PNC visits compared to infants in Gauteng Province, with those in Free State Province almost seven times more likely to receive recommended PNC (aOR = 6.6, 95% CI = 5.8–7.5) compared with Gauteng Province residents.

**Table 2.** Pooled analysis of uptake of  $\geq 3$  postnatal care visits in the first six weeks post delivery

Postnatal care utilization	TOTAL N = 27 699	
	Unweighted N (%)	Weighted % (95% CI)
<b>Percent of participants achieving <math>\geq 3</math> postnatal visits, by province</b>	<b><math>\geq 3</math> PNC</b>	<b><math>\geq 3</math> PNC</b>
<b>South Africa (National)</b>	<b>17 351 (62.64)</b>	<b>59.61 (58.95–60.26)</b>
Gauteng	2238 (47.89)	46.89 (45.43–48.35)
Eastern Cape	1115 (43.55)	43.94 (41.94–45.96)
Free State	2518 (85.44)	85.08 (83.70–86.37)
Kwa-Zulu Natal	2190 (65.26)	66.36 (64.70–67.98)
Limpopo	1776 (57.72)	56.81 (55.02–58.59)
Mpumalanga	2148 (66.90)	67.97 (66.31–69.59)
Northern Cape	768 (62.90)	62.92 (60.16–65.60)
North West	2330 (79.01)	81.13 (79.72–82.47)
Western Cape	2268 (61.31)	61.67 (60.08–63.24)
<b>Percent of participants achieving <math>\geq 3</math> postnatal visits, by HIV exposure</b>		
HIV unexposed	11 661 (61.25)	58.06 (57.26–58.85)
HIV exposed	5076 (65.95)	63.09 (61.86–64.30)
Unknown/Chose not to respond	349 (36.24)	59.30 (55.74–62.77)
<b>Percent of participants achieving <math>\geq 3</math> postnatal visits, by survey year</b>		
2010	5040 (54.32)	46.02 (44.84–47.20)
2011/12	5784 (60.62)	58.19 (57.07–59.30)
2012/13	6527 (73.51)	74.35 (73.37–75.31)
<b>Number of early postnatal care visits (National)</b>		
0 visits	0 (0.00)	0 (0.00)
1 visit	1230 (4.44)	5.63 (5.31–5.96)
2 visits	9118 (32.92)	34.76 (34.13–35.40)
3 visits	13 511 (48.78)	46.95 (46.29–47.62)
4 visits	2678 (9.67)	8.98 (8.61–9.36)
5+ visits	1162 (4.20)	3.67 (3.45–3.92)
<b>Average postnatal visits</b>		
Median (Quartile 1, Quartile 3)	3 (2, 3)	3 (2, 3)

PNC – postnatal care

Mothers who delivered in a clinic vs a hospital, who achieved at least four ANC visits, or who booked an ANC appointment before 12 weeks gestation were far more likely to achieve at least 3 recommended early PNC visits (aOR=1.5, 95% CI=1.3–1.6, ; aOR=1.1, 95% CI=1.04–1.2, and aOR=1.13, 95% CI=1.04–1.23, respectively). However, those who received support during pregnancy and delivery from a community health worker (CHW) were less likely than those who did not receive this support to have at least 3 early PNC visits (aOR=0.9, 95% CI=0.8–0.9). Those who delivered via Caesarean section had lower achievement of early PNC visits than those with vaginal delivery (aOR=0.7, 95% CI=0.7–0.8).

Infants who required hospitalization during the first six weeks of life were more likely than those who were not hospitalized to have 3 early PNC visits (aOR=1.6, 95% CI=1.4–1.8). Similarly, those born prematurely (before 37 weeks gestational age) were more likely to achieve the recommended 3 PNC visits than those born at full term (aOR=1.6, 95% CI=1.4–1.8). Infants classified as Non-Black (Coloured, White, Indian, or other) were more likely to achieve 3 early PNC visits compared with infants classified as Black (aOR=1.2, 95% CI=1.1–1.4).

## DISCUSSION

The WHO recommends that all mother–infant pairs should receive three postnatal care visits during the first 6 weeks of life. Our analysis of data from three nationally representative population–based surveys demonstrated that approximately 60% of infants achieving the 6–week immunization visit access the WHO recommended number of postnatal care visits in the first six weeks of life. Access to early postnatal care visits significantly increased between 2010 and 2012.

**Table 3.** Factors associated with uptake of  $\geq 3$  early PNC visits in the first 6 weeks post delivery

	UNWEIGHTED PROPORTION ACHIEVING $\geq 3$ PNC BY COVARIATE LEVEL (x/n)	WEIGHTED PERCENTAGE ACHIEVING $\geq 3$ PNC BY COVARIATE LEVEL, % (95% CI)	WEIGHTED, UNADJUSTED OR, ( $\geq 3$ PNC VISITS VS $< 3$ PNC VISITS)	95% CI
<b>Infant HIV exposure:</b>				
<b>Mother's self-reported HIV status:</b>				
HIV unexposed	11 661/19 039	58.06 (57.26–58.85)	1.00	
HIV exposed	5076/7697	63.09 (61.86–64.30)	1.23	1.16–1.31
NR/Chose not to respond*	614/963	59.30 (55.74–62.77)	1.05	0.91–1.22
<b>Maternal characteristics:</b>				
<b>Province:</b>				
Gauteng	2238/4673	46.89 (45.43–48.35)	1.00	
Eastern Cape	1115/2560	43.94 (41.94–45.96)	0.89	0.80–0.98
Free State	2518/2947	85.08 (83.70–86.37)	6.46	5.73–7.29
Kwa-Zulu Natal	2190/3356	66.36 (64.70–67.98)	2.23	2.03–2.45
Limpopo	1776/3077	56.81 (55.02–58.59)	1.49	1.36–1.64
Mpumalanga	2148/3211	67.97 (66.31–69.59)	2.40	2.19–2.64
Northern Cape	768/1227	81.13 (79.72–82.47)	1.92	1.69–2.19
North West	2330/2949	61.67 (60.08–63.24)	4.87	4.38–5.42
Western Cape	2268/3699	59.61 (58.95–60.26)	1.82	1.67–1.99
<b>Age:</b>				
$\leq 19$	2575/4188	58.74 (57.03–60.43)	1.00	
20–24	5289/8395	59.82 (58.62–61.01)	1.05	0.96–1.14
25–34	7421/11 763	60.16 (59.15–61.16)	1.06	0.98–1.15
$\geq 35$	2031/3282	58.69 (56.77–60.58)	0.99	0.90–1.11
NR	35/ 71	40.00 (27.94–53.41)	0.47	0.27–0.81
<b>Level of education:</b>				
Grades 1–7	2556/3996	60.44 (58.69–62.16)	1.00	
Grades 8–12	13 783/21 795	60.45 (59.71–61.18)	1.00	0.92–1.08
Completed tertiary/ technical/ university	743/1459	46.24 (43.44–49.05)	0.56	0.49–0.64
None/NR	269/ 449	57.68 (52.44–62.75)	0.89	0.71–1.12
<b>Marital status:</b>				
Single/widowed/divorced/separated	12873/20 393	60.56 (59.80–61.32)	1.00	
Married/co-habiting	4478/7306	56.68 (55.38–57.96)	0.85	0.80–0.91
<b>Parity:</b>				
Multiparous	10 635/16 868	60.27 (59.43–61.10)	1.00	
Primiparous	6716/10 831	58.57 (57.51–59.63)	0.93	0.88–0.99
<b>Correct identification of all MTCT modes:</b>				
Yes	9656/15 166	60.94 (60.06–61.82)	1.00	
No	7695/12 533	57.92 (56.93–58.90)	0.88	0.84–0.93
<b>Infant characteristics:</b>				
<b>Gender:</b>				
Male	8723/13 926	59.38 (58.45–60.30)	1.00	
Female	8628/13 773	59.84 (58.91–60.77)	1.02	0.97–1.08
<b>Age:</b>				
4–5 weeks	1202/1959	58.38 (55.85–60.86)	0.93	0.83–1.03
6 weeks	13 733/21 753	60.25 (59.51–60.99)	1.00	
7–8 weeks	2376/3987	56.44 (54.68–58.18)	0.85	0.79–0.92
<b>Population group:</b>				
Black	15567/24 911	59.51 (58.82–60.20)	1.00	
Non-black	1784/2788	60.83 (58.76–62.85)	1.06	0.96–1.16
<b>Gestational age from the Road to Health Booklet:</b>				
$\geq 37$ weeks	11 433/18 843	57.20 (56.39–58.00)	1.00	
$< 37$ weeks	2249/3646	57.81 (55.97–59.63)	1.03	0.94–1.11
NR	3669/5210	69.92 (68.54–71.26)	1.74	1.62–1.87
<b>Hospitalized within first 6 weeks of life:</b>				
No	15 985/25 708	59.04 (58.36–59.72)	1.00	
Yes	1351/1966	67.03 (64.59–69.38)	1.41	1.26–1.58
NR	15/ 25	58.08 (37.38–76.28)	0.96	0.41–2.23
<b>ANC and delivery characteristics:</b>				
<b>Received support from community health worker:</b>				

**Table 3.** Continued

	UNWEIGHTED PROPORTION ACHIEVING $\geq 3$ PNC BY COVARIATE LEVEL (n/n)	WEIGHTED PERCENTAGE ACHIEVING $\geq 3$ PNC BY COVARIATE LEVEL, % (95% CI)	WEIGHTED, UNADJUSTED OR, ( $\geq 3$ PNC VISITS VS $< 3$ PNC VISITS)	95% CI
No	5713/8953	62.24 (61.10–63.38)	1.00	
Yes	8837/14 419	56.08 (55.15–57.01)	0.77	0.73–0.82
NR	2801/4327	65.33 (63.80–66.84)	1.14	1.05–1.24
<b>ANC visits:</b>				
<4 visits	2571/4382	55.27 (53.61–56.92)	1.00	
$\geq 4$ visits	14 780/23 317	59.57 (58.37–60.76)	1.24	1.15–1.33
<b>Timeliness of first ANC visit:</b>				
>12 weeks	10 404/16 597	59.83 (58.98–60.67)	1.00	
$\leq 12$ weeks	5116/7695	64.69 (63.46–65.89)	1.23	1.15–1.31
NR	1831/3407	47.37 (45.49–49.26)	0.60	0.56–0.66
<b>Delivery location:</b>				
Hospital	13 047/21 466	57.53 (56.78–58.28)	1.00	
Clinic	3326/4724	68.86 (67.35–70.34)	1.63	1.51–1.76
Home/other/NR	978/1509	61.24 (58.37–64.03)	1.17	1.03–1.32
<b>Delivery method:</b>				
Vaginal delivery	13 766/21 424	61.25 (60.51–61.99)	1.00	
Caesarean section	3364/5597	53.79 (52.36–55.22)	0.74	0.69–0.79
DK/NR	221/ 318	62.02 (55.59–68.06)	1.03	0.79–1.35
<b>Survey year:</b>				
2010	5040/9278	46.02 (44.84–47.20)	1.00	
2011/12	5784/9542	58.19 (57.07–59.30)	1.63	1.53–1.74
2012/13	6527/8879	74.35 (73.37–75.31)	3.40	3.17–3.65

CI – confidence interval, PNC – postnatal care, NR – no response, ANC – antenatal care

\*Categories of “Don't know” “No response” “None” and/or “Chose not to answer” were combined in cases where one or both categories were <1%. See **Table 1** for full descriptive report.

This study was the first of its kind to assess achievement of the WHO recommendation for early PNC at a national level in South Africa, and to track changes in access to early PNC over time (2010–2013).

Our findings are consistent with the levels of early PNC uptake documented by the 2014–15 District Health Barometer which, using routine data, demonstrated 74.3% achievement of the first PNC visit within the first 6 days post-delivery [14].

Although approximately 40% of infants did not receive the required number of early PNC visits, our finding that the majority (59.6%) of infants received three or more PNC visits indicates a much higher early PNC coverage in South Africa, an upper middle income country, compared with rural Tanzania and Ghana, low income and lower middle income countries respectively [18]. These countries reported that fewer than 10% of women achieved three or more PNC [19,20]. Identification of factors associated with achieving this recommendation is necessary to inform programmatic efforts that improve coverage in South Africa and other sub-Saharan African countries.

### Factors associated with early PNC uptake

Achievement of recommended interactions with the health system along the continuum of care has been shown to arise from a combination of health system, community, household, and individual-level determinants [21–26]. This study focused on individual-level predictors such as maternal socio-demographic factors, infant HIV exposure, infant socio-demographic factors, and health seeking behavior during pregnancy and delivery.

### Province of residence

Adherence to WHO recommended early PNC was highly influenced by province of residence, with those residing in Gauteng Province less likely than those in almost every other province to achieve at least 3 early PNC visits. Eastern Cape Province was the only province that had lower achievement of at least 3 PNC visits than Gauteng Province. Populations in Gauteng and Eastern Cape Provinces are highly mobile and thus may demonstrate reduced consistency in attendance to health visits [14,27,28]. Populations

**Table 4.** Multivariable logistic regression of factors associated with uptake of  $\geq 3$  early PNC visits in the first 6 weeks post delivery\*

PARAMETER	VARIABLE	$\geq 3$ EARLY PNC VISITS VS $< 3$ PNC VISITS	
		Weighted Adjusted OR	95% CI
<b>Infant HIV exposure:</b>			
Mother's self-reported HIV status (proxy for HIV exposure)	HIV exposed vs HIV unexposed	1.16	1.10–1.23
<b>Maternal characteristics:</b>			
<b>Age (years):</b>			
	20–24 vs $\leq 19$	1.06	0.97–1.17
	25–34 vs $\leq 19$	1.11	1.01–1.21
	$\geq 35$ vs $\leq 19$	1.05	0.93–1.19
<b>Level of education</b>	Grades 8–12 vs Grades 1–7	1.05	0.96–1.14
	Completed tertiary/ technical/ university vs Grades 1–7	0.67	0.58–0.77
	None/NA vs Grades 1–7	0.92	0.72–1.17
<b>Marital status</b>	Married/co-habiting vs Single/widowed/divorced/separated	0.94	0.88–1.01
<b>Correct identification of all MTCT modes</b>	No vs Yes	0.90	0.85–0.96
<b>Province</b>	Eastern Cape vs Gauteng	0.76	0.68–0.85
	Free State vs Gauteng	6.60	5.79–7.52
	Kwa-Zulu Natal vs Gauteng	2.15	1.94–2.37
	Limpopo vs Gauteng	1.27	1.15–1.41
	Mpumalanga vs Gauteng	2.42	2.18–2.68
	Northern Cape vs Gauteng	1.55	1.33–1.80
	North West vs Gauteng	4.22	3.78–4.72
	Western Cape vs Gauteng	1.71	1.52–1.91
<b>Infant characteristics:</b>			
Population group	Non-black vs Black	1.21	1.08–1.37
Gestational age	$< 37$ weeks vs $\geq 37$ weeks	1.22	1.11–1.34
Hospitalized within first 6 weeks of life	Yes vs No	1.62	1.44–1.83
<b>ANC and delivery characteristics:</b>			
<b>Received support from community health worker (CHW)</b>	CHW support vs No CHW support	0.87	0.81–0.93
<b>ANC visits</b>	$\geq 4$ visits vs $< 4$ visits	1.13	1.04–1.23
<b>Timeliness of first ANC visit</b>	$\leq 12$ weeks vs $> 12$ weeks	1.13	1.04–1.23
<b>Delivery location</b>	Clinic vs Hospital	1.45	1.33–1.58
	Home/other vs Hospital	1.14	0.99–1.32
<b>Delivery method</b>	Caesarean vs Vaginal	0.74	0.68–0.79
<b>Survey year</b>	2011/12 vs 2010	1.61	1.49–1.73
	2012/13 vs 2010	3.49	3.21–3.79

CI – confidence interval, PNC – postnatal care, OR – odds ratio

\*All “Don't know” or “No answer” categories were not included since they are not interpretable.

with high mobility have been shown to have reduced health seeking behavior and access to care globally [29]. Such mobility may also contribute to mothers seeking PNC in provinces other than their province of residence, which was not captured in the questionnaire. Differences in uptake of early PNC observed between provinces are likely explained by a complex network of differences in relative strength of the provincial health systems, messaging surrounding the importance of PNC to infant health, the general status of health, and geography.

### HIV exposure status

An infant's exposure to HIV was found to be predictive of meeting recommended early PNC visits, with HIV exposed infants being more likely to receive at least three PNC visits than HIV unexposed infants. Since the early 2000s, South Africa has aggressively adopted each updated WHO recommendation for PMTCT, and has update postnatal care guidelines [30–35]. Programmatic PMTCT messages emphasize the importance of three or more early PNC visits for maternal and infant health [5]. The targeted nature of these messages to HIV positive mothers within PMTCT programming along the continuum of care likely contributes to the differential in PNC uptake among HIV exposed and unexposed infants. The care provided in PMTCT programs likely incentivizes HIV positive mothers to seek PNC at a higher rate than those perceiving themselves as HIV negative. On a population level this is concerning as HIV negative

women comprise the majority of the population. Similarly, HIV positive women have been shown to be more likely to practice safer infant feeding than HIV negative mothers [36].

### **Continuum of care**

Our findings show that a woman's compliance with antenatal recommendations were predictive of early PNC uptake for her infant, as seen in similar studies [19,20]. Uptake of ANC within the first 12 weeks of pregnancy and achievement of at least four ANC visits constitute the recommended timing and frequency of interaction with the health system during pregnancy [37]. This finding indicates that promotion of vital health visits taking place earlier in the continuum of care also work to encourage visits later in the continuum of care.

Receipt of support by a CHW during pregnancy, delivery, and/or postnatal care had the opposite effect on uptake of early PNC. CHW have been increasingly adopted in South Africa as a method to improve access to important health messages and services through lower-level cadres of health workers who are more mobile within the community [38,39]. Results from a randomized control trial in Western Cape, South Africa indicated improved outcomes for mothers and infants from home visits by CHW [40]. By improving health understanding among mothers and by referring infants to higher levels of care when necessary, CHWs may reduce the need for all mother–infant pairs to seek early PNC in a clinic setting [41]. We did not collect information in this national survey regarding frequency and timing of postnatal care provided by CHW. The relationship seen by this analysis indicating lower achievement of 3 early, facility-based PNC visits among those who had received antenatal or delivery care by a CHW could be explained by mothers receiving care from CHW in the early postnatal period instead of in facilities. Replacement of facility visits by CHW visits and the potential associated differences or similarities in health outcomes should be assessed separately to understand whether CHW are having the intended effect of reducing health system burdens and maintaining or improving maternal and child health in South Africa.

Women who delivered in a clinic were more likely than those who delivered in a hospital or outside of a facility to achieve three early PNC visits. While any facility delivery (hospital or clinic) has been observed in similar studies as a predictor of PNC [19,20] compared to non-facility delivery, our findings distinguished that clinic delivery led to higher likelihood of three early PNC visits than hospital delivery. While ANC services are provided at most clinics, only hospitals and maternity outpatient units (MOU) perform deliveries [31]. It is likely that many of the women who indicated delivering in a “clinic” in the current study actually delivered in a MOU and received ANC services there. These women likely returned to the MOU for PNC services. The women who delivered in a hospital were referred to the hospital by the clinic where they received ANC. It is known that 20% of ANC patients are lost in the referral process to delivery care [42], indicating that inconsistent receipt of care along the continuum may influence failure to achieve recommended early PNC. Additional costs incurred from hospital delivery, such as transport costs if the facility is far, may also deter continuation of care with early PNC visits [24-26,31].

Infants delivered through Caesarean section vs vaginal delivery were less likely to achieve 3 early PNC visits. This is surprising due to the high risk nature of Caesarean delivery, yet may be explained by longer duration in the hospital reducing the need for early PNC visits. This further supports the finding that hospital delivery was associated with lower uptake of early PNC visits.

### **Correct identification of MTCT modes and health seeking behavior**

The positive relationship between health knowledge and health seeking behavior by mothers has been shown in multiple studies [43,44] and was supported by our findings. Mothers unable to identify all three modes of MTCT (pregnancy, delivery, and breastfeeding), were less likely to comply with recommendations for early PNC than mothers who correctly identified all MTCT modalities. PMTCT messaging which empowers mothers with knowledge of MTCT modes and preventive behaviors appears to be succeeding in encouraging interaction with the health system during the early postnatal period.

Contrarily, mothers with a high level of education (tertiary/technical, or university education) were less likely than those with a low achievement of education (grade 1–7) to seek at least 3 PNC visits during the first 6 weeks. Studies examining relationships between knowledge, risk perception, and health behaviors indicate that low understanding of a disease may influence perception of high risk and, in turn, perception of high risk motivates health seeking behavior [45–47]. This could imply that mothers with low levels of education, who have low confidence in their own knowledge of health, are more compliant with health professional recommendations, while those with higher education feel confident in their under-

standing of health and seek less health care. Further, mothers with higher education may be employed at higher levels than those with lower education, leading to time conflicts that reduce health seeking.

### Infant and maternal demographic characteristics

Premature infants are at higher risk for adverse health in early infancy and thus are recommended to receive more than three early PNC visits, so this finding that premature infants achieve at least 3 visits at a higher frequency than full-term infants is consistent with expectations [4]. Our finding concerning racial differential in PNC health seeking is consistent with previous findings that children of non-black ethnicity achieve higher levels of health care than black children [48,49]. Health messaging and programming should be equitable across ethnic groups and should stress the importance of early PNC. Health care access, finances, time available, and employment status influence health care seeking and likely contribute to this racial differential.

Women aged 25–34 were more likely than adolescent women (19 and younger) to achieve recommended PNC. Related literature is inconclusive about the relationship between maternal age and health seeking behavior, however our finding is reflective of observations from studies in similar settings indicating that younger and less experienced women are less likely to adhere to recommended health care [50,51]. Attendance to at least three early PNC visits should be emphasized to all mothers regardless of age and perceived experience.

Respondents from the 2011/12 and 2012/13 survey years were much more likely to succeed in reaching the recommendation for early PNC than the 2010 survey year. This suggests a pattern of improvement in compliance with WHO recommended early PNC over time, likely due to the influence of non-governmental organizations partnering with the South African government to implement the PMTCT program. From 2008 to 2010 South Africa revised national PMTCT guidelines, adding specification of a postnatal visit within 3 days of delivery in addition to the 6 week visit specified within guidelines from 2008 and prior. This change may have enhanced health messaging focused on the postnatal period to influence an increase in attendance to early PNC.

Our results highlighted the need to reinforce messages about the importance of completing all health behavior and health visit recommendations along the continuum of care, regardless of infant HIV exposure status. Lessons should be taken from provinces achieving higher uptake of PNC (such as Free State Province and North West Province) to improve uptake in lower performing provinces (such as Gauteng Province and Eastern Cape Province). Messages about the importance of PNC should target those who do not attend at least four ANC, those who initiate ANC after 12 weeks gestation, those who deliver outside of the clinic setting, and those with demographic characteristics predictive of lower PNC utilization.

### Limitations

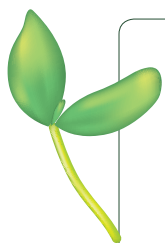
As a secondary data analysis, the original survey was not designed to specifically answer the question addressed by this analysis. By the nature of this study design, responses to the questions about health behavior were retroactive and may have introduced information bias; however there is no reason to believe that maternal memory was affected differentially between the group who met the recommended number of PNC visits and those who did not. Infants that died before 4–8 weeks and those who were severely ill were excluded from the study based on the survey sampling method, even though their cases could have been some of the most informative regarding the importance of PNC.

Only the responses of mothers were included in this study (excluding responses from other caregivers (3.2% of survey participants)). Mothers were excluded from this analysis if they did not provide information about number of PNC visits, however the proportion missing PNC responses was only 6.93% of mothers. . The potential effects of missing data were accounted for in the process of developing weights to analyze this data set as representative of the national population of mothers in South Africa. For some variables, the proportion of “Do not know” or “No answer” responses was greater than 10%, thus interpretability is limited. We did not ascertain information regarding the timing, contents, and quality of PNC visits, thus limiting the extent to which we could assess the productiveness of the care received. The location where ANC, delivery care, and PNC were received was not collected within this study, thus our ability to explain patterns in PNC uptake by district were limited to province of residence. This study focused on health care received in the formal health system and did not include information about home care or traditional care.

## CONCLUSIONS

While South Africa has not yet officially adopted the WHO recommendation for 3 early PNC visits as the national standard of practice, uptake of these important visits is high and coverage appears to have increased between 2010 and 2013. Uptake of at least 2 PNC visits, the current national standard in South Africa, is very high (over 94%), indicating strong implementation of this guideline—an encouraging prospect for potential implementation of the WHO recommendation for 3 early PNC visits. South Africa seen only mild improvements in infant survival over the past decade and has not seen improvements in neonatal survival. Adoption of the global postnatal guideline for 3 early PNC visits, or adoption of more community-based postnatal care, may be the next necessary step toward achieving optimal neonatal and infant health outcomes. Our findings suggest that efforts to encourage a woman's compliance with global recommendations along the continuum of care—such as facility-based delivery, timely initiation of ANC, and achieving 4 or more ANC visits—increase the likelihood of her infant completing the recommended number of early PNC. Messages from the health system should be strengthened during ANC, labor and delivery, and PNC to ensure that all expectant and breastfeeding mothers understand the minimum recommended visits along the continuum of care and work with health care providers to achieve them. These messages should specifically target women who have been non-adherent to at least one recommended health behavior or interaction with a facility.

As HIV prevention moves toward elimination of maternal to child HIV transmission and reduction in neonatal mortality, South Africa must strengthen the implementation of policies and programs aimed at achieving high utilization of early PNC in order to ensure the health and well-being of new generations during the postnatal period and beyond.



**Acknowledgments:** The authors thank all the women and children who participated in this study, all data collectors and supervisors, all provincial maternal and child health/PMTCT managers, all laboratory staff and members of the South Africa PMTCT Evaluation team. The authors thank the South African Medical Research Council for the support and funding to conduct this work and publish this paper.

**Disclaimer:** The findings and conclusions in this paper are those of the authors and do not necessarily represent the official position of the US Centers for Disease Control and Prevention.

**Funding:** The research was performed by the South African Medical Research Council (MRC) through a grant from the President's Emergency Plan for AIDS Relief (PEPFAR) through the Centers for Disease Control and Prevention, under the terms of Cooperative Agreement Number 5U2GPS001137-4. The South African Medical Research Council supported the salaries of some staff members, and funded this publication.

**Authorship contributions:** AG, AL, and MC conceptualized the manuscript objectives, AL carried out statistical analysis with guidance from AG, MC, GA, and TD. AL drafted the manuscript. All authors revised subsequent drafts of the manuscript and approved the final version.

**Competing interests:** The authors have completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author) and declare no conflict of interest.

## REFERENCES

- 1 You D, Wardlaw T, Salama P, Jones G. Levels and trends in under-5 mortality, 1990-2008. *Lancet*. 2010;375:100-3. Medline:19748116 doi:10.1016/S0140-6736(09)61601-9
- 2 Sartorius B, Sartorius K. Global infant mortality trends and attributable determinants – an ecological study using data from 192 countries for the period 1990–2011. *Popul Health Metr*. 2014;12:29. doi:10.1186/s12963-014-0029-6
- 3 World Health Organization. Every Newborn: an action plan to end preventable deaths. 2014. Available: <http://apps.who.int/iris/handle/10665/127938>. Accessed: 9 February 2016.
- 4 World Health Organization. Postnatal care of the mother and newborn 2013. 2013. Available: [http://apps.who.int/iris/bitstream/10665/97603/1/9789241506649\\_eng.pdf](http://apps.who.int/iris/bitstream/10665/97603/1/9789241506649_eng.pdf). Accessed: 8 October 2015.
- 5 World Health Organization. PMTCT strategic vision 2010-2015. World Health Organization: Geneva: 2010.
- 6 Baleta A. South Africa takes steps to reduce perinatal mortality. *Lancet*. 2011;377:1303-4. Medline:21504086 doi:10.1016/S0140-6736(11)60523-0
- 7 Dorrington R, Bradshaw D, Laubscher R, Nannan N. Rapid mortality surveillance report 2013. 2014. Available: <http://www.mrc.ac.za/bod/RapidMortalitySurveillanceReport2014.pdf>. Accessed: 8 October 2015.
- 8 Health Systems Trust. The 2013 National Antenatal Sentinel HIV Prevalence Survey South Africa. 2014. Available: <http://www.hst.org.za/publications/2013-national-antenatal-sentinel-hiv-prevalence-survey-south-africa>. Accessed: 13 December 2015.

- 9 Newell M-L, Coovadia H, Cortina-Borja M, Rollins N, Gaillard P, Dabis F. Mortality of infected and uninfected infants born to HIV-infected mothers in Africa: a pooled analysis. *Lancet*. 2004;364:1236-43. Medline:15464184 doi:10.1016/S0140-6736(04)17140-7
- 10 Goga A, Dinh T-H, Jackson D. Evaluation of the effectiveness of the national Prevention of Mother-to-Child Transmission (PMTCT) programme on infant HIV measured at six weeks postpartum in South Africa. South African Medical Research Council, National Department of Health South Africa and PEPFAR/US Centers for Disease Control & Prevention. 2012. Available: <http://repository.uwc.ac.za/xmlui/handle/10566/462>. Accessed: 9 February 2016.
- 11 Fadnes LT, Jackson D, Engebretsen IMS, Zembe W, Sanders D, Sommerfelt H, et al. Vaccination coverage and timeliness in three South African areas: a prospective study. *BMC Public Health*. 2011;11:404. Medline:21619642 doi:10.1186/1471-2458-11-404
- 12 Bursac Z, Gauss CH, Williams DK, Hosmer DW. Purposeful selection of variables in logistic regression. *Source Code Biol Med*. 2008;3:17. Medline:19087314 doi:10.1186/1751-0473-3-17
- 13 Sines BE, Syed U, Wall S, Worley H. Postnatal Care: A Critical Opportunity to save mothers and newborns. 2006. Available: [http://www.prb.org/pdf07/snl\\_pncbriefinal.pdf](http://www.prb.org/pdf07/snl_pncbriefinal.pdf). Accessed: 8 October 2015.
- 14 Massy N, Peer N, English R, Padarath A, Barron P, Day C, editors. District Health Barometer 2015/16. Durban: Health Systems Trust; 2016.
- 15 Goga A, Jackson D, Singh M, Lombard C. Group S study. 2012-2013 SAPMTCTE Report: Early (4-8 weeks postpartum) Population-level Effectiveness of WHO PMTCT Option A, South Africa. Available: <http://www.mrc.ac.za/healthsystems/SAPMTCTEReport2012.pdf>. Accessed: 5 February 2016.
- 16 Bendel RB, Afifi AA. Comparison of stopping rules in forward "Stepwise" regression. *J Am Stat Assoc*. 1977;72:46-53.
- 17 Mickey RM, Greenland S. The impact of confounder selection criteria on effect estimation. *Am J Epidemiol*. 1989;129:125-37. Medline:2910056 doi:10.1093/oxfordjournals.aje.a115101
- 18 The World Bank Group. Country Income Groups (World Bank Classification), Country and Lending Groups. 2011. Available: <http://data.worldbank.org/about/country-classifications/country-and-lending-groups>. Accessed: 4 March 2016.
- 19 Kanté AM, Chung CE, Larsen AM, Exavery A, Tani K, Phillips JF. Factors associated with compliance with the recommended frequency of postnatal care services in three rural districts of Tanzania. *BMC Pregnancy Childbirth*. 2015;15:341. Medline:26689723 doi:10.1186/s12884-015-0769-8
- 20 Yeji F, Shibanuma A, Oduro A, Debpuur C, Kikuchi K, Owusu-Agei S, et al. Continuum of care in a maternal, newborn and child health program in Ghana: low completion rate and multiple obstacle factors. *PLoS One*. 2015;10:e0142849. Medline:26650388 doi:10.1371/journal.pone.0142849
- 21 Somefun OD, Ibisomi L. Determinants of postnatal care non-utilization among women in Nigeria. *BMC Res Notes*. 2016;9:21. Medline:26754966 doi:10.1186/s13104-015-1823-3
- 22 Titaley CR, Hunter CL, Heywood P, Dibley MJ. Why don't some women attend antenatal and postnatal care services?: a qualitative study of community members' perspectives in Garut, Sukabumi and Ciamis districts of West Java Province, Indonesia. *BMC Pregnancy Childbirth*. 2010;10:61. Medline:20937146 doi:10.1186/1471-2393-10-61
- 23 De Schacht C, Lucas C, Mboa C, Gill M, Macasse E, Dimande SA, et al. Access to HIV prevention and care for HIV-exposed and HIV-infected children: a qualitative study in rural and urban Mozambique. *BMC Public Health*. 2014;14:1240. Medline:25467030 doi:10.1186/1471-2458-14-1240
- 24 Mohan D, Gupta S, LeFevre A, Bazant E, Killewo J, Baqui AH. Determinants of postnatal care use at health facilities in rural Tanzania: multilevel analysis of a household survey. *BMC Pregnancy Childbirth*. 2015;15:282. Medline:26518337 doi:10.1186/s12884-015-0717-7
- 25 Nankwanga A. Factors influencing utilisation of postnatal services in Mulago and Mengo Hospitals Kampala, Uganda. University of the Western Cape. Available: <http://etd.uwc.ac.za/xmlui/handle/11394/237>. Accessed: 27 September 2015.
- 26 Tlebere P, Jackson D, Loveday M, Matizirofa L, Mbombo N, Doherty T, et al. Community-based situation analysis of maternal and neonatal care in South Africa to explore factors that impact utilization of maternal health services. *J Midwifery Womens Health*. 2007;52:342-50. Medline:17603956 doi:10.1016/j.jmwh.2007.03.016
- 27 Landau LB. Gauteng 2055 Trend Paper: Population & Migration. 2008; Available: <http://migration.org.za>. Accessed: 9 February 2016.
- 28 South African Census Bureau. Statistical release (Revised) Census 2011. 2012;78. Available: [www.statssa.gov.za](http://www.statssa.gov.za). Accessed: 9 February 2016.
- 29 Gushulak BD, MacPherson DW. The basic principles of migration health: population mobility and gaps in disease prevalence. *Emerg Themes Epidemiol*. 2006;3:3. Medline:16674820 doi:10.1186/1742-7622-3-3
- 30 Bhardwaj S, Barron P, Pillay Y, Treger-Slavin L, Robinson P, Goga A, et al. Elimination of mother-to-child transmission of HIV in South Africa: Rapid scale-up using quality improvement. *S Afr Med J*. 2014;104:239. Medline:24893500 doi:10.7196/SAMJ.7605
- 31 National Department of Health of South Africa. Guidelines for Maternity Care in South Africa. Pretoria, South Africa; 2007.
- 32 National Department of Health of South Africa. Policy and Guidelines for the Implementation of the PMTC Programme. National Department of Health; Pretoria: 2008.
- 33 The National Department of Health. National Consolidated Guidelines for the Prevention of Mother-To-Child Transmission of HIV (PMTCT) and the Management of HIV in Children, Adolescents and Adults. National Department of Health; Pretoria; 2015.

- 34 National Department of Health of South Africa. The South African Antiretroviral Treatment Guidelines 2013. 2013. Available: [http://www.sahivsoc.org/upload/documents/2013 ART Guidelines-Short Combined FINAL draft guidelines 14 March 2013.pdf](http://www.sahivsoc.org/upload/documents/2013_ART_Guidelines-Short_Combined_FINAL_draft_guidelines_14_March_2013.pdf). Accessed: 25 March 2016.
- 35 National Department of Health of South Africa. The South African Antiretroviral Treatment Guidelines. 2010; Available: <http://apps.who.int/medicinedocs/documents/s19153en/s19153en.pdf>. Accessed: 25 March 2016.
- 36 Goga AE, Doherty T, Jackson DJ, Sanders D, Colvin M, Chopra M, et al. Infant feeding practices at routine PMTCT sites, South Africa: results of a prospective observational study amongst HIV exposed and unexposed infants - birth to 9 months. *Int Breastfeed J*. 2012;7:4. Medline:22472507 doi:10.1186/1746-4358-7-4
- 37 World Health Organization. Pregnancy, childbirth, postpartum and newborn care: a guideline for essential practice. World Health Organization: Geneva; 2006.
- 38 de Wet K, Wouters E, Engelbrecht M. Exploring task-shifting practices in antiretroviral treatment facilities in the Free State Province, South Africa. *J Public Health Policy*. 2011;32 Suppl 1:S94-101. Medline:21730997 doi:10.1057/jphp.2011.30
- 39 National Department of Health of South Africa. National Department of Health Strategic Plan 2010/11-2012/13. Department of Health; Pretoria; 2014.
- 40 le Roux IM, Tomlinson M, Harwood JM, O'Conner MJ, Worthman CM, Mbewu N, et al. Outcomes of home visits for pregnant mothers and their infants: a cluster randomised controlled trial. *AIDS*. 2013;27:1461-71. Medline:23435303 doi:10.1097/QAD.0b013e3283601b53
- 41 Adam MB, Dillmann M, Chen M, Mbugua S, Ndung'u J, Mumbi P, et al. Improving maternal and newborn health: effectiveness of a community health worker program in rural Kenya. *PLoS One*. 2014;9:e104027. Medline:25090111 doi:10.1371/journal.pone.0104027
- 42 Beauclair R, Petro G, Myer L. The association between timing of initiation of antenatal care and stillbirths: a retrospective cohort study of pregnant women in Cape Town, South Africa. *BMC Pregnancy Childbirth*. 2014;14:204. Medline:24923284 doi:10.1186/1471-2393-14-204
- 43 Nabukera SK, Witte K, Muchunguzi C, Bajunirwe F, Batwala VK, Mulogo EM, et al. Use of postpartum health services in rural Uganda: knowledge, attitudes, and barriers. *J Community Health*. 2006;31:84-93. Medline:16737170 doi:10.1007/s10900-005-9003-3
- 44 Mpembeni RN, Killewo JZ, Leshabari MT, Massawe SN, Jahn A, Mushi D, et al. Use pattern of maternal health services and determinants of skilled care during delivery in Southern Tanzania: implications for achievement of MDG-5 targets. *BMC Pregnancy Childbirth*. 2007;7:29. Medline:18053268 doi:10.1186/1471-2393-7-29
- 45 Ferrer R, Klein WM. Risk perceptions and health behavior. *Curr Opin Psychol*. 2015;5:85-9. Medline:26258160 doi:10.1016/j.copsyc.2015.03.012
- 46 Becker M. The Health Belief Model and Sick Role Behavior. *Health Educ Behav*. 1974;2:409-19.
- 47 Brewer NT, Chapman GB, Gibbons FX, Gerrard M, McCaul KD, Weinstein ND. Meta-analysis of the relationship between risk perception and health behavior: The example of vaccination. *Health Psychol*. 2007;26:136-45. Medline:17385964 doi:10.1037/0278-6133.26.2.136
- 48 Burgard SA, Treiman DJ. Trends and racial differences in infant mortality in South Africa. *Soc Sci Med*. 2006;62:1126-37. Medline:16135394 doi:10.1016/j.socscimed.2005.07.025
- 49 South Africa Demographic and Health Survey 2003. Available: <https://dhsprogram.com/pubs/pdf/FR206/FR206.pdf>. Accessed: 28 September 2015.
- 50 Dhakal S, Chapman GN, Simkhada PP, van Teijlingen ER, Stephens J, Raja AE. Utilisation of postnatal care among rural women in Nepal. *BMC Pregnancy Childbirth*. 2007;7:19. Medline:17767710 doi:10.1186/1471-2393-7-19
- 51 Magadi MA, Madise NJ, Rodrigues RN. Frequency and 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



# Quality of antenatal care service provision in health facilities across sub-Saharan Africa: Evidence from nationally representative health facility assessments

Mufaro Kanyangarara, Melinda K Munos, Neff Walker

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

**Background** Utilization of antenatal care (ANC) services has increased over the past two decades. Continued gains in maternal and newborn health will require an understanding of both access and quality of ANC services. We linked health facility and household survey data to examine the quality of service provision for five ANC interventions across health facilities in sub-Saharan Africa.

**Methods** Using data from 20 nationally representative health facility assessments – the Service Provision Assessment (SPA) and the Service Availability and Readiness Assessment (SARA), we estimated facility level readiness to deliver five ANC interventions: tetanus toxoid vaccine for pregnant women, intermittent preventive treatment for malaria in pregnancy (IPTp), syphilis detection and treatment in pregnancy, iron supplementation and hypertensive disease case management. Facility level indicators were stratified by health facility type, managing authority and location, then linked to estimates of ANC utilization in that stratum from the corresponding Demographic and Health Surveys (DHS) to generate population level estimates of the ‘likelihood of appropriate care’. Finally, the association between estimates of the ‘likelihood of appropriate care’ from the linking approach and estimates of coverage levels from the DHS were assessed.

**Findings** A total of 10 534 health facilities were surveyed in the 20 health facility assessments, of which 8742 reported offering ANC services and were included in the analysis. Health facility readiness to deliver IPTp, iron supplementation, and tetanus toxoid vaccination was higher (median: 84.1%, 84.9% and 82.8% respectively) than readiness to deliver hypertensive disease case management and syphilis detection and treatment (median: 23.0% and 19.9% respectively). Coverage of at least 4 ANC visits ranged from 24.8% to 75.8%. Estimates of the likelihood of appropriate care derived from linking health facility and household survey data showed marked gaps for all interventions, particularly hypertensive disease case management and syphilis detection and treatment. There was fairly good concordance between our estimates of high likelihood of appropriate care and DHS estimates of coverage for iron supplementation, IPTp, and tetanus toxoid vaccination.

**Conclusion** Linking household surveys to health facility assessments revealed marked gaps in population-level coverage of quality ANC interventions and underscored the need for a double-pronged approach to increase ANC utilization and improve the quality of ANC services.

## Correspondence to:

Mufaro Kanyangarara  
 Institute for International Programs  
 Johns Hopkins Bloomberg School of Public Health  
 615 North Wolfe Street  
 Baltimore, MD 21205, USA  
 mknyan1@jhu.edu

High quality antenatal care (ANC) can substantially reduce maternal and newborn mortality [1,2]. High population-level coverage of ANC is necessary to deliver proven interventions to improve maternal, newborn and child health (MNCH). Globally, progress to achieve universal ANC coverage is monitored by tracking two indicators: 1) the proportion of women of reproductive age who report at least one ANC visit with a skilled health provider during the most recent live birth (ANC 1+); and 2) the proportion of women of reproductive age who report at least four ANC visits with any provider during the most recent live birth (ANC 4+) [3,4]. These two indicators quantify the number of contacts pregnant women in low- and middle-income countries have with the existing health infrastructure. Contact with the health system does not guarantee the receipt of all or any of the necessary ANC interventions. However, in theory, the more points of contact that a woman has with a well-functioning and effective health system during pregnancy, the greater the opportunity and higher the likelihood that potential pregnancy complications will be prevented, detected, and treated in a timely manner.

Uptake of ANC services has steadily increased over the past two decades according to trends from population-based surveys. Countdown to 2015, a multi-stakeholder collaboration tracking country-level coverage for MNCH, reported that the majority of women (90%) have at least one ANC visit, while only 57% have the recommended four or more ANC visits [5]. As ANC utilization increases, simply reporting on ANC1+ and ANC4+ is insufficient to help countries achieve their maternal health objectives. A country may have high levels of ANC4+ but few women receiving the recommended ANC interventions if the quality of ANC services is poor. Conversely, better service provision is unlikely to improve health outcomes if uptake is low. Service provision and service utilization have often been considered in isolation of each other, providing a fragmented picture of the effectiveness of interventions within a health system. Continued gains in maternal and newborn health will require countries to measure and link both ANC contacts and content in order to inform improvements in access to and quality of ANC services [4,6,7].

Large scale cross-sectional household surveys such as the Demographic and Health Surveys (DHS) and the Multiple Indicator Cluster Surveys (MICS) provide measures of ANC1+ and ANC4+, as well as limited information about physical examinations conducted during ANC [8,9]. These surveys ask women about whether their height, weight and blood pressure were ever measured during ANC, whether urine and blood samples were collected, whether tetanus toxoid vaccination and intermittent preventive treatment of malaria in pregnancy (IPTp) were given, and whether women received counselling on danger signs for pregnancy complications. Multi-country assessments conducted to evaluate the content of ANC visits have found that estimated coverage of these components of ANC was lower than expected based on ANC1+ and ANC4+ coverage due to poor quality of care [5,6,10,11]. An analysis of DHS data from 41 countries showed that the coverage of eight recommended ANC components ranged widely and averaged only 60% among women who reported attending at least four ANC visits [6]. Population-based household surveys may adequately monitor the number of ANC visits and the receipt of simple procedures such as the collection of urine and blood specimens for screening. However, it is unclear whether respondents can correctly recall more complex procedures, and details about examinations or laboratory investigations conducted during one or more ANC visits. The scope and type of ANC interventions that can be reliably tracked using household surveys may be considerably limited. Additionally, the survey reference period often spans up to five years, which calls into question issues related to measurement error or recall bias. Combining service utilization data from household surveys with service provision data from health facility assessments may be a useful strategy to improve coverage measurement for MNCH [12–14].

Health facility assessments are sample surveys or censuses of health facilities designed to assess and monitor the status of the health system. They typically assess service availability (the physical presence or reach of health facilities) and service readiness (the capacity to deliver the services offered) [15]. The availability of trained and supervised staff, guidelines, equipment, diagnostics, medicines and commodities, determined through both self-report and direct observation, form the basis of monitoring and evaluating service-specific readiness [16]. The underlying assumption is that the availability of these elements is a prerequisite for the delivery of high-quality care, and that examining readiness is critical for understanding service quality. Although several previous studies have explored the readiness of health facilities to provide services for ANC [17–19], child health [20], tuberculosis treatment [21], family planning [22,23] and essential medicines [24], only a few have linked estimates of readiness to population-based care-seeking estimates from household surveys to estimate coverage of MNCH interventions [25–27].

The objective of this study was to characterize the likelihood of appropriate care for different ANC interventions across sub-Saharan Africa by linking data from nationally representative health facility assessments to nationally representative household surveys. This study also highlighted the feasibility and methodological challenges of linking health facility assessments to household surveys.

## METHODS

### Data sources

This study relied on three primary data sources: the Service Provision Assessments (SPA), the Service Availability and Readiness Assessments (SARA) and the Demographic and Health Surveys (DHS). The SPA survey, developed by ICF International, is a cross-sectional health facility assessment that aims to systematically review the status of health systems by examining infrastructure, equipment, and supplies considered necessary to provide quality health services [28]. The sampling frame for the survey is the national master facility list, a complete listing of all public and often non-public health facilities in the formal sector ranging from primary to tertiary level of care. Tertiary level hospitals are often oversampled. Typically, SPA surveys adopt a sampling design that allows the calculation of nationally representative indicators by region, health facility type, and managing authority. In countries with a limited number of health facilities or where the budget permits, a full census approach has been used. The core data collection tool, a facility inventory questionnaire, is used for the direct verification of the availability of drugs, commodities, equipment and amenities. Details about the survey design and survey procedures are described in the final survey reports issued by the respective countries. Full survey data sets are available in the public domain [29].

Similar to the SPA, the SARA is a nationally representative assessment of health facilities on the availability and readiness to provide essential health services. The SARA tool was developed in 2008 by a collaboration of the World Health Organization (WHO) and the United States Agency for International Development (USAID) [16]. Building on experiences and best practices from other health facility assessments including the SPA, the SARA was designed to serve as a rapid health facility assessment tool to track quality of health services annually and inform national health planning. The SARA framework is based on a common set of indicators and summary measures formulated to detect changes and monitor progress in health systems strengthening, while enabling comparisons across facilities. Often, random stratified sampling of health facilities by health facility type and managing authority is used to select a nationally representative sample from a national master facility list. Detailed descriptions of the survey process, methodology and data collection instruments are available online [30].

The DHS is a nationally representative household cluster survey used to collect data on sociodemographic characteristics, family planning, fertility, and maternal and child health. The women's module, administered to women aged 15–49 years, collects details on whether ANC was sought and the source of ANC (facility type) for the most recent pregnancy that led to a live birth within the five years preceding the survey. Women also report on the basic components of ANC received, including weight and height measurement, blood pressure measurement, blood and urine specimen collection, iron supplementation, malaria prophylaxis in malaria endemic countries, tetanus toxoid vaccination, and counselling about potential pregnancy complications. To limit recall bias, our analysis was restricted to the sample of women aged 15–49 years who reported a live birth in the three year period preceding the index survey. Similar to the DHS, MICS collect information detailing ANC services received. However, MICS data sets were excluded from this analysis because they do not collect data on the type of health facility where women sought ANC, which is necessary for the linkage.

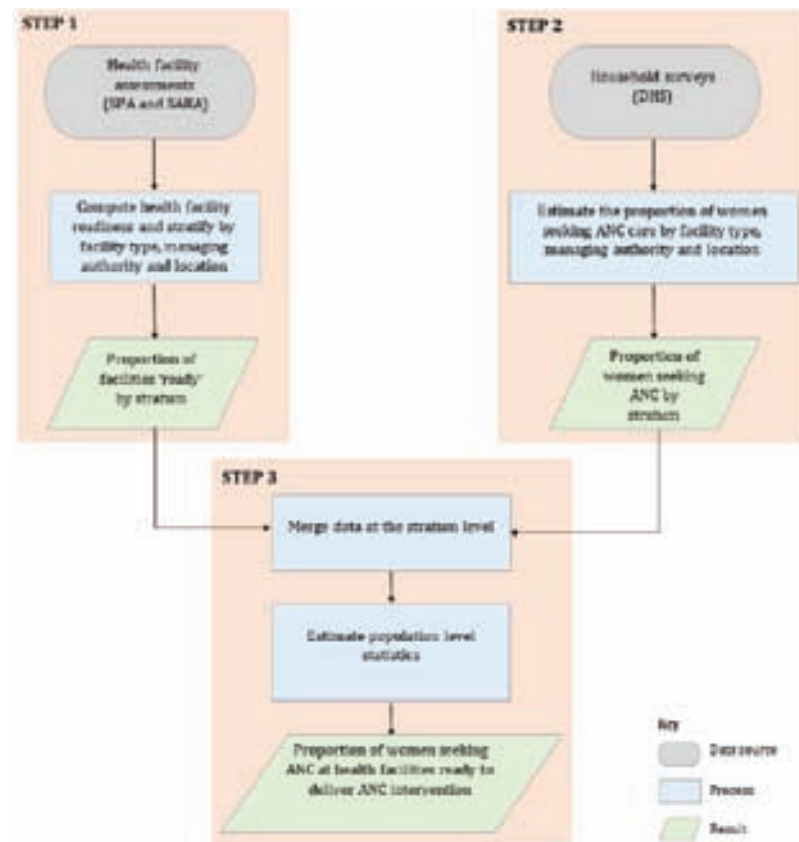
Our analysis included countries in sub-Saharan Africa with at least one SPA or SARA survey conducted in or after 2000, and with a corresponding DHS conducted within 2 years (prior to or after the SPA/SARA). Multiple health facility surveys were available and included for several countries (Democratic Republic of Congo (DRC), Kenya, Senegal, Sierra Leone, Tanzania and Uganda).

### Analysis

The analysis used a three step process to link health facility data from SPAs and SARAs to DHS household survey data (Figure 1). In Step 1, five key ANC interventions were chosen based on the availability of information on the particular intervention components which could be standardized across both the SPA and SARA. The ANC interventions included: (i) tetanus toxoid vaccine for pregnant women; (ii) intermittent preventive treatment of malaria in pregnancy (IPTp); (iii) syphilis detection and treatment in pregnancy; (iv) hypertensive disease case management, including management of pre-eclampsia with magnesium sulphate; and (v) iron supplementation. We defined the components of each intervention and made adjustments to align definitions across the SPAs and SARAs (Table 1). While coverage of prevention of mother-to-child transmission of HIV (PMTCT) services is important in the sub-Saharan Africa

context, the intervention was excluded as most of the health facility surveys only assessed whether ANC facilities offered PMTCT services, and no further verification of the availability of components of PMTCT was conducted.

For each intervention, indicators of readiness to provide the intervention on the day of the health facility assessment were calculated for each health facility that reported offering ANC services. In facilities with multiple pharmacies or medical storage rooms, a medicine was considered available if at least one unit



**Figure 1.** Three-step process to link health facility data (SPA and SARA) to household survey data (DHS).

**Table 1.** Definition of indicators of health facility readiness to deliver key antenatal care (ANC) interventions

DOMAIN	INDICATOR	DEFINITION OF INDICATORS
Staff and guidelines	Staff trained in ANC	• At least one staff member trained in at least one aspect of ANC
	Guidelines on ANC	• Observed or reported the availability of ANC guidelines
Equipment, diagnostics, medicines and commodities	Tetanus toxoid vaccine for pregnant women	• Observed at least one valid unexpired unit of tetanus toxoid vaccine
	Intermittent preventive treatment of malaria in pregnancy	• Observed at least one valid unexpired unit of sulphadoxine/pyrimethamine
	Syphilis detection and treatment in pregnancy	• Observed at least one valid syphilis test <sup>†</sup> • Observed at least one valid unexpired unit of medicine to treat syphilis <sup>‡</sup>
	Hypertensive disease case management	• Observed at least one valid dipstick for urine protein OR acetic acid and flame for heating • Observed at least one functioning blood pressure apparatus <sup>§</sup> • Observed at least one valid unexpired unit of magnesium sulphate
	Iron supplementation	• Observed at least one valid unexpired unit of iron or iron and folic acid tablets

\*Data on receipt of training of ANC services had varying recall periods ranging from 1–3 years preceding the survey.

<sup>†</sup>Syphilis rapid diagnostic test (RDT), Venereal Disease Research Laboratory (VDRL) test or polymerase chain reaction (PCR) or rapid plasma reagin (RPR).

<sup>‡</sup>Injectable benzathine penicillin or procaine penicillin.

<sup>§</sup>Automatic or manual with stethoscope.

with a valid expiration date was observed in any location. Similarly, diagnostic services and equipment were considered available if they were available and functioning in at least one location or service area. Missing values were treated as indicating the unavailability of the component of interest and accounted for <1% of the sample size for all interventions by country.

Facility-level indicators were stratified by health facility type (hospital/ health center/ health post, etc.), managing authority (public/non-public) and location (urban/rural) to obtain stratum-specific proportions of ANC facilities 'ready' to provide each intervention. These stratification variables were selected to account for hypothesized differences in the service environment by health facility type and heterogeneity in health care seeking behaviors by location (urban/rural). Information on the location stratum was missing for SPA surveys conducted prior to 2009, in which case only health facility type and managing authority were used for stratification.

In Step 2, among women aged 15–49 years who reported attending at least four ANC visits (ANC4+) in the DHS survey, source of ANC was categorized as above. For women providing multiple responses about where they sought ANC, only the highest level facility type was included. In the final step, the proportion of women attending ANC4+ in each facility stratum was merged with summary statistics of readiness derived from the SPA/SARA, using facility stratum as the identifier. For each intervention, the expected proportion of women attending ANC at health facilities 'ready' to provide the intervention at the national level was estimated by summing, across strata, the proportion of ANC facilities 'ready' to provide the intervention in that stratum (from the SPA/SARA) multiplied by the proportion of women who reported attending at least four ANC visits in that stratum (derived from the corresponding DHS).

We categorized women attending at least four ANC visits into three groups based on the "likelihood of appropriate care": (i) high likelihood of appropriate care if they attended a health facility type that in the SPA/SARA had the necessary equipment, diagnostics, medicines and commodities in stock, was equipped with ANC guidelines, and had at least one staff member who had been trained recently; (ii) moderate likelihood of appropriate care if they attended a health facility type that in the SPA/SARA had the necessary equipment, diagnostics, medicines and commodities in stock, but the health facility type lacked ANC guidelines or trained staff members (or both); and (iii) low likelihood of appropriate care if they attended a health facility type that in the SPA/SARA did not have the necessary equipment, diagnostics, medicines and commodities in stock, regardless of the availability of ANC guidelines or trained staff members. This classification encompasses all four domains that are required for service specific readiness within the SARA framework – availability of guidelines and trained staff, equipment, diagnostics, and medicines and commodities [16].

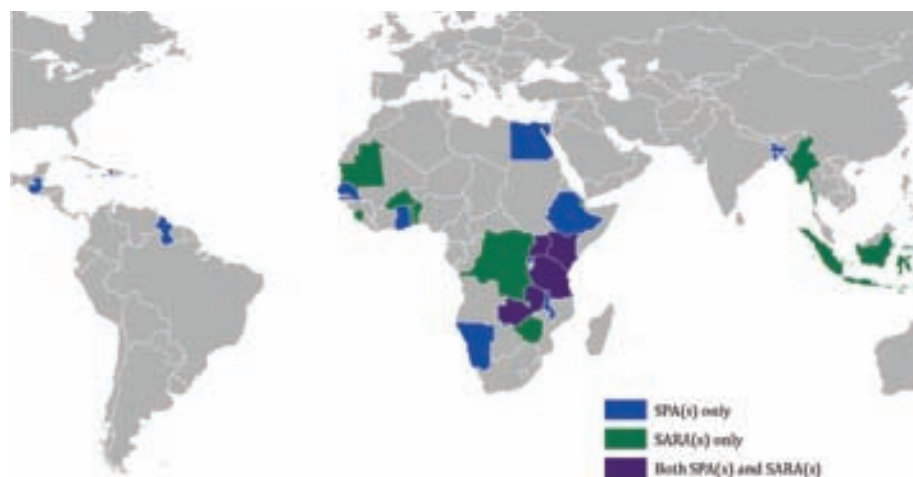
We conducted a sensitivity analysis to check the robustness of the overall results to misclassification of the source of ANC. Two hypothetical cases were considered. The upper bound corresponded to the hypothetical scenario in which all women attending at least four ANC visits were assumed to have sought ANC at health facilities in the stratum with the highest level of readiness for each intervention. Conversely, the lower bound was obtained by assuming all women attending at least four ANC visits sought care at health facilities in the stratum with the lowest level of readiness.

Finally, to assess the plausibility of our results, we compared our results to self-reported receipt of ANC interventions for those interventions where DHS surveys are used to estimate coverage based on maternal recall (tetanus toxoid vaccine, IPTp, and iron supplementation).

All analyses accounted for the complex survey design and adjusted for sampling weights. Analyses were conducted using STATA 14 (College Station, TX).

## RESULTS

A total of 44 health facility assessments were identified, with the majority (36/44, 82%) in sub-Saharan Africa (**Figure 2**). Of the 36 health facility assessments conducted in sub-Saharan Africa, 1 was excluded as it was conducted prior to 2000, 9 were excluded as there was no corresponding DHS within two years of the health facility assessment, and 6 were excluded as they did not contain information on ANC services. The final sample included 20 assessments that reported on a total of 10 534 health facilities, of which 8742 (82.9%) reported offering ANC services. The percentage of facilities offering ANC services ranged from 73.9% to 95.3% across the 20 health facility assessments (**Table 2**).



**Figure 2.** Geographic distribution of Service Provision Assessment (SPA) and Service Availability and Readiness Assessment (SARA) surveys.

**Table 2.** Indicators of health facility readiness to deliver key antenatal care (ANC) interventions from the SPAs and SARAs\*

COUNTRY	YEAR	SURVEY	TOTAL NUMBER OF HEALTH FACILITIES	FACILITIES OFFERING ANC SERVICES (%)	ANC FACILITIES WITH STAFF TRAINED IN ANC (%)	ANC FACILITIES WITH GUIDELINES ON ANC (%)	PERCENTAGE OF ANC FACILITIES WITH THE EQUIPMENT, DIAGNOSTICS, MEDICINES AND COMMODITIES TO DELIVER				
							TT	IPTp	Syph T&T	Hyp	Fe
Benin	2013	SARA	189	83.6	49.9	70.8	68.8	78.5	34.9	34.9	96.9
Burkina Faso	2012	SARA	686	88.8	71.8	94.2	91.5	51.5	4.6	15.4	88.0
DRC	2013	SARA	317	91.8	62.3	75.7	28.0	97.0	11.3	3.6	62.4
DRC	2014	SARA	1555	73.9	45.3	47.3	26.3	63.3	34.3	6.9	59.7
Ghana	2002	SPA	428	87.1	53.8	55.7	77.4	37.4	3.1	65.2	81.9
Kenya	2004	SPA	440	79.3	68.9	33.2	83.6	96.6	41.4	4.2	86.2
Kenya	2010	SPA	695	80.7	78.5	63.5	85.4	95.1	40.7	27.7	52.3
Namibia	2009	SPA	411	73.7	59.2	64.1	97.6	86.8	92.7	50.7	93.9
Rwanda	2007	SPA	538	80.3	79.8	31.2	93.5	90.8	52.1	2.0	78.1
Senegal	2013	SPA	438	78.1	42.6	62.5	79.6	76.6	1.4	22.9	82.7
Senegal	2014	SPA	452	77.0	46.8	62.0	87.6	89.6	4.5	33.0	93.5
Sierra Leone	2011	SARA	207	89.9	96.6	60.3	81.9	79.6	2.9	19.9	83.5
Sierra Leone	2012	SARA	106	91.5	96.3	86.8	81.7	67.4	3.9	46.1	78.1
Sierra Leone	2013	SARA	455	94.1	88.9	85.0	89.0	93.5	5.4	49.0	90.4
Tanzania	2006	SPA	611	82.2	61.6	50.5	83.8	84.9	42.8	3.7	80.1
Tanzania	2014/15	SPA	1200	85.9	30.9	60.7	88.0	61.1	51.5	32.0	94.9
Togo	2012	SARA	100	92.0	31.4	77.3	80.7	87.8	13.1	10.3	93.5
Uganda	2007	SPA	491	81.3	52.1	32.7	70.1	84.4	26.7	9.3	55.6
Uganda	2012	SARA	95	85.3	64.8	59.6	76.0	83.8	30.9	23.1	86.6
Zimbabwe	2014	SARA	275	95.3	99.2	96.2	91.2	59.9	82.5	71.8	96.2
Median			439	84.5	62.0	62.3	82.8	84.1	19.9	23.0	84.9

TT – tetanus toxoid vaccine for pregnant women, IPTp – intermittent preventive treatment of malaria in pregnancy, Syph T&T – syphilis detection and treatment in pregnancy, Hyp – hypertensive disease case management, Fe – iron supplementation

\*Source: Service Provision Assessment (SPA) and Service Availability and Readiness Assessment (SARA) surveys.

## Health facility readiness

Among facilities offering ANC services, the percentage with at least one staff member who had recently received in-service training in ANC varied from 30.9% in Tanzania (2014/15) to 99.3% in Zimbabwe (2014). For ANC guidelines, ANC facilities in Zimbabwe (2014) were most likely (96.2%) to report the availability of ANC guidelines compared to less than a quarter (31.2%) of ANC facilities in Rwanda (2007) (Table 2). Overall, the percentage of facilities with the necessary equipment, diagnostics, medicines and commodities to deliver IPTp, iron supplementation, and tetanus toxoid vaccination was relatively high. The availability of IPTp drugs exceeded 90% in ANC facilities in DRC (2013), Kenya (2004, 2010), Rwanda (2007), and Sierra Leone (2013). Similarly, more than 90% of ANC facilities in Benin (2013), Namib-

ia (2009), Senegal (2014), Sierra Leone (2014), Tanzania (2014/15), Togo (2012), and Zimbabwe (2014) had iron tablets in stock. By contrast, the availability of equipment, diagnostics, medicines and commodities to deliver hypertensive disease case management and syphilis detection and treatment was relatively low. None of the assessments reported more than 75% of ANC facilities in a country with the capacity to manage hypertensive diseases in pregnant women (median: 23.1%; range: 2.0–71.8%). With the exception of Namibia (2009) and Zimbabwe (2014), less than 90% of ANC facilities had a valid syphilis test and penicillin treatment available on the day of assessment, with a low of 9.3% recorded in Burkina Faso (2012).

### ANC coverage and content

The percentage of women attending at least one ANC visit from a skilled provider (ANC1+) was generally high (median: 93.1%; range: 86.0–97.5%) (Table 3). The percentage of women attending at least four ANC visits (ANC4+) ranged from 24.8% in Rwanda (2007/8) to 75.8% in Sierra Leone (2013). Most women first sought care during the second trimester, with the median months of gestation at first ANC visit ranging from 3.7 to 5.9 months. Based on DHS data, more than half of women who attended at least one ANC visit reported blood pressure measurement (median: 93.9%; range: 51.3–99.1%) and receipt of at least one dose of tetanus toxoid vaccine (median: 82.1%; range: 56.7–97.3%). However, the collection of a blood sample (median: 82.7%; range: 28.4–89.3%), the collection of a urine sample (median: 66.9%; range: 11.4–94.1%), the receipt of iron supplementation (median: 79.6%; range: 41.7–94.8%), and the receipt of at least one dose of SP during an ANC visit (median: 41.1%; range: 0.8–73.5%) was much more variable (Table 3).

### Likelihood of appropriate care

Based on the results from linking the SPA/SARA and DHS data, we estimated that on average around one in five women with a recent live birth received ANC at a facility with trained staff and appropriate clinical guidelines and IPTp drugs (20.8%), iron supplements (22.5%), or tetanus toxoid vaccine (21.5%) in stock (Table 4). Despite the availability of these three drugs at health facilities, estimates of the likelihood of appropriate care were lower due to underutilization of ANC services, lack of ANC guidelines, and a shortage of trained staff to deliver interventions (Figure 3). For example, in Rwanda (2007), more than three quarters of facilities offering ANC services had SP (90.8%), iron supplements (78.1%), or tetanus

**Table 3.** DHS estimates of coverage of antenatal care (ANC) and ANC interventions among women who reported attending at least one ANC visit\*

YEAR	ANC1+ (%)	ANC4+ (%)	EARLY ANC ENROLLMENT (%)†	MEDIAN MONTHS OF PREGNANCY AT 1 <sup>ST</sup> ANC VISIT	BLOOD SAMPLE TAKEN (%)	URINE SAMPLE TAKEN (%)	BLOOD PRESSURE MEASURED (%)	WEIGHT MEASURED (%)	IRON TABLETS OR SYRUP GIVEN (%)	ANY OR COMPLETE TETANUS PROTECTION AT BIRTH (%)	ANY SP / FANSIDAR USE DURING ANC VISIT (%)	
Benin	2011/12	86.0	58.3	48.3	3.8	81.3	94.1	97.7	98.3	80.8	78.2	37.3
Burkina Faso	2010	95.5	33.1	40.6	4.3	63.3	84.5	97.2	98.2	93.3	91.0	–
DRC	2013/14	88.5	47.2	16.7	5.4	61.4	52.0	74.1	88.2	58.7	75.0	33.0
Ghana	2003	91.9	68.9	45.8	4.0	85.8	83.7	95.4	94.3	79.4	85.1	0.8
Kenya	2003	87.4	50.9	10.5	5.9	56.4	49.1	82.3	90.9	46.8	85.5	9.9
Kenya	2008/9	91.5	45.7	13.5	5.8	82.7	66.6	84.2	94.0	69.6	88.5	33.6
Namibia	2006/7	94.3	70.2	31.6	4.8	97.3	92.8	97.1	98.0	79.6	56.7	19.6
Rwanda	2007/8	96.1	24.8	23.2	5.3	73.6	17.9	87.0	97.2	41.7	74.1	51.3
Senegal	2012/13	94.2	45.8	54.0	3.7	80.3	86.7	99.1	95.4	93.6	91.2	73.5
Senegal	2014	96.1	47.1	57.8	3.6	85.5	89.3	99.6	95.4	94.8	90.7	71.1
Sierra Leone	2013	97.5	75.8	44.7	4.1	89.3	72.4	93.9	88.7	94.1	97.3	62.1
Tanzania	2004/5	92.3	59.2	13.6	5.5	52.6	39.7	64.0	93.8	61.0	78.8	49.5
Tanzania	2015/16	97.9	49.2	23.1	5.0	87.3	58.6	69.5	93.8	81.9	74.4	68.3
Togo	2013/14	91.9	56.4	27.0	4.9	87.4	88.4	97.0	N/A	85.8	85.5	58.7
Uganda	2006	93.1	45.9	16.2	5.6	28.4	11.4	51.3	76.5	63.2	76.1	33.5
Uganda	2011	94.9	47.2	20.7	5.2	83.4	22.0	58.0	77.5	75.8	80.9	44.9
Zimbabwe	2015	92.0	74.1	37.4	4.5	98.7	66.9	97.3	94.0	83.9	82.1	13.3
Median		93.1	49.2	27.0	4.9	82.7	66.9	93.9	94.0	79.6	82.1	41.1

N/A – no data available

\*Reference period: 3 years.

†Early ANC enrollment was defined as first ANC visit before 16 weeks of gestation.

toxoid vaccine (93.5%) in stock (Table 2). However, ANC 4+ coverage was 24.8% (Table 3), and only 31.2% of ANC facilities had ANC guidelines (Table 2). As a result, very few women had a high likelihood of appropriate care for these three interventions (IPTp: 5.6%, iron supplementation: 4.8%, tetanus toxoid vaccination: 5.8%) (Figure 3, Table 4).

Across countries, the percentage of women attending ANC at facilities ready to perform hypertensive disease case management and syphilis detection and treatment was lower relative to the other interventions, representing substantial missed opportunities for women seeking ANC care. For example, in Sierra Leone (2013) where there was the most widespread utilization of ANC services (ANC4+ coverage: 75.8%) (Table 3), only 12.0% of all women with a recent live birth had a high likelihood of appropriate care for syphilis detection and treatment. In countries with multiple health facility assessments (DRC, Kenya, Uganda, Senegal, Sierra Leone and Tanzania), trends over time suggested that there were improvements in the availability of diagnostics for syphilis detection and treatment (Table 2), and subsequent improve-

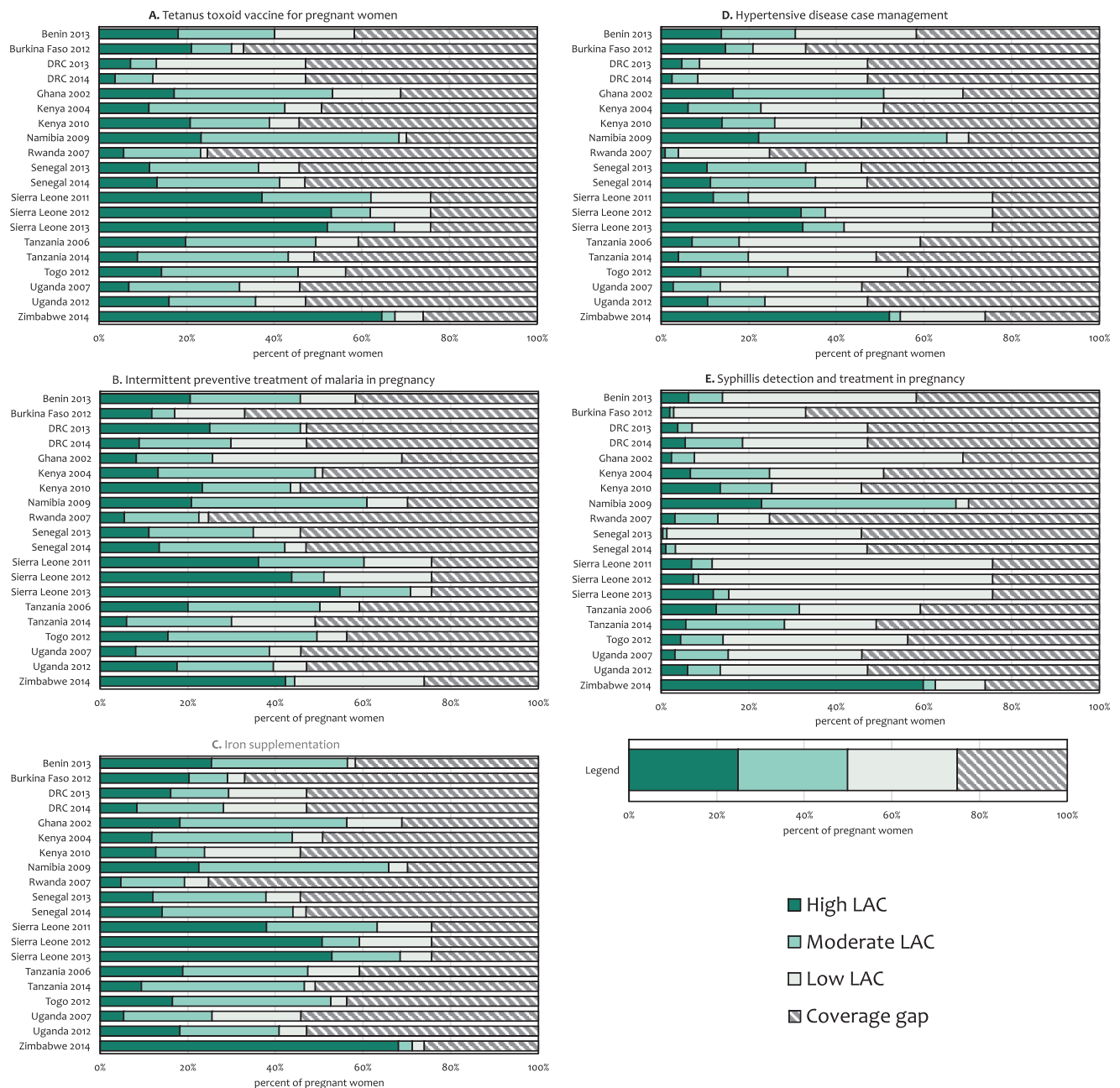


Figure 3. Likelihood of appropriate care (LAC) for antenatal care interventions by country and year.

**Table 4.** Percentage of all pregnant women with a high likelihood of appropriate care for antenatal care (ANC) interventions\*

COUNTRY	YEAR	SURVEY	TT	IPTp	SYPH T&T	HYP	Fe
Benin	2013	SARA	18.1	20.6	6.4	9.2	25.5
Burkina Faso	2012	SARA	21.2	11.9	2.0	3.6	20.3
DRC	2013	SARA	7.2	25.1	3.9	0.9	16.1
DRC	2014	SARA	3.7	8.9	5.5	1.0	8.4
Ghana	2002	SPA	17.3	8.4	2.5	14.5	18.3
Kenya	2004	SPA	11.5	13.3	6.7	0.6	11.9
Kenya	2010	SPA	20.9	23.3	13.5	6.8	12.8
Namibia	2009	SPA	23.4	20.8	23.0	12.2	22.5
Rwanda	2007	SPA	5.8	5.6	3.2	0.1	4.8
Senegal	2013	SPA	11.7	11.2	1.1	3.4	12.1
Senegal	2014	SPA	13.3	13.6	3.4	5.0	14.2
Sierra Leone	2011	SARA	37.3	36.3	7.0	9.0	38.0
Sierra Leone	2012	SARA	53.0	43.8	7.3	29.9	50.7
Sierra Leone	2013	SARA	52.1	54.8	12.0	28.7	52.9
Tanzania	2006	SPA	19.8	20.1	12.7	0.9	18.9
Tanzania	2014/15	SPA	8.8	6.1	5.7	3.2	9.5
Togo	2012	SARA	14.3	15.6	4.5	1.8	16.5
Uganda	2007	SPA	6.8	8.2	3.2	0.9	5.4
Uganda	2012	SARA	16.0	17.6	6.1	4.9	18.2
Zimbabwe	2014	SARA	64.6	42.4	59.9	50.8	68.1
Mean			21.3	20.4	8.4	9.4	22.3

TT – tetanus toxoid vaccine for pregnant women, IPTp – intermittent preventive treatment of malaria in pregnancy, Syph T&T – syphilis detection and treatment in pregnancy, Hyp – hypertensive disease case management, Fe – iron supplementation

\*Source: Service Provision Assessment (SPA) and Service Availability and Readiness Assessment (SARA) surveys.

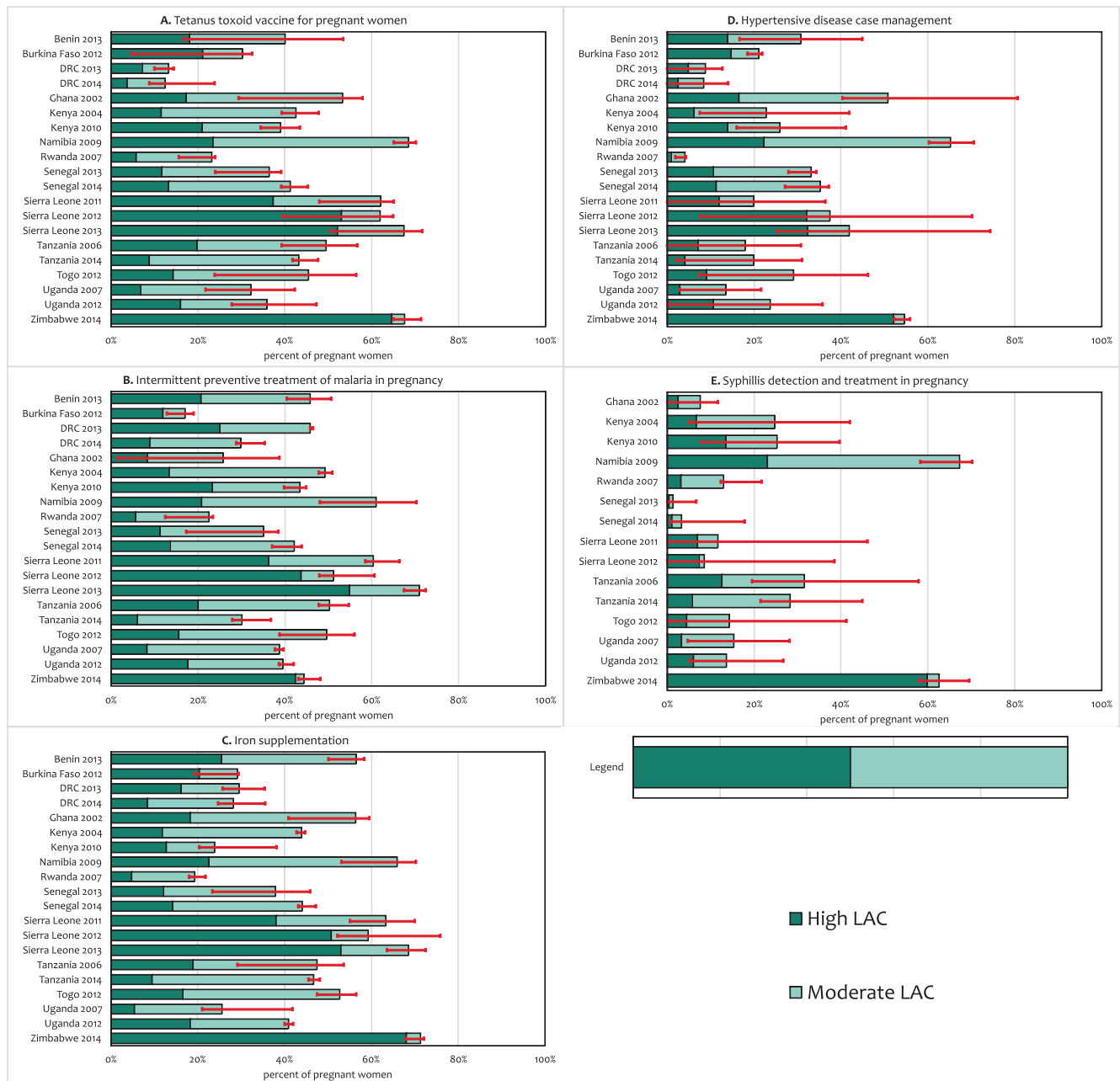
ments in the likelihood of appropriate care for syphilis testing and treatment (Table 4). Similar improvements over time were observed for hypertensive disease case management.

Sensitivity analyses showed that the overall results were stable and not unduly influenced by potential misclassification of strata (Figure 4). However, the sensitivity bounds were wider for syphilis detection and hypertensive disease case management, suggesting a greater heterogeneity in readiness to deliver these interventions by stratum.

For three interventions, we were able to directly compare estimates of the likelihood of appropriate care based on the linking approach and coverage estimates derived from mother's recall from the DHS (Figure 5). If women sought care at facilities ready to deliver the intervention, and then received it, one would expect "perfect" correlation between these two estimates, assuming women were able to accurately report receipt of the intervention in the household survey. Our estimates of high likelihood of appropriate care correlated relatively well with the DHS coverage estimates for all three interventions (iron supplementation: Pearson correlation 0.52, p value 0.02; tetanus toxoid vaccination: Pearson correlation 0.46, p value 0.04; IPTp: Pearson correlation 0.64, p value 0.003). Our estimates of high likelihood of appropriate care tended to underestimate coverage levels obtained from the DHS for iron supplementation and tetanus toxoid vaccination while our estimates overestimated coverage levels obtained from the DHS for IPTp (Figure 5).

## DISCUSSION

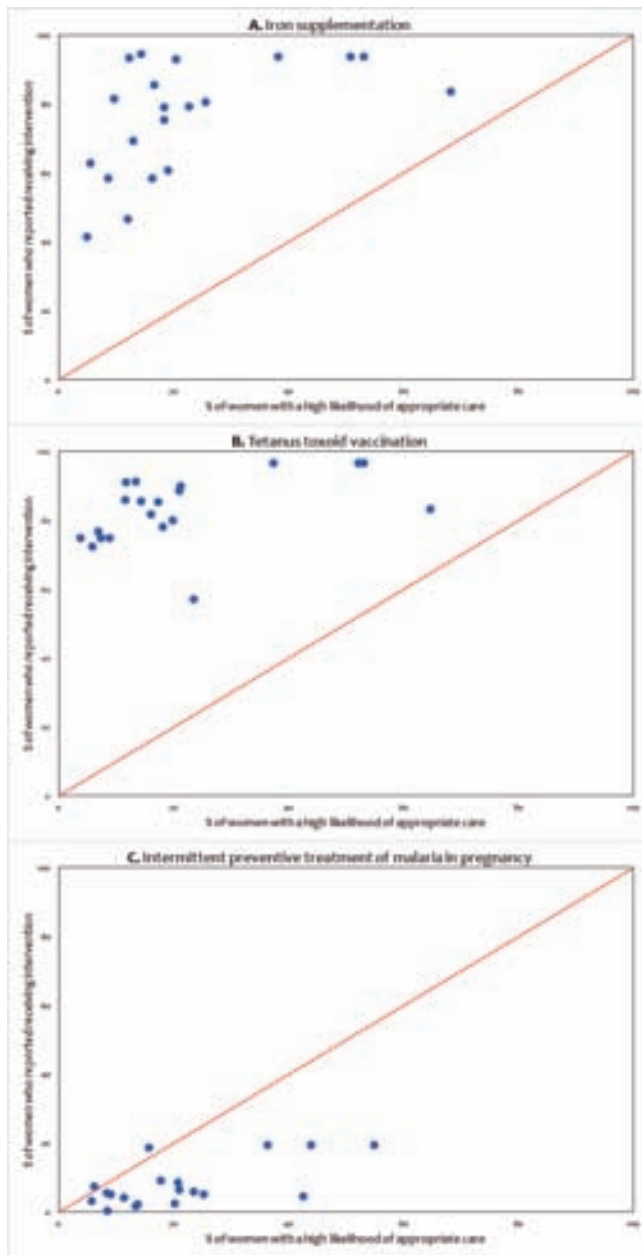
This study systematically linked 20 large-scale health facility assessments (SPAs and SARAs) and household surveys (DHS) conducted in sub-Saharan Africa to estimate the percentage of women who attended health facilities which were 'ready' to deliver five key ANC interventions. While the immediate goal of our framework was to determine the likelihood of appropriate care for key ANC interventions in several countries in sub-Saharan Africa, our analysis also offers a strategy to identify opportunities for strengthening ANC at the national level. Furthermore, this study underscores the need for a two-pronged approach to improving coverage of ANC interventions: namely, to improve ANC utilization and quality of ANC services.



**Figure 4.** Results of sensitivity analysis to assess the impact of misclassification of stratum on the likelihood of appropriate care (high and moderate).

Household surveys are useful for monitoring trends in ANC coverage based on self-reported utilization patterns. However, ANC coverage does not provide insight about the actual delivery of health services nor does it serve as a reliable measure of coverage of interventions delivered during ANC. Health facility assessments, on the other hand, can provide insight on service availability and readiness at health facilities, but not on population-based coverage. Linking information about service provision or readiness to service utilization data offers great potential in identifying barriers to achieving high population coverage of interventions [27]. For tetanus toxoid vaccine, IPTp, and iron supplementation, our comparison of DHS coverage estimates and linked estimates in this analysis correlated relatively well. While there were biases in the two approaches to measuring coverage of these three interventions, the fairly good concordance suggests that the linking approach may be useful in estimating coverage for those ANC interventions for which DHS provides no coverage estimate.

The percentage of women attending health facilities ready to deliver tetanus toxoid vaccination, IPTp, and iron supplementation was relatively higher than the more complex and multi-step interventions –



**Figure 5.** Comparison of estimated coverage (DHS) and estimated high likelihood of appropriate care (LAC) based on the ecological linking for iron supplementation, tetanus toxoid vaccination and intermittent preventive treatment of malaria in pregnancy.

hypertensive disease case management and syphilis detection and treatment. Hypertension is a major cause of maternal mortality and syphilis is estimated to be associated with 11% of stillbirths [31,32]. Our findings highlight the need to improve commodity-driven logistical barriers for the delivery of ANC services, notably the supply of syphilis tests for screening injectable penicillin for syphilis treatment, dipsticks for urine protein and magnesium sulphate for pre-eclampsia. Health facility readiness varies by region, health facility type and managing authority, and country specific information is available in the final survey reports [29,30].

There are several limitations when linking health service provision to utilization of health services at an ecological level, as in this analysis. The underlying assumption of our ecological analysis was that women seeking ANC experienced the “average” level of health facility readiness in that stratum, which was not the case. Second, although ideally we would have weighted our readiness estimates by health facility utilization to better reflect the “average” ANC experience, health facility utilization data are not collected in the SPAs or SARAs. Third, the linkage approach relied on women’s self-reported source of ANC care, including the type of health facility and managing authority. To minimize recall bias, our analysis was restricted to women reporting the most recent live birth in the 3 years preceding the survey, but this constraint does not ensure the validity of all responses. Finally, health facility assessments and household surveys may have been conducted at different time points, so that the survey reference period for both sources may not be perfectly matched or temporally aligned. To address this concern, surveys were restricted to occur within two years of each other, with the underlying assumption that ANC utilization patterns were unlikely to change drastically during the two year period. However, given a DHS survey reference period of up to 3 years, it is possible to have a total lag period of up to 5 years between the time a woman had her most recent live birth and the time the health facility assessment was conducted.

Our analysis focused on health facility readiness to provide interventions, which is a minimum requirement for but not a guarantee of the delivery of quality ANC services. While facilities that lack the necessary drugs, equipment, diagnostics or trained staff cannot provide an intervention,

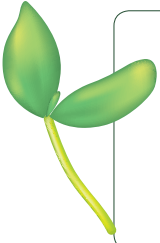
‘ready’ facilities may also fail to provide quality services in part due to lack of provider knowledge, motivation, supervision or increased workloads [27]. Although SPAs provide further information on other dimensions of quality based on observations of ANC consultations and client exit interviews, the SARAs only provide information on service availability and readiness. To allow the use both surveys we focused on health facility readiness. Future research should assess other measures of quality.

Our sample included surveys that were conducted as early as 2002, in order to aggregate a data set for proof of concept, but we acknowledge that changes in the health systems may have occurred in the past decade and recommend caution in drawing inferences based on the cross-country comparisons. Furthermore, health facility readiness can change rapidly over a short time period. Our results may not adequately reflect these shifts or the current situation across all countries in sub-Saharan Africa. As nationally representative surveys are conducted more routinely across more countries, tracking of improvements in the

quality of care over time will be possible. Lastly, our analysis did not account for the timing of ANC interventions. Early screening and treatment is a critical consideration for interventions such as syphilis detection and treatment. Our analysis assumed that interventions were appropriately delivered in a timely manner to yield the intended health gains.

Linking ANC readiness and care-seeking data at the population-level was valuable in characterizing the quality of ANC care available to pregnant women and highlighting deficiencies in the provision of ANC services across health facilities in sub-Saharan Africa. The use of standardized survey instruments facilitated the pooling of data from multiple countries and years. We believe that the linking approach that ties utilization to readiness holds great promise to produce estimates of coverage of effective interventions that cannot be or at least are not measured in household surveys.

While there will continue to be work to refine and improve estimates of coverage and service quality with analyses linking household and health facility surveys, two changes could facilitate the easier linking of data from these two sources. First, the inclusion of a question about the source of ANC in the MICS would make allow the inclusion of another standardized source of data on care-seeking patterns. Second, harmonization of the two health facility assessments (SPA and SARA) would further aid linking studies and foster new strategies in order to ultimately improve measurement of coverage in MNCH.



**Ethics approval:** Not applicable.

**Funding:** None declared.

**Authorship declaration:** Conceived of the study design: MK, NW and MKM. Analysed the data: MK. Wrote the manuscript: MK, MKM and NW.

**Competing interests:** The authors have completed the Unified Competing Interest form at [www.icmje.org/coin\\_disclosure.pdf](http://www.icmje.org/coin_disclosure.pdf) (available on request from the corresponding author) and declare no conflicts of interests.

## REFERENCES

- 1 Darmstadt GL, Bhutta ZA, Cousens S, Adam T, Walker N, de Bernis L, et al. 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
- 2 McDonagh M. Is antenatal care effective in reducing maternal morbidity and mortality? *Health Policy Plan*. 1996;11:1-15. Medline:10155875 doi:10.1093/heapol/11.1.1
- 3 Indicators for monitoring the Millennium Development Goals: definitions, rationale, concepts and sources. United Nations Publications; Available:<http://mdgs.un.org/unsd/mi/wiki/MainPage.ashx>. Accessed 10 September 2016.
- 4 Saad-Haddad G, Dejong J, Clara M, Méndez R, Vaz L, Amouzou A, et al. Patterns and determinants of antenatal care utilization: analysis of national survey data in seven countdown countries. *J Glob Health*. 2016;6:010404. Medline:27231540 doi:10.7189/jogh.06.010404
- 5 Victora CG, Requejo JH, Barros AJD, Berman P, Bhutta Z, Boerma T, et al. Countdown to 2015: a decade of tracking progress for maternal, newborn, and child survival. *Lancet*. 2016;387:2049-59. Medline:26477328 doi:10.1016/S0140-6736(15)00519-X
- 6 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
- 7 Moxon SG, Ruysen H, Kerber KJ, Amouzou A, Fournier 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 Demographic and Health Surveys. Available: <http://www.dhsprogram.com>. Accessed: 21 September 2016.
- 9 Statistics and Monitoring: Multiple Indicator Cluster Survey website. Available: <http://mics.unicef.org/>. Accessed: 21 September 2016.
- 10 Marchant T, Tilley-Gyado RD, 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
- 11 Kinney MV, Kerber KJ, Black RE, Cohen B, Nkrumah F, Coovadia H, et al. Sub-Saharan Africa's mothers, newborns, and children: Where and why do they die? *PLoS Med*. 2010;7: e1000294. Medline:20574524 doi:10.1371/journal.pmed.1000294
- 12 Bryce J, Arnold F, Blanc A, Hancioglu A, Newby H, Requejo J, et al. 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

- 13 Munos MK, Stanton C, Bryce JCG; The Core Group for Improving Coverage Measurement for MNHC. Improving coverage measurement for reproductive, maternal, neonatal and child health: gaps and opportunities. *J Glob Health*. 2017;7:010801. Medline:28607675 doi:10.7189/jogh.07.010801
- 14 Do M, Micah A, Brondi L, Campbell H, Marchant T, Eisele T, et al. Linking household and facility data for better coverage measures in reproductive, maternal, newborn, and child health care: systematic review. *J Glob Health*. 2016;6:020501. Medline:27606060 doi:10.7189/jogh.06.020501
- 15 World Health Organization. Service availability and readiness assessment (SARA): an annual monitoring system for service delivery: reference manual. 2013. Available: [http://www.who.int/healthinfo/systems/SARA\\_Reference\\_Manual\\_Full.pdf](http://www.who.int/healthinfo/systems/SARA_Reference_Manual_Full.pdf). Accessed: 10 September 2017
- 16 O'Neill K, Takane M, Sheffel A, Abou-zahr C, Boerma T. Monitoring service delivery for universal health coverage: the Service Availability and Readiness Assessment. *Bull World Health Organ*. 2013;91:923-31. Medline:24347731 doi:10.2471/BLT.12.116798
- 17 Lee E, Madhavan S, Bauhoff S. Levels and variations in the quality of facility-based antenatal care in Kenya: evidence from the 2010 service provision assessment. *Health Policy Plan*. 2016;31:777. Medline:26879091 doi:10.1093/heapol/czv132
- 18 Kyei NNA, Chansa C, Gabrysch S. Quality of antenatal care in Zambia: a national assessment. *BMC Pregnancy Childbirth*. 2012;12:151. Medline:23237601 doi:10.1186/1471-2393-12-151
- 19 Sipsma HL, Curry LA, Kakoma J-B, Linnander EL, Bradley EH. Identifying characteristics associated with performing recommended practices in maternal and newborn care among health facilities in Rwanda: a cross-sectional study. *Hum Resour Health*. 2012;10:13. Medline:22776289 doi:10.1186/1478-4491-10-13
- 20 K'Oyugi BO. Access to child health care, medical treatment of sick children and childhood mortality relationships in Kenya. *Health (Irvine Calif)*. 2014;6:1152-64.
- 21 Getachew T, Bekele A, Defar A, Tadesse M. Tuberculosis Service Provision in Ethiopia: Health Facility Assessment. pp. 145-159.
- 22 Thatte N, Choi Y. Does human resource management improve family planning service quality? Analysis from the Kenya Service Provision Assessment 2010. *Health Policy Plan*. 2015;28:356-67. Medline:24740709 doi:10.1093/heapol/czu019
- 23 Jayachandran V, Chapotera G, Stones W. Quality of facility-based family planning services for adolescents in Malawi: findings from a national census of health facilities. *Malawi Med J*. 2016;33:48-52. Medline:27895828 doi:10.4314/mmj.v28i2.4
- 24 Choi Y, Ametepi P. Comparison of medicine availability measurements at health facilities: evidence from Service Provision Assessment surveys in five sub-Saharan African countries. *BMC Health Serv Res*. 2013;13:266. Medline:23837467 doi:10.1186/1472-6963-13-266
- 25 Campbell OMR, Aquino EML, Vwalika B, Gabrysch S. Signal functions for measuring the ability of health facilities to provide abortion services: an illustrative analysis using a health facility census in Zambia. *BMC Pregnancy Childbirth*. 2016;16:105. Medline:27180000 doi:10.1186/s12884-016-0872-5
- 26 Kruk ME, Leslie HH, Verguet S, Mbaruku GM, Adanu RMK, Langer A. Quality of basic maternal care functions in health facilities of five African countries: an analysis of national health system surveys. *Lancet Glob Health*. 2016;4:e845-e855. Medline:27670090 doi:10.1016/S2214-109X(16)30180-2
- 27 Baker U, Peterson S, Marchant T, Mbaruku G, Temu S, Manzi F, et al. Identifying implementation bottlenecks for maternal and newborn health interventions in rural districts of the United Republic of Tanzania. *Bull World Health Organ*. 2015;93:380-9. Medline:26240459 doi:10.2471/BLT.14.141879
- 28 The DHS Program - Service Provision Assessments (SPA). Available: <http://dhsprogram.com/What-We-Do/Survey-Types/SPA.cfm>. Accessed: 22 September 2016.
- 29 The DHS Program - Available Datasets. Available: <http://dhsprogram.com/data/available-datasets.cfm>. Accessed: 22 September 2016.
- 30 WHO. | Service availability and readiness assessment (SARA). Available: [http://www.who.int/healthinfo/systems/sara\\_introduction/en/](http://www.who.int/healthinfo/systems/sara_introduction/en/). Accessed: 22 September 2016.
- 31 de Bernis L, Kinney MV, Stones W, ten Hoope-Bender P, Vivio D, Leisher SH, et al. Stillbirths: ending preventable deaths by 2030. *Lancet*. 2016;387:703-16. Medline:26794079 doi:10.1016/S0140-6736(15)00954-X
- 32 Say L, Chou D, Gemmill A, Tunçalp Ö, Moller A-B, Daniels J, et al. Global causes of maternal death: a WHO systematic analysis. *Lancet Glob Health*. 2014;2:e323-33. Medline:25103301 doi:10.1016/S2214-109X(14)70227-X

# Development and validation of a simplified algorithm for neonatal gestational age assessment – protocol for the Alliance for Maternal Newborn Health Improvement (AMANHI) prospective cohort study

AMANHI (Alliance for Maternal and Newborn Health Improvement); Abdullah Baqui<sup>1</sup>, Parvez Ahmed<sup>1</sup>, Sushil Kanta Dasgupta<sup>1</sup>, Nazma Begum<sup>1</sup>, Mahmoodur Rahman<sup>1</sup>, Nasreen Islam<sup>1</sup>, Mohammad Quaiyum<sup>1</sup>, Betty Kirkwood<sup>2</sup>, Karen Edmond<sup>2</sup>, Caitlin Shannon<sup>2</sup>, Samuel Newton<sup>2</sup>, Lisa Hurt<sup>2</sup>, Fyezah Jehan<sup>3</sup>, Imran Nisar<sup>3</sup>, Atiya Hussain<sup>3</sup>, Naila Nadeem<sup>3</sup>, Muhammad Ilyas<sup>3</sup>, Anita Zaidi<sup>3</sup>, Sunil Sazawal<sup>4</sup>, Saikat Deb<sup>4</sup>, Arup Dutta<sup>4</sup>, Usha Dhingra<sup>4</sup>, Said Moh'd Ali<sup>4</sup>, Davidson H. Hamer<sup>5</sup>, Katherine EA Semrau<sup>5</sup>, Marina Straszak-Suri<sup>5</sup>, Caroline Grogan<sup>5</sup>, Godfrey Biemba<sup>5</sup>, Anne CC Lee<sup>6</sup>, Blair J Wylie<sup>6</sup>, Alexander Manu<sup>7</sup>, Sachiyo Yoshida<sup>7</sup>, Rajiv Bahl<sup>7</sup>

<sup>1</sup> AMANHI Gestational Age Study Group, Bangladesh (Sylhet)

<sup>2</sup> AMANHI Gestational Age Study Group, Ghana

<sup>3</sup> AMANHI Gestational Age Study Group, Pakistan (Karachi)

<sup>4</sup> AMANHI Gestational Age Study Group, Tanzania (Pemba)

<sup>5</sup> AMANHI Gestational Age Study Group, Zambia

<sup>6</sup> Brigham & Women's Hospital, Massachusetts General Hospital, Boston, Massachusetts, USA

<sup>7</sup> World Health Organization (MCA/MRD), Geneva, Switzerland

## Correspondence to:

Dr Rajiv Bahl & Alexander Ansah Manu  
MCA/MRD, World Health Organization  
Avenue Appia 20  
CH-1211 Geneva 27  
Switzerland  
bahlr@who.int  
makmanu128@gmail.com

**Objective** The objective of the Alliance for Maternal and Newborn Health Improvement (AMANHI) gestational age study is to develop and validate a programmatically feasible and simple approach to accurately assess gestational age of babies after they are born. The study will provide accurate, population-based rates of preterm birth in different settings and quantify the risks of neonatal mortality and morbidity by gestational age and birth weight in five South Asian and sub-Saharan African sites.

**Methods** This study used on-going population-based cohort studies to recruit pregnant women early in pregnancy (<20 weeks) for a dating ultrasound scan. Implementation is harmonised across sites in Ghana, Tanzania, Zambia, Bangladesh and Pakistan with uniform protocols and standard operating procedures. Women whose pregnancies are confirmed to be between 8 to 19 completed weeks of gestation are enrolled into the study. These women are followed up to collect socio-demographic and morbidity data during the pregnancy. When they deliver, trained research assistants visit women within 72 hours to assess the baby for gestational maturity. They assess for neuromuscular and physical characteristics selected from the Ballard and Dubowitz maturation assessment scales. They also measure newborn anthropometry and assess feeding maturity of the babies. Computer machine learning techniques will be used to identify the most parsimonious group of signs that correctly predict gestational age compared to the early ultrasound date (the gold standard). This gestational age will be used to categorize babies into term, late preterm and early preterm groups. Further, the ultrasound-based gestational age will be used to calculate population-based rates of preterm birth.

**Importance of the study** The AMANHI gestational age study will make substantial contribution to improve identification of preterm babies by frontline health workers in low- and middle- income countries using simple evaluations. The study will provide accurate preterm birth estimates. This new information will be crucial to planning and delivery of interventions for improving preterm birth outcomes, particularly in South Asia and sub-Saharan Africa.

Accurately identifying preterm births around the time of delivery increases opportunities for using targeted high impact interventions to reduce mortality and morbidity but requires that gestational age at birth be known. Preterm birth is a global challenge affecting both developed and developing countries [1,2]. It contributes annually to approximately 35% of neonatal and 75% of perinatal mortality worldwide [3–5]. In addition, survivors suffer long-term respiratory, gastrointestinal and neurodevelopmental morbidities with consequences on health, growth, psychosocial functioning and economic viability in later life [6–8]. The greatest burden of mortality and morbidity from preterm births occur in low and middle income settings (LMICs) [9]. In 2005, almost 13 million children were born preterm and 85% of these occurred in Africa and Asia [10].

Interventions that are feasible to implement and have the potential to prevent or help manage preterm birth complications in LMICs exist, but their effectiveness has not been rigorously evaluated [11]. The critical step to deploying such interventions is identification of preterm births. In high-income settings, early pregnancy ultrasound is used to enhance accurate gestational age estimation. Preterm births are therefore anticipated right at the onset of labour and interventions such as antenatal corticosteroids are administered to improve outcomes. In contrast, pregnancy ultrasound for assessing gestational age is often not available or affordable to women in LMICs [12]. Pregnant women also seek care quite late; by which time most ultrasound scans do not yield reliable gestational age estimates. To help frontline health workers identify preterm babies for targeted interventions, simpler methods of gestational age assessment are needed.

One option is to train them to identify preterm babies immediately after they are born. The Ballard [13,14] and Dubowitz [15,16] scoring systems use physical and neuromuscular signs to estimate newborn gestational maturity after birth. Whilst they are potentially feasible to implement in LMICs, they are rarely used at scale partly because they have not been rigorously evaluated in comparison with a reliable “gold standard” or because they include too many component signs that are difficult to assess on all babies. For instance, five [14,17–20] of six studies that evaluated the accuracy of these scores compared them with women’s last menstrual period (LMP) that is itself not reliable for gestational age assessment. Meanwhile, portable ultrasound machines are now available, relatively cheap and can be deployed to allow improved gestational age assessment at population level or within research settings in LMICs. In addition, advances in computer programming and “machine learning” methods enable simpler algorithms with equivalent validity to be derived from complex patterns and inter-relationships between signs.

The Alliance for Maternal and Newborn Health Improvement (AMANHI) gestational age study is being implemented in five countries in South Asia and sub-Saharan Africa. It aims to develop and validate simple, programmatically feasible methods (using non-clinically-trained research workers) to accurately assess gestational age of babies at the population level in comparison to early pregnancy ultrasound (gold standard). The study will also use these accurate gestational ages to evaluate the risk of mortality and morbidity, by gestational age, among neonates. This manuscript describes the protocol for the harmonized implementation of the study.

## METHODOLOGY

### Study design

The AMANHI gestational age study is a multi-centre, population-based prospective development and evaluation of the diagnostic accuracy of simple methods for gestational age assessment (including reported LMP, physical, neuromuscular, feeding assessments and anthropometry). These will be compared with gestational age from early pregnancy ultrasound scans. The Maternal, Newborn, Child and Adolescent Health department of the World Health Organization (WHO/MCA) is coordinating the study.

### Objectives

The objective of the AMANHI gestational age study is to develop and validate a programmatically feasible and simple approach to accurately assess the gestational age of babies after they are born. The study will provide accurate, population-based rates of preterm birth in different settings and quantify the risks of (i) neonatal mortality by gestational age and birthweight, and (ii) maternal and neonatal morbidity by gestational age from population-based pregnancy cohorts.

## Study settings

Five sites in South Asia (Sylhet, Bangladesh and Karachi, Pakistan) and sub-Saharan Africa (Kintampo, Ghana; Pemba, Tanzania; and Southern Province, Zambia) are using harmonized protocols to implement the AMANHI gestational age study. Across the sites, study populations are predominantly rural with low literacy levels. Health facilities of various types provide a range of services from basic to comprehensive emergency obstetric and newborn care. All sites, except Zambia, have an on-going one to three-monthly community-based pregnancy surveillance in which trained fieldworkers conduct home visits to all women of reproductive age to identify pregnancies for enrolment and follow-up until the end of pregnancy. In Zambia, the study is recruiting pregnant women at antenatal clinics.

## Identifying eligible pregnancies and obtaining consent

**Eligibility.** Trained fieldworkers use a variety of methods to identify pregnant women for the study. Women either directly report their pregnancies or the fieldworkers elicit missed periods from women's reported dates of their LMP. They also conduct pregnancy tests, if required, to confirm pregnancies. In Zambia, over 96% of all pregnant women attend antenatal care (ANC) clinics so study participants are recruited from these clinics. The study enrolls women only if their pregnancies are between 8–19 completed weeks by ultrasound examination. This window ensures that pregnancies are early enough to give accurate ultrasound gestational age estimates but also excludes very early pregnancy losses due to genetic or chromosomal abnormalities. Enrolled women should also be willing to be available for the gestational age assessment to be conducted on their babies within 72 hours after birth.

**Consenting.** Fieldworkers consent pregnant women, in their local or preferred languages, to undergo a screening ultrasound examination to establish gestational age and subsequently enrol women if eligibility is confirmed (pregnancies between 8–19 weeks' gestation). Consent for additional procedures varies slightly across sites. In Ghana, for instance, fieldworkers obtain initial consent for the screening ultrasound. They only consent eligible women after scan and again before the gestational age assessment when the baby is born. In other sites, women are comprehensively consented at the initial contact to cover the scan and follow-up. In Zambia in case the pregnant woman is a "minor" between 15 years to 18 years, assent is first obtained from women themselves before their legal guardians provide written informed consent for their participation.

## Harmonization of implementation

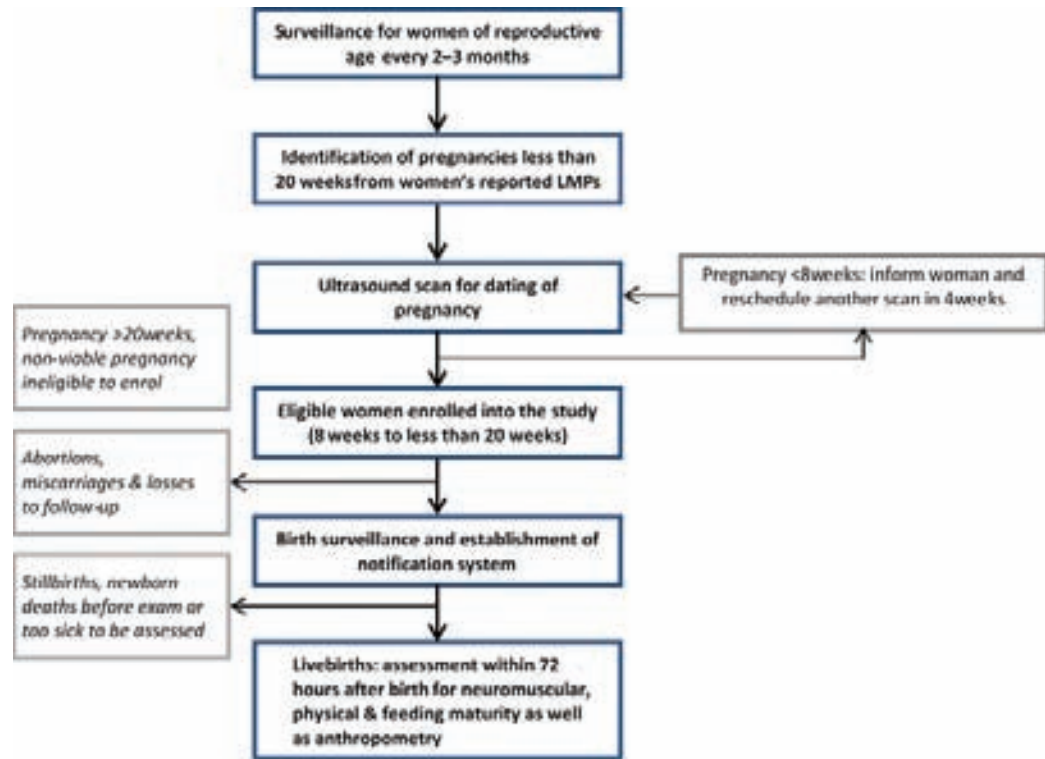
Uniform protocols are being used across sites to conduct the screening ultrasound scan, pregnancy follow-up, newborn gestational age assessments and quality assurance. The principal investigators from the sites, WHO/MCA team and selected global newborn experts supported the development of these protocols for use at all sites. Centralised training sessions were organised for investigators from the sites to ensure harmonised implementation.

### *Standard operating procedures*

Principal investigators from the sites developed generic standard operating procedures (SOP) for implementing the study. This document outlines the strategy for implementation as illustrated in **Figure 1**. In summary, it specifies that the gold standard gestational age will be derived from an early pregnancy ultrasound scan (<20 weeks) conducted as part of the study by trained study staff. The study enrolls pregnant women during the gestational age window of 8 to 19 completed weeks. Biometric parameters measured include fetal crown-rump length (CRL), bi-parietal diameter (BPD) and femur length (FL) as appropriate. Trained non-clinical workers conduct the gestational age assessment of newborns. With this cadre of workers, the AMANHI gestational age study ensures that minimally trained frontline health workers can implement postnatal gestational age assessment, at scale. A common manual for implementation and core variable table have been developed, which contains modules on the standards for the dating ultrasound, newborn gestational age assessment and data collected during follow-up home visits.

### *Centralised standardised training and validation of trainers for newborn assessment*

AMANHI organised a 3-day training of trainers in Sylhet (Bangladesh) for two physician participants from each site, using the AMANHI gestational age study manual. The training, coordinated by WHO/MCA, aimed to standardize these participants for harmonised implementation across sites. Two experts, a WHO Medical Officer in newborn health (AM) and a paediatrician specialist in newborn maturity assessment



**Figure 1.** Flow diagram showing overview of the AMANHI gestational age study.

(ACL) facilitated the training. It involved theoretical instructions, videos and practice sessions. After showing videos of the assessment, the facilitators demonstrated the assessment modalities to the participants on mannequins, with emphasis on the assessment principles. Participants in turn practiced on these mannequins to acquaint themselves with the assessment modalities. The last two days of training involved practice at the paediatric ward of the Osmani Medical College in Sylhet. The trainers again demonstrated assessments to participants using newborn babies on the wards. They supervised the participants to take turns, practicing initially with term and later with preterm babies. Participants were closely supervised and provided additional support to ensure that they became confident and proficient in the assessments. For the proficiency validation and standardization, each participant independently assessed five or more babies and recorded their findings. Trained facilitators also independently assessed these same babies and the findings were compared with the participants. Participants were only certified as having attained mastery for any particular sign if their assessment findings did not vary by more than one point from the findings from the independent assessment by the facilitators.

#### ***Ultrasonography for accurately dating pregnancies (gold standard data)***

The SOP for the AMANHI gestational age study specifies standardized procedures for measuring fetal biometric parameters, and was developed in consultation with a maternal–fetal medicine specialist (BJW). Scanning is strictly trans–abdominal. When a potential participant comes to a scanning centre to have her pregnancy date assessed, the choice of which biometric parameter to measure depends on the estimated gestational age; this was to ensure that gestational ages estimated from these parameters have the highest precision (Table 1). AMANHI measures only the crown–rump length (CRL) if pregnancy is less than 14 weeks, both bi–parietal diameter (BPD) and femur length (FL) if more than 14 weeks and all three if within the 14th week. Three measurements are independently taken and recorded for each parameter. Pregnancies that are less than 8 weeks are re–scheduled for repeat scans after four weeks. Women with pregnancies at 20 weeks or more are counselled to continue routine antenatal clinic attendance but are not enrolled into the AMANHI gestational age study. Major abnormalities or intra–uterine fetal deaths are promptly referred to health facilities for appropriate management.

**Quality control.** The team developed a common checklist, in consultation with the MFM expert and with adaptations from other international ultrasound quality guidelines [21], which specifies the mini-

**Table 1.** Choice of fetal biometric parameter by estimated gestational age in AMANHI

GESTATIONAL AGE OF FETUS	WHAT BIOMETRIC PARAMETER TO MEASURE
8 weeks 0 days – 13 weeks 6 days	Crown–rump length only
14 weeks 0 days – 14 weeks 6 days	Crown–rump length, bi–parietal diameter & femur length
15 weeks 0 days and beyond	Bi–parietal diameter and femur length

mum acceptable quality standards for measuring each biometric parameter (**Table 2**). This checklist serves as a reference (and harmonization) document for the sonographers at the sites as well as a yardstick for independent assessment of image quality and includes guidance on appropriate image zooming, obtaining the correct plane and placement of callipers.

Key quality control measures include:

- i. The MFM expert independently reviews all images taken for 20 among the first 50 study participants recruited in each site and provides direct feedback on each individual image and the overall performance of sonographers. She also makes recommendations on how to further improve image quality.
- ii. Checking images throughout the duration of enrolment;
  - Images taken for 10% of participants are randomly selected and sent to a specialist sonographer (local to each site but independent of the AMANHI study team) for review. This independent sonographer provides direct feedback to inform refresher–training needs of study sonographers within each site.
  - Images taken for 5% of participants are randomly selected and sent for central review and validation by the MFM expert (BJW).

### **Morbidity and birth outcome surveillance**

Trained surveillance fieldworkers visit women in their homes, immediately after enrolment, to collect baseline socio–demographic, medical and obstetric histories. They make three antenatal visits at 24–28 weeks, 32–36 weeks and after 38 weeks to collect data on morbidity and mortality during the pregnancy and pregnancy outcomes using the common core variable table. Women's blood pressure is measured and their urine is tested for proteins. In some sites, women's symphysis–fundal height is also measured. During these visits, fieldworkers establish a birth notification system so that families will notify the study whenever the pregnancy ends – irrespective of the outcome. For live births, immediately after the notification, research assistants are dispatched to reach babies and conduct AMANHI gestational age assessment within 72 hours of birth. The surveillance fieldworkers also increase their frequency of home visits to all women whose pregnancies are 28 weeks or older and leave their phone numbers for families to contact them whenever the pregnancy ends. When any pregnancy ends, additional home visits are made to ascertain the status of newborns and mothers at the end of the first 28 and 42 to 60 days, days respectively. For every maternal, fetal or neonatal death, a verbal autopsy form is completed to obtain detailed information on the circumstances leading to the death; causes of death are assigned by trained and AMAN-

**Table 2.** Indicators for ensuring optimal quality of ultrasound images for each biometric parameter

QUALITY MODALITY	BIOMETRIC PARAMETER		
	Crown rump length	Bi–parietal diameter	Femur length
<b>Zooming</b>	Good magnification (50% of image size)	Good magnification (50% of image size)	Good magnification (50% of image size)
<b>Frozen in correct plane</b>	Neutral position–not hyperflexed or hyperextended	Skull is oval and bone visible throughout	Femur imaged side–to–side on screen
	Fetus horizontal (side–to–side on screen)	Thalamus is visible	Only one bone in this portion of extremity
	Full extent of crown visible	Skull side to side on screen	Upper femur measured
	Full extent of rump visible		Full extent of femur visualized (solid straight line)
<b>Caliper placement</b>	Crown caliper at edge of head	Calipers placed perpendicular to long axis of skull	Calipers placed at edge of echogenic bone (outer to outer)
	Rump caliper placement at lower spine	Top caliper placed on outer portion of the skull	Secondary ossification centres not measured
		Bottom caliper placed on inner portion of skull	

HI-accredited physicians. The accreditation process involves training for harmonized assignment of causes of deaths and passing of a centralised standardised test conducted by the WHO [22].

### Newborn assessment

**Assessors.** Non-clinician field workers (at supervisory grade and with <sup>3</sup>10 years formal education) were trained and standardized to conduct the AMANHI newborn gestational age assessments. These assessors are blinded to babies' ultrasound-estimated gestational ages.

**Inclusion and exclusion criteria.** The assessors screen for newborn illnesses before conducting the newborn assessment. Families are engaged throughout the assessment process. The assessors ask to know about the families' perception on the general health of the newborn. If babies are reportedly unwell, they use the newborn assessment module of WHO's integrated management of childhood illness (IMCI) guidelines to assess severity. Babies with severe illness are excluded and referred to health facilities for care. Even if the family thinks the baby is well but the assessors consider them unwell, they also refer to health facilities for care. In general, assessors advise care-seeking at health facilities even when the IMCI algorithm shows the newborn is not ill but the family thinks otherwise.

**Assessment modalities.** Table 3 shows signs being assessed for in the AMANHI gestational age study. These signs were selected based on a systematic literature review for evidence of their use in assessing newborn gestational maturity and a pragmatic consideration of how easily they could be assessed by front-line health workers. There are six neuromuscular signs which test passive flexor tone or flexibility and 10 physical signs. The physical and neuromuscular signs were selected from the original Ballard score [14]. Each sign relates to a particular body part (eg, skin), under which are sub-categories such as colour, opacity, texture, etc. The separate sub-categories were adapted from the Dubowitz score [15,16] because they are separated and relatively easier to differentiate. After discussions with Dubowitz (personal communication), ankle dorsiflexion was added to the neuromuscular signs as a measure of joint flexibility. The "square window" sign was dropped due to parental concerns during piloting.

Seven anthropometric measures (head, chest and mid-upper arm circumference, infant length, birth weight, breast-bud diameter, infant foot length and (optionally) women's symphysis-fundal height) in

**Table 3.** Assessment modalities with acceptable precision (for anthropometric measurements)

ASSESSMENT	SIGNS	PRECISION (IF APPLICABLE)
Neuro-muscular signs	<ul style="list-style-type: none"> <li>• Posture</li> <li>• Arm recoil</li> <li>• Scarf sign</li> <li>• Popliteal angle</li> <li>• Heel-to-ear test</li> <li>• Ankle dorsiflexion</li> </ul>	
Physical signs	<ul style="list-style-type: none"> <li>• <b>Skin:</b> colour, texture, opacity and presence of lanugo</li> <li>• <b>Ear:</b> shape and recoil</li> <li>• <b>Breast:</b> nipple-areola development</li> <li>• <b>Male genitalia:</b> testes and scrotum</li> <li>• <b>Female genitalia:</b> labia and clitoris</li> <li>• <b>Foot:</b> plantar creases</li> </ul>	
Anthropometry	<ul style="list-style-type: none"> <li>• Head circumference (cm)</li> <li>• Chest circumference (cm)</li> <li>• Breast bud diameter (mm)</li> <li>• Mid-upper arm circumference (cm)</li> <li>• Foot length (mm)</li> <li>• Infant length (cm)</li> <li>• Symphysis-fundal height (cm)</li> <li>• Weight (g)</li> </ul>	<ul style="list-style-type: none"> <li>±0.5 cm</li> <li>±0.5 cm</li> <li>±1 mm OR ±0.1 cm</li> <li>±2 mm OR ±0.2 cm</li> <li>±2 mm OR ±0.2 cm</li> <li>±0.5 cm</li> <li>±1.0 cm</li> <li>±40 g if digital scale &amp; 100 g if suspension scale</li> </ul>
Feeding	<ul style="list-style-type: none"> <li>• Check whether the infant is able to attach well to the breast by observing four signs: mouth open; lower lip everted; chin touching the breast below; greater part of the areola in mouth and more areola seen above than below.</li> <li>• Observation: Suckling behaviour of the baby at the breast – deep, slow continuous suck with swallowing in between or shallow quick sucks.</li> <li>• Observe: Duration the infant was able to stay attached to the breast continuously during the feed?</li> <li>• Observation: The longest continuous burst of suckling (number of sucks)</li> <li>• While observing breastfeeding, how many times did the baby suck before swallowing?</li> </ul>	

pregnancy are taken. Foot length (right foot) is measured from base of the heel to the tip of the halux, based on methods proposed by Marchant et al [23]. For all the measurements, two are taken independently and if their difference exceeds a set acceptable margin of error (Table 3), a third measurement is taken. All three measurements are recorded. Five signs of feeding maturity, adapted from WHO's infant feeding assessment guide with additions (on suckling bursts and suck-to-swallow ratios) from the Nvquist [24] preterm feeding questionnaire are also assessed. Assessors record their findings on a standard AMANHI form.

### ***Supervision of assessment***

This involves direct observation, scheduled and un-scheduled visits to study assessors. Study coordinators supervise assessors on the field. They directly observe 5% of assessments conducted by the Research Assistants in the communities as part of the supervision. These 5% selected observations are spread over the entire duration of the study to ensure assessors' skills are being maintained. The coordinators have been purposively selected from a pool of supervisors who have performed excellently in previous engagements as supervisors in the sites. Their qualifications range from post-secondary to university graduates. As well as their extensive previous experience in field research, they were provided additional training in the conduct of the gestational age assessment and then on how to supervise and provide feedback to the assessors. The coordinators also make scheduled and unscheduled visits to assessors in the field to observe their work. In Tanzania and Ghana, geographical information systems are used to locate assessors for these supervisory visits. After each supervisory visit session, feedback on performance is provided to the assessor.

In some sites including Ghana, a pilot implementation phase followed the assessors' training during which the study clinician/trainer conducted intensive supervision of assessors and directly observed five or more of their newborn assessments. That phase was also used to train field coordinators on the practicality of supervising the assessors.

### ***Professional clinician validation of assessment***

In order to validate the assessment by the non-clinicians, a trained professional clinician/trainer independently validates a random 5–15% sub-sample of assessments conducted by the assessors. In some sites, the process is centrally coordinated and an automated system or a study coordinator selects these babies, informs the physician of the deliveries and provides them with women's location to facilitate the validation assessment. The physician visits the woman at the place of delivery, also within 72 hours, to independently assess babies using the AMANHI assessment forms.

### ***Study materials/equipment***

The AMANHI gestational age study teams procured uniform equipment and materials for the study across sites. Where materials could not be procured from a common source, the desired quality was specified in order to provide comparable error margins. Weighing scales, non-stretchable measuring tapes, transparent rulers and infantometers were procured and are used for the assessment across sites.

### ***Supportive activities***

In all the sites, the AMANHI study built on existing long-standing relationship of trust and mutual respect between investigators, communities and health facilities. They held meetings with community members (and their leaders) to introduce the study and to enhance acceptability of assessments. Separate orientation sessions were organised in health facilities to solicit their support and discuss the potential impact of the AMANHI gestational age study on facility workload. These included discussions on referral of complicated pregnancies for appropriate management. In some sites, a separate training was conducted for facility staff to assess selected physical signs on babies considered too sick to be assessed by the non-clinician assessors.

### ***Sample size estimation***

Previously pooled data from the participating sites showed that approximately 10% of all births were preterm. In the AMANHI gestational age study, it was assumed that the simpler methods will detect this prevalence with  $\pm 5\%$  (absolute) precision and achieve 80% sensitivity and specificity in comparison to the early pregnancy ultrasound gestational age estimate. With an additional 1.5-fold adjustment for de-

**Table 4.** Study sites and contribution to the overall study sample

STUDY SITES	MINIMUM SAMPLE SIZE CONTRIBUTION FROM STUDY SITES
Bangladesh (Sylhet)	2200
Ghana (Kintampo)	1000
Pakistan (Karachi)	2200
Tanzania (Pemba)	2200
Zambia (Southern Province, Zambia)	1000

sign effect (to account for contextual differences between study sites) and 20% attrition rate (due to abortions/miscarriages, stillbirths, early neonatal deaths and losses to follow-up) [25], a minimum of 5740 pregnant women will be required to be enrolled. Machine learning approaches use separate pools of data from the same population for the identification of the algorithm (two thirds of total) and for the validation of this algorithm (one third of total). This increases the total targeted sample size to 8610, 5740 for machine learning and 2870 for validation. **Table 4** shows proposed sample size contributions from each site.

## Data processing and management

All data are double-entered by two independent clerks at all sites except in Zambia, where they are scanned directly using TeleForms software (Hewlett Packard, Sunnyvale, CA). Principal investigators in the sites have excellent track records for management of large volumes of data with stringent protocols. They are locally responsible for maintaining the data quality in their sites, running range and consistency checks and conducting periodic reviews of distributions of the responses to identify outliers in the data to address promptly. The WHO/MCA also developed a data quality checking programme to be applied on all data to supplement the routine checks conducted by individual sites. Any inconsistencies were flagged and resolved in consultation with data managers and investigators from the sites.

## Plan for data analysis

The pooled data from the study sites will be split into equal thirds. Computer machine learning techniques will be employed on two-thirds of the data to find a group of signs that accurately and independently predicts gestational age or identifies preterm babies <34 weeks and/or <37 weeks compared to ultrasound dates (gold standard). The difference in gestational age estimates (in completed weeks) between the gold standard and the simpler methods (LMP and/or newborn assessments) will be plotted against the mean gestational ages (of gold standard and simplified method). Any identified algorithm that meets the requisite validity criteria will then be applied to the remaining one-third of data to assess validity (sensitivity, specificity and receiver operating characteristic curves).

Predicted gestational ages from the new algorithm will be used to categorize babies into those <34 weeks, 34–36 weeks and those ≥37 weeks or a dichotomy of preterm (<37 weeks) vs term. Chance-corrected agreement and kappa statistics will be generated between these categories and those obtained from gestational age cut-offs based on ultrasound dating and interpreted using Landis and Koch [26] criteria. Regression models will also be fitted to determine association between gestational age and adverse maternal and neonatal outcomes. Analyses will be done using Stata® Version 14.0 (StataCorp., Tx, USA) and other specialized software.

## Ethical considerations

All study protocols have been approved by ethical review committees of the WHO and appropriate institutions in each of the participating sites.

## Dissemination of study findings

Study findings will be shared promptly with the Technical Steering Committee comprising the WHO/MCA coordinating team, the principal investigators from the five sites and the Bill & Melinda Gates Foundation. Local dissemination meetings will be held with community members, policy-makers, health managers and administrators in all sites. A detailed implementation evaluation report, including lessons learned will be shared with the WHO and Gates Foundation. Policy briefs on outcomes will be disseminated nationally and internationally to relevant policy and donor organisations. Study findings will also be disseminated in scientific/technical forums and in peer-reviewed journals.

## CONTRIBUTION OF THE AMANHI GESTATIONAL AGE STUDY TO PRETERM BIRTH IDENTIFICATION

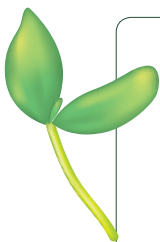
We described the design, organization and implementation strategy of a harmonized study to identify a suitable algorithm comprising simple signs that could be feasibly used by frontline workers to accurately

predict gestational age. If successful, this method comprising simple signs that do not require intense clinical training to assess, will help health workers in lower level facilities to identify preterm births in these settings where the results will be most relevant (South Asia and sub-Saharan Africa). The results of this study will contribute to mechanisms of targeting interventions aimed at reducing risk of death and significant disability from complications of preterm birth in LMICs [8,27].

The AMANHI gestational age study has many strengths: it is being conducted in five countries in South Asia and sub-Saharan Africa and has adequate power to yield very precise estimates on the validity of these simpler methods; it is one of the few LMIC studies with early ultrasound dating as the gold standard; it targets the identification of preterm babies at the population/community level with large sample size, using non-clinician research workers rather than health professionals in health facilities. The findings will therefore be feasible to implement programmatically at first level facilities or in communities by frontline workers with the barest minimum of professional clinical training. Since many of these preterm births are born at home [28–30] or in these lower level facilities in many LMICs, equipping this cadre of staff with a tool to identify preterm births will be a significant step towards ensuring survival and well-being of preterm babies because tested interventions can be deployed in time for their care.

Despite these strengths, the study also has potential challenges. First, with many deliveries occurring at home, it may be possible that assessors will not be able to reach babies within the narrow 72-hour window particularly since social seclusion after delivery (where women and the babies are secluded from public for up to 40 days) is pervasive in these settings [31–34]. Second excluding babies who are deemed “unwell” may selectively bias the assessments towards term babies because preterm babies are more likely to be perceived as unwell by families and hence assessments may not be done. If assessments are mainly conducted on term babies, component signs may lose their discriminatory power to identify preterm babies. The study sites are therefore using innovative methods to reach babies immediately after they are born. These methods include providing incentives to families for prompt birth notification, regular phone calls to women around term and increasing frequency of home visits to women whose pregnancies were older than 34 weeks. In some sites, selected nurses within health facilities are trained to conduct full assessments on babies who are apparently unwell and cannot be examined by research workers. These nurses use the same forms being used by the AMANHI research assistants. If babies are truly unwell to undergo the full assessments, a select limited number of physical signs and anthropometric measures are assessed by these nurses.

In conclusion, if successful, the AMANHI gestational age study will make a significant contribution to improving the ability of frontline health workers in LMICs to accurately assess gestational age in order to discriminate preterm from term infants. This will allow for more effective targeting of interventions for preterm babies in order to improve their survival. Although identification of preterm babies is the most crucial step for now in the planning and delivery of interventions, it remains only the first step and must be accompanied by concurrent deployment of tested interventions to reduce mortality and residual disabilities in survivors.



**Acknowledgements:** The authors acknowledge the contribution made by the AMANHI study staff in host institutions—both local and abroad—and study participants including women, children and their families in the included countries. We also thank other governmental and non-governmental institutions including Ministries of Health, district and provincial governments and other agencies who provided advice, ethical reviews and support to ensure smooth implementation of the AMANHI gestational age study.

**Disclaimer:** RB, AM and SY 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:** The AMANHI gestational age study was funded by the Bill & Melinda Gates Foundation through a grant to the World Health Organization. The funders have played no role in the drafting of the manuscript and the decision to submit for publication.

**Authorship declaration:** All authors participated in the design of the AMANHI gestational age study. RB and AM wrote the first draft of the manuscript based on discussions among the authors. All authors reviewed the manuscripts and made inputs into it. All authors reviewed the final manuscript and agreed to its submission.

**Competing interests:** The authors have completed the Unified Competing Interest form at [www.icmje.org/doi\\_disclosure.pdf](http://www.icmje.org/doi_disclosure.pdf) (available on request from the corresponding author) and declare no competing interests.

- 1 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
- 2 McCormick MC. The contribution of low birth weight to infant mortality and childhood morbidity. *N Engl J Med*. 1985;312:82-90. Medline:3880598 doi:10.1056/NEJM198501103120204
- 3 Liu L, Oza S, Hogan D, Chu Y, Perin J, Zhu J, 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
- 4 Liu L, Johnson HL, Cousens S. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000 (vol 379, pg 2151, 2012). *Lancet*. 2012;380:1308. Medline:23063283
- 5 Kinney MV, Lawn JE, Howson CP, Belizan J. 15 Million preterm births annually: what has changed this year? *Reprod Health*. 2012;9:28. Medline:23148557 doi:10.1186/1742-4755-9-28
- 6 Wang ML, Dorer DJ, Fleming MP, Catlin EA. Clinical outcomes of near–term infants. *Pediatrics*. 2004;114:372-6. Medline:15286219 doi:10.1542/peds.114.2.372
- 7 Woythaler MA, McCormick MC, Smith VC. Late preterm infants have worse 24–month neurodevelopmental outcomes than term infants. *Pediatrics*. 2011;127:e622-9. Medline:21321024 doi:10.1542/peds.2009-3598
- 8 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
- 9 Lawn JE, Cousens SN, Darmstadt GL, Bhutta ZA, Martines J, Paul V, et al. 1 year after The Lancet Neonatal Survival Series—was the call for action heard? *Lancet*. 2006;367:1541-7. Medline:16679168 doi:10.1016/S0140-6736(06)68587-5
- 10 Beck S, Wojdyla D, Say L, Betran AP, Merialdi M, Requejo JH, et al. The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity. *Bull World Health Organ*. 2010;88:31-8. Medline:20428351 doi:10.2471/BLT.08.062554
- 11 Iams JD, Romero R, Culhane JE, Goldenberg RL. Primary, secondary, and tertiary interventions to reduce the morbidity and mortality of preterm birth. *Lancet*. 2008;371:164-75. Medline:18191687 doi:10.1016/S0140-6736(08)60108-7
- 12 Belizán JM, Hofmeyr J, Buekens P, Salaria N. Preterm birth, an unresolved issue. *Reprod Health*. 2013;10:58. Medline:24228615 doi:10.1186/1742-4755-10-58
- 13 Ballard JL, Khoury JC, Wedig K, Wang L, Eilers–Walsman BL, Lipp R. New Ballard Score, expanded to include extremely premature infants. *J Pediatr*. 1991;119:417-23. Medline:1880657 doi:10.1016/S0022-3476(05)82056-6
- 14 Ballard JL, Novak KK, Driver M. A simplified score for assessment of fetal maturation of newly born infants. *J Pediatr*. 1979;95:769-74. Medline:490248 doi:10.1016/S0022-3476(79)80734-9
- 15 Dubowitz LM, Dubowitz V, Palmer P, Verghote M. A new approach to the neurological assessment of the preterm and full–term newborn infant. *Brain Dev*. 1980;2:3-14. Medline:7416439 doi:10.1016/S0387-7604(80)80003-9
- 16 Dubowitz LM, Dubowitz V, Goldberg C. Clinical assessment of gestational age in the newborn infant. *J Pediatr*. 1970;77:1-10. Medline:5430794 doi:10.1016/S0022-3476(70)80038-5
- 17 Eregie CO, Muogbo DC. A Simplified Method of Estimating Gestational–Age in an African Population. *Dev Med Child Neurol*. 1991;33:146-52. Medline:2015982 doi:10.1111/j.1469-8749.1991.tb05093.x
- 18 Finnström O. Studies on maturity in newborn infants. II. External characteristics. *Acta Paediatr Scand*. 1972;61:24-32. Medline:5018575 doi:10.1111/j.1651-2227.1972.tb15899.x
- 19 Narayanan I, Dua K, Gujral VV, Mehta DK, Mathew M, Prabhakar AK. A Simple Method of Assessment of Gestational–Age in Newborn–Infants. *Pediatrics*. 1982;69:27-32. Medline:7054757
- 20 Parkin JM, Hey EN, Clowes JS. Rapid assessment of gestational age at birth. *Arch Dis Child*. 1976;51:259-63. Medline:1275536 doi:10.1136/adc.51.4.259
- 21 Sarris I, Ioannou C, Ohuma EO, Altman DG, Hoch L, Cosgrove C, et al. Standardisation and quality control of ultrasound measurements taken in the INTERGROWTH–21st Project. *BJOG*. 2013;120 Suppl 3:33-7. Medline:23841486 doi:10.1111/1471-0528.12315
- 22 Amanhi study group. Maternal nd, stillbirths mortality cohort s. Burden, timing and causes of maternal and neonatal deaths and stillbirths in sub–Saharan Africa and South Asia: protocol for a prospective cohort study. *J Glob Health*. 2016;6:020602. Medline:27648257
- 23 Marchant T, Jaribu J, Penfold S, Tanner M, Armstrong Schellenberg J. Measuring newborn foot length to identify small babies in need of extra care: a cross sectional hospital based study with community follow–up in Tanzania. *BMC Public Health*. 2010;10:624. Medline:20959008 doi:10.1186/1471-2458-10-624
- 24 Nyqvist KH. Early attainment of breastfeeding competence in very preterm infants. *Acta Paediatr*. 2008;97:776-81. Medline:18460108 doi:10.1111/j.1651-2227.2008.00810.x
- 25 StataCorp. Stata Statistical Software: Release 11.2. College Station, TX: StataCorp; 2009.
- 26 Landis JR, Koch GG. An application of hierarchical kappa–type statistics in the assessment of majority agreement among multiple observers. *Biometrics*. 1977;33:363-74. Medline:884196 doi:10.2307/2529786
- 27 Blencowe H, Lee AC, Cousens S, Bahalim A, Narwal R, Zhong N, et al. Preterm birth–associated neurodevelopmental impairment estimates at regional and global levels for 2010. *Pediatr Res*. 2013;74 Suppl 1:17-34. Medline:24366461 doi:10.1038/pr.2013.204
- 28 Martines J, Paul VK, Bhutta ZA, Koblinsky M, Soucat A, Walker N, et al. Neonatal survival: a call for action. *Lancet*. 2005;365:1189-97. Medline:15794974 doi:10.1016/S0140-6736(05)71882-1

- 29 Lawn JE, Cousens S, Zupan J, Lancet Neonatal Survival Steering T. 4 million neonatal deaths: When? Where? Why? *Lancet*. 2005;365:891-900. Medline:15752534 doi:10.1016/S0140-6736(05)71048-5
- 30 Darmstadt GL, Bhutta ZA, Cousens S, Adam T, Walker N, de Bernis L, et al. 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
- 31 Choi Y, El Arifeen S, Mannan I, Rahman SM, Bari S, Darmstadt GL, et al. Can mothers recognize neonatal illness correctly? Comparison of maternal report and assessment by community health workers in rural Bangladesh. *Trop Med Int Health*. 2010;15:743-53. Medline:20406425 doi:10.1111/j.1365-3156.2010.02532.x
- 32 Bazzano AN, Kirkwood BR, Tawiah-Agyemang C, Owusu-Agyei S, Adongo PB. Beyond symptom recognition: care-seeking for ill newborns in rural Ghana. *Trop Med Int Health*. 2008;13:123-8. Medline:18291010 doi:10.1111/j.1365-3156.2007.01981.x
- 33 Bang AT, Bang RA, Baitule S, Deshmukh M, Reddy MH. Burden of morbidities and the unmet need for health care in rural neonates—a prospective observational study in Gadchiroli, India. *Indian Pediatr*. 2001;38:952-65. Medline:11568371
- 34 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



# Understanding biological mechanisms underlying adverse birth outcomes in developing countries: protocol for a prospective cohort (AMANHI bio-banking) study

AMANHI (Alliance for Maternal and Newborn Health Improvement) Bio-banking Study group); Abdullah H Baqui<sup>1</sup>, Rasheda Khanam<sup>1</sup>, Mohammad Sayedur Rahman<sup>1</sup>, Aziz Ahmed<sup>1</sup>, Hasna Hena Rahman<sup>1</sup>, Mamun Ibne Moin<sup>1</sup>, Salahuddin Ahmed<sup>1</sup>, Fyezah Jehan<sup>2</sup>, Imran Nisar<sup>2</sup>, Atiya Hussain<sup>2</sup>, Muhammad Ilyas<sup>2</sup>, Aneeta Hotwani<sup>2</sup>, Muhammad Sajid<sup>2</sup>, Shahida Qureshi<sup>2</sup>, Anita Zaidi<sup>2</sup>, Sunil Sazawal<sup>3</sup>, Said M Ali<sup>3</sup>, Saikat Deb<sup>3</sup>, Mohammed Hamad Juma<sup>3</sup>, Usha Dhingra<sup>3</sup>, Arup Dutta<sup>3</sup>, Shaali Makame Ame<sup>3</sup>, Caroline Hayward<sup>4</sup>, Igor Rudan<sup>4</sup>, Mike Zangenberg<sup>7</sup>, Donna Russell<sup>7</sup>, Sachiyo Yoshida<sup>7</sup>, Ozren Polasek<sup>5,6</sup>, Alexander Manu<sup>7</sup>, Rajiv Bahl<sup>7</sup>

<sup>1</sup> Center for Maternal and Newborn Health, Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

<sup>2</sup> Department of Pediatrics and Child Health, Aga Khan University, Pakistan

<sup>3</sup> Centre for Public Health Kinetics, India and Public Health Laboratory-IdC, Pemba, Tanzania

<sup>4</sup> University of Edinburgh: Centre for Global Health Research, Old Medical School, Edinburgh, Scotland, UK

<sup>5</sup> University of Split School of Medicine, Split, Croatia

<sup>6</sup> Gen-info Ltd, Zagreb, Croatia

<sup>7</sup> World Health Organization (MCA/MRD), Geneva, Switzerland

## Correspondence to:

Drs Rajiv Bahl & Alexander Ansah Manu  
MCA/MRD, World Health Organization  
Avenue Appia 20  
CH-1211 Geneva 27  
Switzerland  
bahlr@who.int  
makmanu128@gmail.com

**Objectives** The AMANHI study aims to seek for biomarkers as predictors of important pregnancy-related outcomes, and establish a biobank in developing countries for future research as new methods and technologies become available.

**Methods** AMANHI is using harmonised protocols to enrol 3000 women in early pregnancies (8–19 weeks of gestation) for population-based follow-up in pregnancy up to 42 days postpartum in Bangladesh, Pakistan and Tanzania, with collection taking place between August 2014 and June 2016. Urine pregnancy tests will be used to confirm reported or suspected pregnancies for screening ultrasound by trained sonographers to accurately date the pregnancy. Trained study field workers will collect very detailed phenotypic and epidemiological data from the pregnant woman and her family at scheduled home visits during pregnancy (enrolment, 24–28 weeks, 32–36 weeks & 38+ weeks) and postpartum (days 0–6 or 42–60). Trained phlebotomists will collect maternal and umbilical blood samples, centrifuge and obtain aliquots of serum, plasma and the buffy coat for storage. They will also measure HbA1C and collect a dried spot sample of whole blood. Maternal urine samples will also be collected and stored, alongside placenta, umbilical cord tissue and membrane samples, which will both be frozen and prepared for histology examination. Maternal and newborn stool (for microbiota) as well as paternal and newborn saliva samples (for DNA extraction) will also be collected. All samples will be stored at –80°C in the biobank in each of the three sites. These samples will be linked to numerous epidemiological and phenotypic data with unique study identification numbers.

**Importance of the study** AMANHI biobank proves that biobanking is feasible to implement in LMICs, but recognises that biobank creation is only the first step in addressing current global challenges.

Sub-Saharan Africa and south Asia are the sub-regions with the highest proportion of the global burden of over 289 000 maternal deaths, 6 million child deaths and 2.6 million stillbirth [1–3]. To improve survival and secure attainment of developmental poten-

tial of both mothers and their babies, holistic approaches that have future applicability are warranted. A better understanding of the biological mechanisms underlying adverse birth outcomes (such as eclampsia, intrauterine growth restriction, preterm births and stillbirths) and their relationships with various phenotypic, epidemiologic and more importantly epigenetic characteristics will provide a gateway to addressing these challenges.

Evidence from high-income settings suggests that biobanks, which are repositories of biological samples with data linked to individual subjects' characteristics, may provide a sustained platform with infrastructure for research and discovery of biological mechanisms underlying those leading causes of deaths [4–6]. These mechanisms may have roots in endogenous and exogenous factors (eg, genetic composition, nutrition, environment, etc.) [7–9]. Biological molecules comprising glycomes, proteomes, lipidomes, and other metabolomes that circulate in the blood, other human tissues and body fluids have been linked with detection of risks of adverse maternal and fetal outcomes. For instance, soluble fms-like tyrosine kinase (sFlt-1 or sVEGFR-1) have been found to blunt the beneficial effects of proangiogenic factors on maternal endothelium with consequences such as proteinuria and hypertension (pre-eclampsia) [10–12]. Other biomarkers such as soluble endoglin (s-Eng), P-selectin, Cell free fetal DNA (cfDNA), placental protein 13 (PP-13) are being evaluated as predictors of pre-eclampsia and intrauterine growth restriction (IUGR) [13–18].

However, these factors may vary between developed and developing country populations. If these hypothesized biomarkers prove to be important predictors for these adverse outcomes in developing country settings too, they could potentially allow early risk assessment of pregnant woman in order to promote timely referral and optimal management. The current distribution of bio-repositories is skewed to high income countries (HICs) and their focus of research on ageing and chronic diseases, arguably, may have very limited immediate value for low- and middle-income countries (LMICs) [19,20]. Translating knowledge acquired from developed country settings to implement interventions in developing country settings without testing could be a risky investment. Biobanks are largely lacking in sub-Saharan Africa and South Asia.

Prior to recent initiatives such as the Human Heredity and Health in Africa (H3Africa) by the National Institute of Health, USA [21,22], only few African countries had biobanks [23,24] and almost none in south Asia. In developing country biobanks, biological samples are collected at clinics for focused research into specific infectious diseases and, coupled with weak health systems and poor access to health care, data are rarely systematically collected to make them useful in describing population dynamics, disease and death. Large-scale, population-level epidemiological research with capability to acquire biological specimen that can be linked to morbidity and present and future mortality are crucial if the advantages provided by newly-available high throughput analytical technologies are to be exploited to maximise the public health and clinical relevance of research activities [20].

There are suggestions that developing countries do not have the capacity (the legislation, human resource and logistics such as reliable power supply) to establish and maintain biobanks [25]. Whilst this may be true for now, context also plays an important role in diseases and deaths and so capacity building in these settings is paramount. Fortunately, with the ever-increasing availability of and great advances in high throughput technologies at progressively decreasing cost [20], such biological specimen assays to identify biological markers that predict or are associated with pregnancy-related outcomes, growth and development will have direct global relevance especially for LMICs.

The Alliance for Maternal and Newborn Health Improvement (AMANHI) [26,27] initiative aims to establish the best-characterized cohort of pregnant women and their babies in sub-Saharan Africa and south Asia. This study will contribute to advancing knowledge on key pregnancy and birth outcomes on a sustained research platform, prove the principle that such initiatives are feasible to implement in LMICs and develop local capacity around biobanking and the use to explore future hypotheses.

## METHODS

### Study design

This is a population-based, prospective cohort study to collect detailed epidemiological and biological data.

## Objectives

1. To uncover biological markers as predictors of important maternal and foetal outcomes. To do this AMANHI biorepository study will:
  - i. Conduct case–control studies to identify biomarkers that can predict (pre–) eclampsia, preterm births, IUGR, and stillbirths.
  - ii. Replicate the role of genetic variants which have been identified in high–income settings as important determinants of these outcomes through genome–wide association and candidate gene approaches e.g. PSG11, INHBB, ACVR2A, etc. (pre–eclampsia) [11,17,18] and ADCY5, CDKAL1, HHX–IDE, GCK, etc. (preterm births/low birthweight) [13,28,29].
  - iii. Evaluate the validity of the most commonly proposed existing biochemical, nutritional and inflammatory biomarkers (in serum, plasma or urine), which have been identified from hospital–based studies in high income countries (such as PAPP–A, AFP and inhibin–A (INH), homocysteine, sFlt–1, etc.) [14,30]
2. To facilitate future discoveries in maternal, foetal and neonatal health as new and more feasible methods become available, by establishing a bank/repository of biological samples which collects, stores and maintains samples in a harmonized way across sites in developing countries.

## Overview of the AMANHI Biobanking Study protocol

An on–going surveillance system with longitudinal follow–up of a dynamic cohort of women of reproductive age has been established in all sites–Bangladesh (Sylhet), Pakistan (Karachi) and Tanzania (Pemba Island)–in South Asia and sub–Saharan Africa. These sites were selected because they represent predominantly rural populations within geographic regions of the world with the highest maternal and foetal mortality burden and centres have proven excellent track–record of accomplishment of international research, good leadership and on–going trials will not interfere with the AMANHI biobanking study protocols. All resident women with early pregnancies (before 20 gestational weeks), who intend to stay in the study areas for the entire duration of follow–up, are consented for collection of epidemiological data as well as biological samples for the study. The overall sampling frame is 3000 trios (1000 per site), and the collection process is expected to be functional from August 2014 to June 2016. Unpublished results from the parent studies show less than 5% attrition rates in the population. Many study procedures are similar across sites but the following are site–specific settings for the study:

In Bangladesh this study is being implemented within the Bangladesh Maternal Infection and Preterm Birth (MIST) study by the Project for Advancing the Health of Newborns and Mothers (PROJAHNMO) and the John Hopkins University. All pregnant women are identified through monthly pregnancy surveillance by community health workers (CHWs) and village health workers (VHWs). All pregnancies are confirmed via a strip–based pregnancy tests (Diaspot, marketed by BRESTA) administered by CHW. CHWs, supported by senior staff and study physicians (where needed) consent and enrol pregnant women for the study. All enrolled mothers are placed under e–surveillance using cell phone. Contact numbers are exchanged between women and respective CHWs for ease of communication. The study area is served by six health facilities operated by the Bangladesh Ministry of Health and Family Welfare (MOHFW). Four of them are first level health facilities staffed by a physician or paramedic; these facilities provide a range of preventive care services including antenatal care and normal delivery care. The remaining two are 31 bed hospitals with doctors, nurses, in–patient care and basic laboratory facilities.

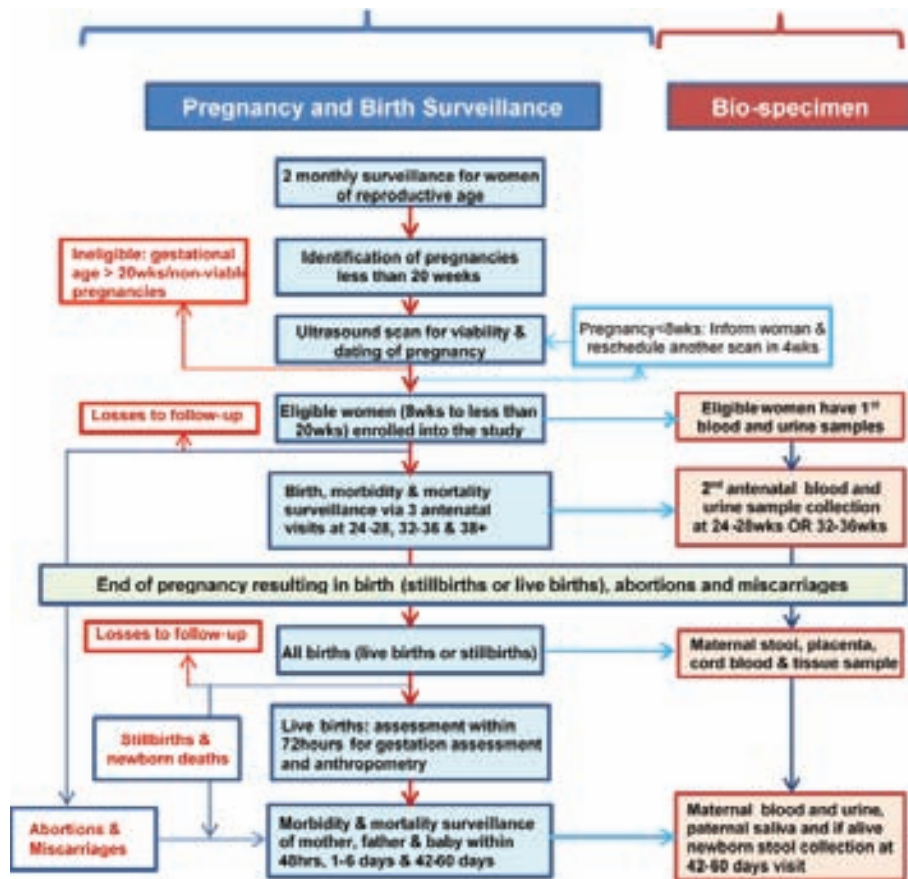
In Pakistan, the Biobanking study will be conducted in two peri–urban communities of Karachi, Ibrahim Hyderi (IH) Goth and Bilal Colony Rehri (RG) Goth within the pregnancy and newborn surveillance as part of the Aetiology of Newborn Infections Study in Asia (ANISA) study. Trained CHWs conduct 3–monthly surveillance to identify pregnancies and conduct confirmatory urine dipstick test. Women with confirmed pregnancies are recruited for an ultrasound scan in a study facility whereupon they are enrolled if found eligible. CHWs will be supported by research assistants and study investigators will follow–up on all enrolled women to recruit them to the study centres for sample collection. Phenotypic and epidemiological data will be collected at the homes of the participants. The study area is served by several health facilities–primary to tertiary–but sample collection will be done at IH health centre and the Aga Khan University Hospital.

In Tanzania, the study is being carried out in two of the four districts or “Shehias” of Pemba, the smaller of the two islands of the Zanzibar archipelago. It is a collaboration between the from Johns Hopkins Uni-

versity and local Investigators from Pemba Health Laboratory–IDC and Ministry of Health, Zanzibar. There is an ongoing delivery and neonatal surveillance at facilities and within communities throughout the island as part of the ongoing chlorhexidine (CHx) application to the umbilical cord trial. Consequently, close contacts and communication has been established between the study team, all the Maternal and Child Health (MCH) staff, TBAs and all facilities where deliveries occur on the island. There is a 2-monthly surveillance for identifying pregnancies by CHWs. CHWs provide menstrual calendars to women to record menstrual periods. If the period is missed at two consecutive months, pregnancy tests are performed by local MCH staff or TBA together with a study supervisor. Women are consented for screening ultrasound at study centres and enrolled if found to be eligible. An elaborate system of tracking of pregnant women with the aid of geographical information systems and exchange of mobile contact numbers between participants’ families and study staff. Immediate information about all deliveries is communicated to a central study call centre and the informant is directly linked to the appropriate study team member and designated MCH staff is arranged. This call also helps to get all information regarding stillbirths and location/directions to the household for home deliveries. Mobile clinics are used for following up enrolled women. Two main facilities in the district are being used for biological sample collection.

**Figure 1** shows an algorithm of the strategy for implementing the AMANHI study. The key components of the study protocol are pregnancy and birth surveillance, ultrasound scan to date pregnancy and to enrol women, epidemiological and phenotypic data collection, and biological sample collection, processing and storage.

*Pregnancy and birth surveillance:* Site specific protocols are as described above. Essentially, in all the three AMANHI sites, all households have been provided unique identification (ID) numbers and have had their geographical coordinates collected and linked to a database. Trained fieldworkers (FWs), predominantly women, perform home visits every 2–3 months to all women of reproductive age in the study area to enquire about pregnancy. If a woman reports or suspects a pregnancy, FWs ascertain the gestational age us-



**Figure 1.** Overview of the study protocol for surveillance and bio-specimen collection within the AMANHI bio-repository study.

ing the date of her last menstrual bleeding and conduct a urine pregnancy test to confirm. Pregnant women who provide consent undergo a screening ultrasound to date the pregnancy more precisely.

**Consenting:** Fieldworkers consent pregnant women, in their local or preferred languages, to undergo a screening ultrasound scan to date pregnancies accurately. They enroll women if the ultrasound–estimated gestational age of the pregnancy is within the eligibility cut-offs of 8 to 19 weeks. Women are consented for the screening scan, follow-up and biosample collection. However, at each sample collection contact, women are required to provide consent for their babies and their own sample taking. Women's husbands (fathers of the babies) also consent for their saliva sample collection.

**Early pregnancy dating ultrasound (gold standard for gestational age) and enrolment:** This is the entry point into the study. Recruited women undergo this dating scan at designated AMANHI facilities to determine foetal viability and measure foetal biometric parameters for gestational age determination. Women whose pregnancies are found to be between 8 and 19 weeks of gestation are considered eligible for enrolment. Harmonized protocols have been developed by a consultant in obstetrics and gynaecology and all study sonographers (medically-trained professionals with additional specialist training in ultrasonography) were trained in strategies to improve image quality and to obtain accurate measurements of the biometric parameters. Crown–rump length (CRL) is the biometric measure of choice for foetuses less than 14 weeks whereas bi–parietal diameter (BPD) and femur length (FL) are taken for foetuses at 14 weeks of gestation or more. Quality assurance measures include expert review of images internally (within site) for a random 10% of study subjects and externally (centrally by the consultant in obstetrics and gynaecology) for a random 5% of subjects. Each image is scored based on a quality checklist and sonographers are provided routine feedback for quality improvement.

**Late Pregnancy Ultrasound:** On a subset of mothers, we are also conducting late pregnancy biometry and testing the accuracy of transcerebellar diameter measurements in late pregnancy for gestational age assessment, as compared with early ultrasound based assessment. In the late pregnancy scans, the biometric measurements include bi–parietal diameter (BPD), head circumference (HC), femoral length (FL), abdominal circumference (AC), and transcerebellar diameter (TCD). For this sub–study, all images are being reviewed with stringent quality control and scoring, and routine feedback provided to sonographers.

## Phenotypic and other epidemiological data collection

Trained study fieldworkers (FWs) conduct four home visits to the enrolled women; at baseline (immediately after enrolment), and at 24–28 weeks, 32–36 weeks and after 37 completed weeks of pregnancy to collect routine study data as detailed in **Table 1**. After delivery, three additional visits are made, within 72 hours, on the fourth to the seventh day after birth and after 42 days [26]. Women or newborns with

**Table 1.** Type of epidemiological and phenotypic data being collected during home visits

MODALITY	DETAILS BEING COLLECTED IN AMANHI BIOREPOSITORY STUDY	TIME OF COLLECTION
Background characteristics	Demographic, socio–economic, other characteristics of the woman and her household and an asset inventory to be used in constructing an asset index for classifying women into wealth quintiles	Baseline visit
Medical history	Previous obstetric and gynaecological history, history of birth defects and congenital anomalies among previous babies, stillbirths and IUGRs; previous medical and surgical history including medicinal prescription drugs taken or being taken for chronic diseases and periodontal diseases	Baseline and antenatal (AN) visits
Risk factors and exposures	Cigarette smoking, alcohol ingestion, smoke from biomass cooking fuels, occupational chemical exposures, strenuous physical work and the use of narcotics and other drugs	Baseline visit
Depression screening	Depression screening using the 9–question patient health questionnaire (PHQ–9)	AN and postnatal (PN) visits
Anthropometry	Maternal weight and height, maternal mid–upper arm circumference and abdominal girth	Baseline, AN and PN visits
Reported morbidity	Questions about any illness or complications during pregnancy, childbirth and the postpartum period. Care–seeking, hospitalizations and treatment received for any morbidity	AN and PN visits
Assessment for pre–eclampsia	Measurement of blood pressure using a digital (Microlife®) sphygmomanometer (calibrated for hyperdynamic circulation in pregnancy) and testing urine for proteins (Uristix®)	All visits except delivery visits
Nutritional assessment	Food frequency questionnaire	One AN visit and Post day 42 PN visit

a morbidity or an abnormal urine test and BP results are referred for appropriate care within health facilities. Core variable tables containing the minimum data that should be collected across all sites have been developed. They ensure uniform data are collected across all sites and specify the type of question, the response options and how the data are presented in the databases (text, numeric or string). These core variables tables were translated into questionnaires for phenotypic and epidemiological data collection across sites.

There is a passive surveillance system in health care facilities where trained hospital FWs collect clinical data on any facility attendance by women enrolled in AMANHI. These data cover the reason for attending the facility, the details of the treatment given and the outcomes.

Harmonized protocols have been used across sites for training study supervisors (who are experienced fieldworkers who have had experience working in health facilities and provided additional training for newborn handling and assessment) to conduct neuromuscular, physical and feeding maturity assessment within 72 hours of the birth on all babies born to women enrolled in the AMANHI study. They also take neonatal anthropometric measurements (head, chest and mid-upper arm circumference, foot length and breast bud diameter) on the baby during these visits. Trained study clinicians validate 5%–15% assessments through repeat assessments also within the 72-hour window period after birth.

After 42 days of the birth, fieldworkers visit families to collect epidemiological and phenotypic data and whereupon women exit the study.

In case of a maternal, foetal (stillbirth) or neonatal death, uniform protocols and training have been provided to supervisors in all sites to conduct home visits to conduct verbal autopsy interviews with reliable informants to elicit the circumstances leading to the death and any relevant information that may help identify the medical cause of death. Harmonized methods (based on principles of the International Classification of Diseases) are being used to assign the causes of deaths and complete death certificates for each death on a specially designed software platform [27].

### Biospecimen collection and processing

A sampling scheme with a sequence of time points to obtain maternal blood and urine, maternal stool, umbilical cord blood and tissue, placenta tissue and membranes, newborn stool and saliva samples (where cord blood is not available), and paternal saliva samples are used. Samples are collected at enrolment, at either 24–28 weeks or 32–36 weeks antenatal visit, at delivery and after 42 days of the delivery. Participants are randomised for antenatal maternal blood and urine collection at either 24–28 weeks or 32–36 weeks gestation in a ratio 2:1. **Table 2** shows the samples collected, the timing of collection and the main extraction from these samples in AMANHI.

Standardized protocols are being implemented across all sites. Blood samples of the mother and from the umbilical cord are collected into pre-labelled tubes, centrifuged and serum, plasma and buffy coat aliquots obtained. Aliquots of whole blood samples are also used to perform HbA1C assay and spots were also placed in Whatman cards and dried. Maternal urine samples are similarly centrifuged and RNALater is mixed with sediments and aliquots taken for storage.

**Table 2.** Timing of collection, processing and/or main extraction for AMANHI biological samples

SAMPLE TYPE	TIMING OF COLLECTION	PROCESSING & PLANNED USE OF THE SAMPLE
Maternal blood	Enrolment, 24–28 weeks or 32–36 weeks, postnatal day 42–60	For DNA extraction, HbA1C analysis, and serum/plasma extraction, aliquoting and storage
Maternal urine	Enrolment, 24–28 weeks or 32–36 weeks, postnatal day 42–60	Uncentrifuged and centrifuged sample, biochemical and pathological analysis
Cord blood sample	At birth	DNA extraction, HbA1C analysis, and serum/plasma extraction, aliquoting and storage
Cord tissue samples	At birth	RNALater, alcohol, flash frozen and formalin sample
Placenta tissue samples	At birth	RNALater, alcohol, flash frozen and formalin sample
Placenta membrane samples	At birth	RNALater, alcohol, flash frozen and formalin sample
Maternal faeces	At birth	Maternal faecal microbiome
Paternal saliva	Antenatal or postnatal	Paternal DNA
Fetal faeces	Postnatal day 42–60	Newborn faecal microbiome

Placenta samples are harvested and processed within 30 minutes of delivery and photographs of the surfaces taken. Full thickness tissues samples are harvested in four areas, three of which have a thin layer of maternal tissues sliced off the surface. Samples of the membranes and umbilical cord are also taken. Placental tissue samples are stored in RNALater, alcohol, or are flash frozen. A sample is also stored in formalin solution for histology. The placenta is then weighed and a third photograph is taken before safe discarding.

Maternal and newborn stool samples are being also collected to assess microbiota around the time of delivery and when newborn feeding is established, respectively. A single sample of paternal taken either during one of the antenatal or postnatal visits) and two of newborn saliva (are taken at 42–60 days post-partum from babies whose cord blood could not be obtained at the time of birth) are collected using an Oragene DNA collection kit for DNA extraction. All biological samples are stored at  $-80^{\circ}\text{C}$ .

**Outcomes.** The main adverse pregnancy outcomes being evaluated in the AMANHI bio-repository study are (pre-) eclampsia, intrauterine growth restriction, preterm birth and stillbirths.

**Sample size considerations.** In this study, 3000 pregnant women will be recruited from the three participating sites (1000 per site) over a period of one year. Sample size and power considerations were based on the 3% prevalence of pre-eclampsia (the rarest outcome) in the population. The 3000 women from the three sites is only sufficient to detect a 1% (absolute) change in the prevalence of pre-eclampsia with 90% power and at 5% significance level if analyses adopt a case-control design with a 1:3 ratio of cases to controls.

**Confidentiality.** All participants are provided with unique study IDs with which they and their families are identified in the study. All data collected from participants are being kept confidential; hard copies of study-related forms are stored in locked cabinets and soft copies are securely stored on dedicated, password-protected servers. These are only accessible to the principal investigators and approved co-investigators. After completion of the study, any sample to be used for analyses will be de-linked from participant's identity and only alphanumeric identification numbers will be used.

**Safety issues in sample collection and processing.** All the personnel involved in sample collection and processing are well trained by the sites. AMANHI-specific standard operating procedures (SOP) have been developed and all phlebotomists and laboratory scientist involved have been trained in study participant care, sample collection, sample processing, bio-safety in sample handling and safe disposal of instruments and materials (**Online Supplementary Document**).

### Bio-specimen storage and security

**Administration.** The principal investigators of the three sites are responsible for the implementation and management of the study. These principal investigators will be responsible for maintenance of optimal quality of the biological samples and will lead all analyses to be conducted as part of the study. They have, in consultation with governmental agencies and ministries of health, academia and key stakeholders, constituted a governance council to take over the management of the biobanks after the main AMANHI analyses are completed. These governing councils will meet at agreed times and at various frequencies to receive, review and approve protocols for studies that require the use of the AMANHI biological samples in the biobank.

**Temperature maintenance.** Maintenance of optimal storage temperature of all biological samples is a key issue in the biobanks and is a core component of the common protocols being used across sites. There are minor site-specific adaptations; for instance, in Karachi and Pemba, to minimise the travel distance between the families and the biobank, field laboratories have been set up and samples are transited through shuttle freezers (also maintained at between  $-76^{\circ}\text{C}$  to  $-86^{\circ}\text{C}$ ) to the Biobank. Temperature logs are maintained for shuttle freezers during transport and main freezers at the biobank.

**Power supply and alert systems.** Availability of reliable power supply is cited as one of the reasons why biobanking may not be feasible in LMICs and so premium was placed on this [25]. All sites have therefore procured additional power back-up generators as well as high capacity uninterruptible power supply (UPS) systems to protect sensitive equipment from power trips, surges and fluctuations. For instance, the Pemba site has installed a 18KV solar power system with 2 days autonomy as one of the power back-up system other than national grid and generators. The biobanks have also been fitted with security alarm systems that immediately report through text messages and emails (to the principal investigators and appropriate technical persons for immediate redress) any temperature fluctuations and power trips irrespec-

tive of whether the back-up systems are activated. All these have been piloted and are being closely monitored.

**AMANHI Biobanking software.** A special Windows-based software has been developed for the capture of the AMANHI biospecimen collection. This software which was developed by the Pemba team is aimed first at reducing transcription errors in the data capture and to align the sample collection process with the common protocol. It has in-built range and consistency checks and will only accept data appropriate for the field in which it is being entered. All samples and study materials are provided labels with encrypted 2D digital signature codes that are read by a 2D scanner into the software. The software has undergone several rounds of testing before deployment to all the sites for use. Training has been provided to all the sites on the installation and the use of the software for data capture and real-time troubleshooting is being done for all sites because individual sites do not have the facility to edit the database. In case of any errors they have to inform the WHO coordinator for the study (Leader, MCA/MRD, Newborn Health Research) who has the password. A specified software expert does the required changes if it is essential and then the database is again locked.

**Coordination, monitoring and quality assurance.** The Maternal, Newborn, Child and Adolescent Health department (MCA) of the World Health Organization (WHO) is centrally coordinating the study. This involves technical input into the implementation as well as raising contracts with the constituent sites for the implementation. The WHO team has also procured the services of various experts from the United States of America, the University of Edinburgh in Scotland and Croatian National Biobank [31] to make specific input into various aspects of the study and to maintain the quality of implementation. Weekly (initially, but now fortnightly) teleconferences are held between team members from all the sites, the WHO team and the experts to discuss implementation challenges and to make decisions on the progress and strategies for the implementation. In preparation for these, sites present a progress report to the WHO which is discussed during the teleconferences. The experts are also commissioned at quarterly intervals to make visits to the sites to monitor the progress and quality of the implementation. After each visit, a detailed feedback is provided to the site and a report is also submitted to the MCA. In between site visits, the experts conduct videoconferences with the sites to monitor implementation. Other monitoring processes have been described in the other AMANHI publications [26,27].

**Ethical considerations.** The biomarker study has received ethical approval from the local and institutional ethics committees of all the three sites: ICDDR,B and John Hopkins University for Bangladesh, Aga Khan University for Pakistan and ZAMREC and John Hopkins University for Tanzania. The protocols were also approved by the WHO Ethics Review Committee and continuing approvals are obtained each new year. There will be no direct benefits of the study to the participants. They will be compensated for the time contribution to the sample collection at the health facilities.

**Plan for analyses.** Analyses of the epidemiological data will be carried out to characterize the women in the study and to link the various reported and measured exposures to the main outcomes. These will include principal components analyses (PCA) at each site to generate asset indices from the inventory of assets collected. These indices will be ranked and divided into quintiles and individual women will be assigned the wealth quintiles for their household. Associations will be explored using simple cross tabulations and tested with either a  $\chi^2$  or Fisher exact tests for their significance. Where applicable, models with robust standard errors will be fitted to explore relationships. Multivariable regression models will be fitted in the risk factors analyses and likelihood ratio tests used to assess statistical significance.

**STEP 1: Testing existing hypotheses.** Simple, highly focused analyses to identify current hypothesized biomarkers associated with the risk of the main outcomes will be conducted using small aliquots of the biological samples. The analyses will seek to:

1. replicate the role of previously identified genetic variants which have been identified as important determinants of these outcomes in high income countries (HICs) through genome-wide association and candidate gene approaches and explore a panel of candidate genetic variants;
2. evaluate the validity of the most commonly proposed existing biochemical markers (in serum, plasma or urine) of these outcomes.

**STEP 2: Exploratory “hypotheses free” research using the biobanks.** Using high throughput “hypothesis free” approaches will be done to advance the science of biomarker-disease pathway discovery. This will include analyses of data from whole-genome arrays, high throughput data on many “-omics” traits, relevant maternal and newborn health outcomes and disease phenotypes. These will be “data driven”

analyses in which novel high throughput technologies will be employed to yield high dimensional genomic, proteomic, lipidomic, glycomic and epigenetic data on large sample sizes to discover and identify entirely new associations and biomarkers. These analyses will require substantial financial investment and will be conducted through collaborations with other experts within countries and around the globe.

**Plan for dissemination of findings.** The results of the AMANHI biomarker study will be disseminated among the public health and maternal and newborn health community of researchers, policy-makers and programme managers. Channels for dissemination will include peer-reviewed journals, print and electronic media and through oral and poster presentations at appropriate fora. In each participating country, there will be extensive briefing on their country-specific and overall study results with interpretation of the potential implications for health programmes to the country.

## DISCUSSION

We present here the protocols used for the harmonised implementation of, to the best of our knowledge, the first population-based harmonized multi-country bio-repository study to be set up in developing country settings. With data on phenotypic and epidemiological characteristics and most importantly epigenetic and biochemical information, this study may have the best characterized cohort of pregnant women and their newborns in the entire developing world; certainly so in south Asia and sub-Saharan Africa. Successful implementation of the study affords the opportunity to explore hypothesized risk factors for adverse pregnancy outcomes and also allow further explorations into biomarker-disease pathway along several dimensions in LMICs.

Hitherto, biobanking has largely been the preserve of developed country settings [19,20]. There is however a systematic challenge in translating research results from such settings to developing country settings. Also whilst non-communicable diseases tend to be the main agenda for biobanks in HICs and hence a lot of emphasis is on adult health outcomes, infectious diseases as well as preventable maternal, newborn and child deaths will require much greater focus in developing country settings [20]. There are also systematic differences in exposures to various risk factors due to differences in culture, workplace environment, lifestyle and access to health care. In contrast to HICs, vital registration systems are non-existent in LMICs, health systems are weak and majority of the population do not have access to health services and hence a facility-based bio-repository will have significant biases and may not be representative. It is believed that when biobanks are sited in LMICs, they will likely address specific needs and equity considerations [20,32,33].

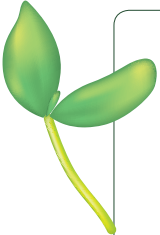
A review by McKinnon et al [7] on birth cohort studies in South East Asia and Eastern Mediterranean found only few studies that saved biological samples. Most of these studies had small sample sizes and were of relatively low quality. Only one of the studies had DNA samples stored for later analyses and none took genetic material from family members. Other logistical challenges such as availability of reliable power supply, the ability to assemble sufficient human resource with the capability of setting up such biobanks in LMICs are often thought of as reasons why such biobanks will not be feasible in LMICs [25] which carry over 95% of the world's burden of morbidity and mortality around childbirth [34].

The AMANHI study proves the principle that such biobanks are feasible to set up in developing country settings. The AMANHI sites are collecting population level data that is very representative of rural populations in the study settings. The inclusion of genetic material and data from family members that can be used directly or "as proxies of exposure" for the identification of "parent-to-origin effects and de novo mutations" is quite novel in these developing country settings. It will therefore contribute to knowledge in the testing of existing hypothesis around risk factors for adverse pregnancy and birth outcomes in these LMICs of sub-Saharan Africa and south Asia who carry the highest burden of these. It will establish a platform for further exploration of new hypotheses and technologies and build local capacity in these low-resource settings for high quality research and ultra-high throughput analyses. Its analytical approach including "hypothesis-free" analyses has been described as one that is potentially free of human biases and may create opportunities for breakthrough discoveries in the biomarker-disease pathways [19]. Its implementation provides a model for adaptation in conducting high impact research in similar settings.

The key desirable attributes of such studies are clear aims and objectives, cultural and social acceptability to both participants and staff involved and low attrition rates. AMANHI has all these attributes. The harmonized implementation including the use of uniform protocols, centralised procurement of equipment, materials and reagents and the strict and rigorous quality control measures will allow for compa-

rability of data across all sites which will allow for pooling and consequent use in the analyses of rare outcomes. A significant limitation of the AMANHI bio-repository study is its inability to inform on biological mechanisms underlying childhood linear growth and neurodevelopment. With the best characterized population based cohort in these LMICs, following up children into the 2nd and 3rd years after birth would have provided opportunities to examine epigenetic factors in pregnancy or early childhood that predict stunting and impaired neurodevelopment, the origins and risk factors of susceptibility to infectious agents and non-communicable diseases.

Establishing these biobanks is only the first step and will be of little value if it is not utilized to address current global challenges. Strategic investment to maximise the utility and gains from this infrastructure is an ethical call. This investment will allow for the use of existing and newly developed ultra-high throughput technologies and develop local capacity to participate and use such technologies to address global health challenges.



**Acknowledgements:** The authors acknowledge the contribution made by the AMANHI study staff in host institutions – both local and abroad – and study participants including women, children and their families in the included countries. We also thank other governmental and non-governmental institutions including Ministries of Health, district and provincial governments and other agencies who provided advice, ethical reviews and support to ensure smooth implementation of AMANHI.

**Disclaimer:** RB, AM, MZ and SY 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:** The AMANHI mortality study was funded by the Bill & Melinda Gates Foundation through a grant to the World Health Organization. The funders have played no role in the drafting of the manuscript and the decision to submit for publication.

**Authorship declaration:** All authors participated in the design of the AMANHI mortality study. The corresponding authors wrote the first draft of the manuscript based on discussions among the authors. All authors reviewed the manuscripts and made inputs into it. All authors reviewed the final manuscript and agreed to its submission.

**Competing interests:** The authors have completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author) and declare no competing interests. IR is the 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.

## REFERENCES

- 1 Wang H, Liddell CA, Coates MM, Mooney MD, Levitz CE, Schumacher AE, et al. Global, regional, and national levels of neonatal, infant, and under-5 mortality during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2014;384:957–79. Medline:24797572 doi:10.1016/S0140-6736(14)60497-9
- 2 World Health Organization. Maternal mortality fact sheet [http://apps.who.int/iris/bitstream/10665/112318/1/WHO\\_RHR\\_14.06\\_eng.pdf?ua=1](http://apps.who.int/iris/bitstream/10665/112318/1/WHO_RHR_14.06_eng.pdf?ua=1). 2014.
- 3 World Health Organization. Trends in maternal mortality 1990–2013 [http://apps.who.int/iris/bitstream/10665/112682/2/9789241507226\\_eng.pdf?ua=1](http://apps.who.int/iris/bitstream/10665/112682/2/9789241507226_eng.pdf?ua=1). 2014.
- 4 Deschenes M, Cardinal G, Knoppers BM, Glass KC. Human genetic research, DNA banking and consent: a question of ‘form’? *Clin Genet*. 2001;59:221–39. Medline:11298677 doi:10.1034/j.1399-0004.2001.590403.x
- 5 Elger BS, Caplan AL. Consent and anonymization in research involving biobanks: differing terms and norms present serious barriers to an international framework. *EMBO Rep*. 2006;7:661–6. Medline:16819458 doi:10.1038/sj.embo.7400740
- 6 Godard B, Schmidtke J, Cassiman JJ, Ayme S. Data storage and DNA banking for biomedical research: informed consent, confidentiality, quality issues, ownership, return of benefits. A professional perspective. *Eur J Hum Genet*. 2003;11 Suppl 2:S88–122. Medline:14718939 doi:10.1038/sj.ejhg.5201114
- 7 McKinnon R, Campbell H. Systematic review of birth cohort studies in South East Asia and Eastern Mediterranean regions. *J Glob Health*. 2011;1:59–71. Medline:23198103
- 8 Batty GD, Alves JG, Correia J, Lawlor DA. Examining life-course influences on chronic disease: the importance of birth cohort studies from low- and middle- income countries. An overview. *Braz J Med Biol Res*. 2007;40:1277–86. Medline:17876486 doi:10.1590/S0100-879X2007000900015
- 9 Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med*. 2006;3:e442. Medline:17132052 doi:10.1371/journal.pmed.0030442
- 10 Maynard SE, Min JY, Merchan J, Lim KH, Li J, Mondal S, et al. Excess placental soluble fms-like tyrosine kinase 1 (sFlt1) may contribute to endothelial dysfunction, hypertension, and proteinuria in preeclampsia. *J Clin Invest*. 2003;111:649–58. Medline:12618519 doi:10.1172/JCI17189

- 11 Troisi R, Braekke K, Harsem NK, Hyer M, Hoover RN, Staff AC. Blood pressure augmentation and maternal circulating concentrations of angiogenic factors at delivery in preeclamptic and uncomplicated pregnancies. *Am J Obstet Gynecol*. 2008;199: 653.e1-10. Medline:18722574 doi:10.1016/j.ajog.2008.06.030
- 12 Khalil A, Muttukrishna S, Harrington K, Jauniaux E. Effect of antihypertensive therapy with alpha methyl dopa on levels of angiogenic factors in pregnancies with hypertensive disorders. *PLoS One*. 2008;3:e2766. Medline:18648513 doi:10.1371/journal.pone.0002766
- 13 Romero R, Nien JK, Espinoza J, Todem D, Fu W, Chung H, et al. A longitudinal study of angiogenic (placental growth factor) and anti-angiogenic (soluble endoglin and soluble vascular endothelial growth factor receptor-1) factors in normal pregnancy and patients destined to develop preeclampsia and deliver a small for gestational age neonate. *J Matern Fetal Neonatal Med*. 2008;21:9-23. Medline:18175241 doi:10.1080/14767050701830480
- 14 Signore C, Mills JL, Qian C, Yu K, Lam C, Epstein FH, et al. Circulating angiogenic factors and placental abruption. *Obstet Gynecol*. 2006;108:338-44. Medline:16880304 doi:10.1097/01.AOG.0000216014.72503.09
- 15 Signore C, Mills JL, Qian C, Yu KF, Rana S, Karumanchi SA, et al. Circulating soluble endoglin and placental abruption. *Prenat Diagn*. 2008;28:852-8. Medline:18702104 doi:10.1002/pd.2065
- 16 Stroustrup Smith A, Huang WY, Wong G, Levine D. Placental chorangiomas associated with markedly elevated maternal chorionic gonadotropin. A case report. *J Reprod Med*. 2003;48:827-30. Medline:14619654
- 17 Thadhani R, Mutter WP, Wolf M, Levine RJ, Taylor RN, Sukhatme VP, et al. First trimester placental growth factor and soluble fms-like tyrosine kinase 1 and risk for preeclampsia. *J Clin Endocrinol Metab*. 2004;89:770-5. Medline:14764795 doi:10.1210/jc.2003-031244
- 18 Levine RJ, Thadhani R, Qian C, Lam C, Lim KH, Yu KF, et al. Urinary placental growth factor and risk of preeclampsia. *JAMA*. 2005;293:77-85. Medline:15632339 doi:10.1001/jama.293.1.77
- 19 Rudan I, Marusic A, Campbell H. Developing biobanks in developing countries. *J Glob Health*. 2011;1:2-4. Medline:23198094
- 20 Sgaier SK, Jha P, Mony P, Kurpad A, Lakshmi V, Kumar R, et al. Public health. Biobanks in developing countries: needs and feasibility. *Science*. 2007;318:1074-5. Medline:18006727 doi:10.1126/science.1149157
- 21 Haines R, Calvert PJ. Is isolation a problem? Issues faced by rural libraries and rural library staff in South Australia. *Aust Libr J*. 2009;58:400-15. doi:10.1080/00049670.2009.10735928
- 22 NIH, Wellcome Trust. NIH and Wellcome Trust Announce Partnership to support Population-based Genome Studies in Africa. [cited; Available from: <http://www.genome.gov/27540084>
- 23 Blankenship L, Haines A. The Institutional Repository: Storing and Sharing Digital Scholarship. *Art Documentation. Bulletin of the Art Libraries Society of North America*. 2008;27:22-5.
- 24 Sirugo G, Schim van der Loeff M, Sam O, Nyan O, Pinder M, Hill AV, et al. A national DNA bank in The Gambia, West Africa, and genomic research in developing countries. *Nat Genet*. 2004;36:785-6. Medline:15284842 doi:10.1038/ng0804-785
- 25 Staunton C, Moodley K. Challenges in biobank governance in Sub-Saharan Africa. *BMC Med Ethics*. 2013;14:35. Medline:24025667 doi:10.1186/1472-6939-14-35
- 26 AMANHI Maternal Morbidity Study group. Burden of severe maternal morbidity and association with adverse birth outcomes in sub-Saharan Africa and south Asia: protocol for a prospective cohort study. *J Glob Health*. 2016;6:020601. Medline:27648256
- 27 AMANHI Mortality Study group. Burden, timing and causes of maternal and neonatal deaths and stillbirths in sub-Saharan Africa and South Asia: protocol for a prospective cohort study. *J Glob Health*. 2016;6:020602. Medline:27648257
- 28 Plunkett J, Muglia LJ. Genetic contributions to preterm birth: implications from epidemiological and genetic association studies. *Ann Med*. 2008;40:167-95. Medline:18382883 doi:10.1080/07853890701806181
- 29 Plunkett J, Doniger S, Orabona G, Morgan T, Haataja R, Hallman M, et al. An evolutionary genomic approach to identify genes involved in human birth timing. *PLoS Genet*. 2011;7:e1001365. Medline:21533219 doi:10.1371/journal.pgen.1001365
- 30 Roberts JM. Preeclampsia: what we know and what we do not know. *Semin Perinatol*. 2000;24:24-8. Medline:10709854 doi:10.1016/S0146-0005(00)80050-6
- 31 Rudan I, Marusic A, Jankovic S, Rotim K, Boban M, Lauc G, et al. "10001 Dalmatians:" Croatia launches its national biobank. *Croat Med J*. 2009;50:4-6. Medline:19260138 doi:10.3325/cmj.2009.50.4
- 32 Singer PA, Daar AS. Harnessing genomics and biotechnology to improve global health equity. *Science*. 2001;294:87-9. Medline:11588248 doi:10.1126/science.1062633
- 33 Daar AS, Thorsteinsdottir H, Martin DK, Smith AC, Nast S, Singer PA. Top ten biotechnologies for improving health in developing countries. *Nat Genet*. 2002;32:229-32. Medline:12355081 doi:10.1038/ng1002-229
- 34 UNICEF. Maternal and newborn health: where we stand: UNICEF; 2009. Report No.: 978-92-806-4318-3.

Ahmad Zia Shams, Ulrike Haug Strategies for prevention of gastrointestinal cancers in developing countries: a systematic review	105
Deborah A Gust, Yi Pan, Fred Otieno, Tameka Hayes, Tereza Omoro, Penelope A Phillips–Howard, Fred Odongo, George O Otieno Factors associated with physical violence by a sexual partner among girls and women in rural Kenya	125
Nigel Unwin, Christina Howitt, Angela MC Rose, T Alafia Samuels, Anselm JM Hennis, Ian R Hambleton Prevalence and phenotype of diabetes and prediabetes using fasting glucose vs HbA1c in a Caribbean population	136
Emily White Johansson, Humphreys Nsona, Liliana Carvajal-Aguirre, Agbessi Amouzou, Helena Hildenwall Determinants of Integrated Management of Childhood Illness (IMCI) non-severe pneumonia classification and care in Malawi health facilities: Analysis of a national facility census	147
Lorraine Sherr, Mark Tomlinson, Ana Macedo, Sarah Skeen, Imca Sifra Hensels, Lucie Dale Cluver Can cash break the cycle of educational risks for young children in high HIV-affected communities? A cross-sectional study in South Africa and Malawi	160
Alberto Barcelo, Armando Arredondo, Amparo Gordillo-Tobar, Johanna Segovia, Anthony Qiang The cost of diabetes in Latin America and the Caribbean in 2015: Evidence for decision and policy makers	170
Saara Parkkali, Bright I Nwaru, Orvalho Augusto, Fatima Abacassamo, Julie Cliff, Elina Hemminki Causes of death among women aged 17–49 years between 2007 and 2010 in Maputo, Mozambique	185
Rima A Abdul-Khalek, Andrea J Darzi, Mohammad W Godah, Lama Kilzar, Chantal Lakis, Arnav Agarwal, Elias Abou-Jaoude, Joerg J Meerpohl, Wojtek Wiercioch, Nancy Santesso, Hneine Brax, Holger Schünemann, Elie A Akl Methods used in adaptation of health-related guidelines: A systematic survey	193
Markus Chan, John J Park, Ting Shi, Federico Martínón-Torres, Louis Bont, Harish Nair; Respiratory Syncytial Virus Network (ReSViNET) The burden of respiratory syncytial virus (RSV) associated acute lower respiratory infections in children with Down syndrome: A systematic review and meta-analysis	202
Tegene Regassa Luba, Zhanchun Feng, Simon Afewerki Gebremedhin, Asfaw N Erena, Abdulsalam MA Nasser, Ghose Bishwajit, Shangfeng Tang Knowledge about mother-to-child transmission of HIV, its prevention and associated factors among Ethiopian women	212
Ravi Prakash Upadhyay, Ranadip Chowdhury, Sarmila Mazumder, Sunita Taneja, Bireswar Sinha, Jose Martines, Rajiv Bahl, Nita Bhandari, Maharaj Kishan Bhan Immunization practices in low birth weight infants from rural Haryana, India: Findings from secondary data analysis	221
Meta van den Heuvel, Wieger Voskuil, Kate Chidzalo, Marko Kerac, Sijmen A Reijneveld, Robert Bandsma, Melissa Gladstone Developmental and behavioural problems in children with severe acute malnutrition in Malawi: A cross-sectional study	235
<b>RESEARCH THEME 1: MEASURING COVERAGE OF ESSENTIAL MATERNAL AND NEWBORN CARE INTERVENTIONS: AN UNFINISHED AGENDA</b>	
Liliana Carvajal-Aguirre, Agbessi Amouzou, Vrinda Mehra, Meng Ziqi, Nabila Zaka, Holly Newby Gap between contact and content in maternal and newborn care: An analysis of data from 20 countries in sub-Saharan Africa	245

Agbessi Amouzou, Vrinda Mehra, Liliana Carvajal–Aguirre, Shane M. Khan, Deborah Sitrin, Lara ME Vaz Measuring postnatal care contacts for mothers and newborns: An analysis of data from the MICS and DHS surveys	253
Deborah Sitrin, Jamie Perin, Lara ME Vaz, Liliana Carvajal–Aguirre, Shane M Khan, Joy Fishel, Agbessi Amouzou Evidence from household surveys for measuring coverage of newborn care practices	263
Agbessi Amouzou, Meng Ziqi, Liliana Carvajal, John Quinley Skilled attendant at birth and newborn survival in Sub–Saharan Africa	274
Kavita Singh, Shane M Khan, Liliana Carvajal–Aguirre, Paul Brodish, Agbessi Amouzou, Allisyn Moran The importance of skin–to–skin contact for early initiation of breastfeeding in Nigeria and Bangladesh	285
Shane M Khan, Ilene S Speizer, Kavita Singh, Gustavo Angeles, Nana AY Twum–Danso, Pierre Barker Does postnatal care have a role in improving newborn feeding? A study in 15 sub–Saharan African countries	294
Kavita Singh, Paul Brodish, Mahbub Elahi Chowdhury, Taposh Kumar Biswas, Eunsoo Timothy Kim, Christine Godwin, Allisyn Moran Postnatal care for newborns in Bangladesh: The importance of health–related factors and location	308
Liliana Carvajal–Aguirre, Vrinda Mehra, Agbessi Amouzou, Shane M Khan, Lara Vaz, Tanya Guenther, Maggie Kalino, Nabila Zaka Does health facility service environment matter for the receipt of essential newborn care? Linking health facility and household survey data in Malawi	320
Rebecca Winter, Jennifer Yourkavitch, Wenjuan Wang, Lindsay Mallick Assessment of health facility capacity to provide newborn care in Bangladesh, Haiti, Malawi, Senegal, and Tanzania	331
<b>RESEARCH THEME 2: HEALTH POLICY AND SYSTEMS RESEARCH</b>	
Kerri Wazny “Crowdsourcing” ten years in: A review	342
<b>RESEARCH THEME 3: HEALTH TRANSITIONS IN CHINA</b>	
Stéphane Verguet, Gillian Tarr, Cindy L Gauvreau, Sujata Mishra, Prabhat Jha, Lingrui Liu, Yue Xiao, Yingpeng Qiu, Kun Zhao Distributional benefits of tobacco tax and smoke–free workplaces in China: A modeling study	356
Rongjing Ding, Ping Ye, Shuiping Zhao, Dong Zhao, Xiaowei Yan, Yugang Dong, Jihu Li, Yuqin Ran, Dayi Hu; on behalf of the DYSIS–China Study Investigators Effect of physician characteristics and knowledge on the quality of dyslipidemia management and LDL–C target goal achievement in China: Subgroup analysis of the Dyslipidemia International Study	369
Peige Song, Yuhang Du, Kit Yee Chan, Evropi Theodoratou, Igor Rudan; on behalf of the Global Health Epidemiology Research Group (GHERG) The national and subnational prevalence and burden of age–related macular degeneration in China	381
Kit Yee Chan, Xue Li, Wanjing Chen, Peige Song, Nuen Wing Katy Wong, Adrienne N Poon, Wenyan Jian, Ireneous Y Soyiri, Simon Cousens, Davies Adeloje, Aziz Sheikh, Harry Campbell, Igor Rudan; on behalf of Global Health Epidemiology Research Group (GHERG) Prevalence of chronic obstructive pulmonary disease (COPD) in China in 1990 and 2010	397
Peige Song, Jiawen Wang, Kajo Bucan, Evropi Theodoratou, Igor Rudan, Kit Yee Chan; on behalf of the Global Health Epidemiology Research Group (GHERG) National and subnational prevalence and burden of glaucoma in China: A systematic analysis	408

- Peige Song, Yezhou Liu, Xinwei Yu, Jingjing Wu, Adrienne N Poon, Alessandro Demaio, Wei Wang, Igor Rudan, Kit Yee Chan; on behalf of the Global Health Epidemiology Research Group (GHERG)  
Prevalence of epilepsy in China between 1990 and 2015: A systematic review and meta-analysis 426
- RESEARCH THEME 4: ACCELERATING KANGAROO MOTHER CARE**
- Tanya Guenther, Sarah Moxon, Bina Valsangkar, Greta Wetzel, Juan Ruiz, Kate Kerber, Hannah Blencowe, Queen Dube, Shashi N Vani, Donna Vivio, Hema Magge, Socorro De Leon-Mendoza, Janna Patterson, Goldy Mazia  
Consensus-based approach to develop a measurement framework and identify a core set of indicators to track implementation and progress towards effective coverage of facility-based Kangaroo Mother Care 437
- Kondwani Chavula, Dyson Likomwa, Bina Valsangkar, Richard Luhanga, Lydia Chimtembo, Queen Dube, Wasihun Andualem Gobezie, Tanya Guenther  
Readiness of hospitals to provide Kangaroo Mother Care (KMC) and documentation of KMC service delivery: Analysis of Malawi 2014 Emergency Obstetric and Newborn Care (EmONC) survey data 449
- RESEARCH THEME 5: INVESTING IN GLOBAL HEALTH AND DEVELOPMENT**
- Jasmine Turner, Hugo Pigott, Mark Tomlinson, Mark JD Jordans  
Developmental assistance for child and adolescent mental health in low- and middle-income countries (2007–2014): Annual trends and allocation by sector, project type, donors and recipients 458
- RESEARCH THEME 6: MONITORING EMTCT IN SOUTH AFRICA**
- Anna Larsen, Mireille Cheyip, Getahun Aynalem, Thu-ha Dinh, Debra Jackson, Nobubelo Ngandu, Witness Chirinda, Mary Mogashoa, Gupreet Kindra, Carl Lombard, Ameena Goga  
Uptake and predictors of early postnatal follow-up care amongst mother-baby pairs in South Africa: Results from three population-based surveys, 2010–2013 469
- RESEARCH THEME 7: IMPROVING COVERAGE MEASUREMENT**
- Mufaro Kanyangarara, Melinda K Munos, Neff Walker  
Quality of antenatal care service provision in health facilities across sub-Saharan Africa: Evidence from nationally representative health facility assessments 483
- RESEARCH THEME 8: RESEARCH THEME 8: ALLIANCE FOR MATERNAL AND NEWBORN HEALTH IMPROVEMENT**
- AMANHI (Alliance for Maternal and Newborn Health Improvement); Abdullah Baqui, Parvez Ahmed, Sushil Kanta Dasgupta, Nazma Begum, Mahmoodur Rahman, Nasreen Islam, Mohammad Quaiyum, Betty Kirkwood, Karen Edmond, Caitlin Shannon, Samuel Newton, Lisa Hurt, Fyezah Jehan, Imran Nisar, Atiya Hussain, Naila Nadeem, Muhammad Ilyas, Anita Zaidi, Sunil Sazawal, Saikat Deb, Arup Dutta, Usha Dhingra, Said Moh'd Ali, Davidson H. Hamer, Katherine EA Semrau, Marina Straszak-Suri, Caroline Grogan, Godfrey Bemba, Anne CC Lee, Blair J Wylie, Alexander Manu, Sachiyo Yoshida, Rajiv Bahl  
Development and validation of a simplified algorithm for neonatal gestational age assessment – protocol for the Alliance for Maternal Newborn Health Improvement (AMANHI) prospective cohort study 496
- AMANHI (Alliance for Maternal and Newborn Health Improvement) Bio-banking Study group); Abdullah H Baqui, Rasheda Khanam, Mohammad Sayedur Rahman, Aziz Ahmed, Hasna Hena Rahman, Mamun Ibne Moin, Salahuddin Ahmed, Fyezah Jehan, Imran Nisar, Atiya Hussain, Muhammad Ilyas, Aneeta Hotwani, Muhammad Sajid, Shahida Qureshi, Anita Zaidi, Sunil Sazawal, Said M Ali, Saikat Deb, Mohammed Hamad Juma, Usha Dhingra, Arup Dutta, Shaali Makame Ame, Caroline Hayward, Igor Rudan, Mike Zangenberg, Donna Russell, Sachiyo Yoshida, Ozren Polašek, Alexander Manu, Rajiv Bahl  
Understanding biological mechanisms underlying adverse birth outcomes in developing countries: protocol for a prospective cohort (AMANHI bio-banking) study 507



*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, Managing Editor

Edinburgh University Global Health Society

Centre for Global Health Research

Usher Institute of Population Health Sciences and Informatics

University of Edinburgh

Teviot Place

Edinburgh EH8 9AG, Scotland, UK

Phone +44 (0) 131 651 5142 Fax +44 (0) 131 650 6909

Rachael.Atherton@ed.ac.uk

**Journal design:** Snježana Engelman Džafić for LASERplus, Zagreb, Croatia

**Realisation:** LASERplus, Zagreb, Croatia, [www.laser-plus.hr](http://www.laser-plus.hr)

EDITORIAL

Liliana Carvajal–Aguirre, Lara ME Vaz, Kavita Singh, Deborah Sitrin, Allisyn C Moran, Shane M Khan, Agbessi Amouzou  
*Measuring coverage of essential maternal and newborn care interventions: An unfinished agenda* 2

Yan Guo  
*Development of global health research in China* 7

NEWS

*Regions* 9

*Agencies* 18

*Resources* 27

VIEWPOINTS

David Beran, Margaret Ewen, François Chappuis, Tim Reed, Hans Hogerzeil  
*Pharmaceutical industry, non-communicable diseases and partnerships: More questions than answers* 37

Margaret MacDonald  
*Why ethnography matters in global health: The case of the traditional birth attendant* 40

Gunisha Kaur  
*Chronic pain in refugee torture survivors* 44

Vidushi Neergheen-Bhujun, Almas Taj Awan, Yusuf Baran, Nils Bunnefeld, Kit Chan, Thomas Edison dela Cruz, Dilfuzza Egamberdieva, Simon Elsässer, Mari-Vaughn V. Johnson, Shoji Komai, Andrey L. Konevega, John H. Malone, Paul Mason, Rothsophal Nguon, Ross Piper, Uttam Babu Shrestha, Milica Pešić, Alexander Kagansky  
*Biodiversity, drug discovery, and the future of global health: Introducing the biodiversity to biomedicine consortium, a call to action* 47

Rebecca Heidkamp; NEP Working Group  
*The National Evaluation Platform for Maternal, Newborn, and Child Health, and Nutrition: From idea to implementation* 52

Emily Wilson, Lois A. Park, Scott Zeger, Timothy Robertson  
*Reimagining statistical analysis for evidenced-based policy making: Early experiences using Stats Report* 57

Allisyn Moran, Tanya Marchant  
*Measuring coverage of essential maternal and newborn care interventions: An unfinished agenda to define the data matrix for action in maternal and newborn health* 62

PAPERS

Victoria B Chou, Ingrid K Friberg, Mervyn Christian, Neff Walker, Henry B Perry  
*Expanding the population coverage of evidence–based interventions with community health workers to save the lives of mothers and children: an analysis of potential global impact using the Lives Saved Tool (LiST)* 64

Rohan J Vilms, Lotus McDougal, Yamini Atmavilas, Katherine Hay, Daniel P Triplett, Jay Silverman, Anita Ra  
*Gender inequities in curative and preventive health care use among infants in Bihar, India* 75

Brian T Chan, Amrose Pradeep, Lakshmi Prasad, Vinothini Murugesan, Ezhilarasi Chandrasekaran, Nagalingeswaran Kumarasamy, Kenneth H Mayer, Alexander C Tsai  
*Association between internalized stigma and depression among HIV-positive persons entering into care in Southern India* 85

Sapna Desai, Ajay Mahal, Tara Sinha, Joanna Schellenberg, Simon Cousens  
*The effect of community health worker–led education on women’s health and treatment–seeking: A cluster randomised trial and nested process evaluation in Gujarat, India* 94

(continued on the inside)

