Global Health Action

2014

Special Issue: Epidemiological Transitions – Beyond Omran's Theory

Guest Editor: Barthélémy Kuate Defo

COACTION PUBLISHING

UMEÅ UNIVERSITET Umeå Centre for Global Health Research

Global Health Action

Special Issue: Epidemiological Transitions – Beyond Omran's theory

Guest Editor: Professor Barthélémy Kuate Defo, University of Montreal, Canada

Published: 2014

CONTENTS

Global Health Action

PART I

Beyond the 'transition' frameworks: the cross-continuum of health, disease and mortality framework

Barthélémy Kuate Defo*

Public Health Research Institute and Department of Demography, University of Montreal, Montreal, Quebec, Canada

*Correspondence to: Barthélémy Kuate Defo, Public Health Research Institute and Department of Demography, University of Montreal, C.P. 6128 Succursale Centre-ville, Montreal, Quebec H3C 3J7, Canada, Email: barthelemy.kuate.defo@umontreal.ca

This paper is part of the Special Issue: *Epidemiological Transitions – Beyond Omran's Theory*. More papers from this issue can be found at http://www.globalhealthaction.net

The planning, development and sustainable implementation of health policies and health systems ought to be based on precise measurements and understandings of prevalence and incidence of communicable and non-communicable diseases, accidents and other disabilities, given past and current demographic and epidemiological profiles in societies as well as how they are predicted to change over time. Equally crucial is the need to understand and appreciate the underling mechanisms and influential factors of these changes, and their monetary and non-monetary costs and implications to individuals, families, communities and governments in the global context. In specific contexts, it may be the interactions between factors from different levels and categories of determinants, and their timing and sequencing during the life courses, which are critical to the health of individuals and populations and how the health care system responds to health problems. To advance knowledge and promote action, various theoretical perspectives, notably the epidemiological transition theory (1, 2), have been used in an attempt to both describe and understand local, national and global patterns in demographic and epidemiological profiles within and across societies, given the multiple domains of health (3).

The epidemiological transition theory was first formulated by Abdel R. Omran (1) to describe quite accurately the shift in demographic and disease profiles reflecting historical experiences of populations in Europe and North America from the mid-18th century through the 1950s. Over the past four decades, the focus in numerous academic and research settings has been on this theoretical perspective for training and research. It has also been used extensively for research and discussion on the changing demographic and epidemiological profiles in developing countries. By and large, it has been used as a main conceptual framework in discussing how disease patterns change over time from predominantly infectious diseases to chronic non-communicable diseases. Since the 1980s, this theory has been challenged on its applicability in low- and middle-income countries (LMICs) where valid and reliable morbidity and mortality data over long time periods are often lacking or incomplete. The validity of Omran's model has also been questioned for failing to recognize and analyze the importance of cultural and social beliefs and values, political forces and health policy in understanding epidemiological profiles, especially in developing countries. Understandably, with improvements in survival at early ages in tandem with increasingly growing proportions of adult and elderly populations as well as the emergence of new infectious diseases, such as HIV/AIDS and the re-emergence of old ones such as tuberculosis, cholera, polio and dengue fever, disease and mortality patterns in populations of LMICs have been changing in unprecedented ways. These countries are largely faced with scarcity of adequate data for health policy and planning, with the double burden of communicable and non-communicable diseases, with their health systems still mainly ill-prepared to face the challenges of quality care and affordable health care services, with a sizeable proportion of their populations living in chronic poverty and unmet basic needs, with no access to clean water, and with inexistent or sub-standard sanitation systems. Over the past 20 years or so, these compounding situations have given rise to renewed interest in the patterns of demographic and epidemiological profiles of developing countries and whether existing theoretical

Citation: Glob Health Action 2014, **7**: 24804 - http://dx.doi.org/10.3402/gha.v7.24804

perspectives in global health research can provide some broad guidance. In response to those concerns, the United States National Academy of Sciences organized two workshops on the epidemiological transition in developing countries. One was in 1991 and led to the publication of the workshop proceedings (4); another was convened in 2011 on the topic of epidemiological transition in sub-Saharan Africa and resulted in a workshop summary (5). These two workshops and resulting publications highlight the continued interest in the epidemiological transition as a broad research theme in global health.

This Special Issue of *Global Health Action* is timely to reassess the epidemiological transition theory coined by Omran over 40 years ago and to consider whether it still serves the purpose it was intended for. In doing so, this Special Issue contributes most directly to the fast-growing literature on global health research in the field of population health. It contains 12 articles including this editorial, which provide the latest evidence and brings forth issues related to the epidemiological transition, in response to a call for papers by *Global Health Action* for the Special Issue on 'Epidemiological Transitions: Beyond Omran's Theory'.

This editorial first gives an overview of the other 11 articles published in the Special Issue. These articles review the original formulation of Omran's theory and its applications, reappraise its utility and relevance to contemporary developing countries, and consider its usefulness for current and future demographic and epidemiological changes. Then, with reference to developing countries and African countries more specifically, we discuss the inadequacies of theoretical perspectives – demographic transition, epidemiological transition and health transition – since these perspectives have generally been used to inform the understanding of global health researchers regarding the multilevel influences on population change and epidemiological landscape in societies over time. Next, we propose a new framework that is better suited for guiding and illuminating historical, contemporary and future demographic and epidemiological changes in human populations, especially in LMICs. Special reference is made of African countries where individuals and populations, on average, have been falling excruciatingly behind by most indicators of well-being and development. Finally, we discuss ways in which the new framework and consequent multilevel life course data collection and analyses might inform understandings on underlying mechanisms of demographic and epidemiological changes and the responses of the health system to them.

Together with the editorial, these articles cover the several broad topics related to the epidemiological transition. The first three papers clarify the concepts of the epidemiological transition in parallel with those of the demographic and health transitions, followed by reviews and empirical assessments of their relevance in Africa and other parts of the developing world. My paper starts by presenting similarities and differences between the three perspectives of the demographic transition, the epidemiological transition and the health transition (6). It then considers seven specific conjectures emanating from these perspectives and empirically tests each of them with time series data covering 60 years of population change and mortality statistics for all five regions and 57 countries of Africa, along with cause-of-death data. It is shown that existing concepts provide inadequate frameworks for describing and understanding population health trends in Africa in general and specifically in sub-Saharan African countries. With the notable exception of island African countries such as Mauritius and probably some northern African countries, the overwhelming evidence indicates that over the past 60 years, African countries have not experienced any sustained shift from one epidemiological regime to another nor seen demographic changes and health improvements as predicted from the perspectives of the demographic, epidemiological, and health transition, in contrast to prevailing situations in other developing countries outside of Africa.

In a review of published studies from 1971 to 2013 on mortality transition and associated epidemiological changes in diverse contexts of LMICs, Santosa et al. document substantial variation in empirical evidence supporting and contradicting Omran's original propositions, and underscore the critical role of social determinants of health in contributing to deviations from these propositions (7). They synthesize some new evidence of such deviations, such as in Nauru where high mortality from infectious diseases gave way to quite high mortality from diabetes, circulatory disorders and accidents over a short period without any appreciable increases in life expectancy; in Mexico with overlapping burdens of disease and an increasing trend of non-communicable diseases at younger ages during 1922–1955 due to poverty and unaffordability of healthcare; and in the native Indian population in Canada following a different epidemiological profile from the general population. They stress the need for a new evidence-based framework of patterns of changes in causes of death and disruptions in health due to emerging risks, which focuses on the underlying mechanisms and cause-specific mortality changes that result.

Zuckerman et al. argue that from anthropological and epidemiological perspectives, the original epidemiological transition theory first formulated by Omran has several limitations (8). Hence, they modify it within an expanded evolutionary framework. They label the 'first epidemiological transition' to coincide with the Neolithic Period and the Agricultural Revolution, the 'second epidemiological transition' to typify the Omran's classic formulation of the epidemiological transition theory,

Citation: Glob Health Action 2014, **7**: 24804 - http://dx.doi.org/10.3402/gha.v7.24804

and the 'third epidemiological transition' to represent the situation of emerging and re-emerging infectious diseases occurring in the modern era. They do their categorization by using the hygiene hypothesis to explain the increased incidence of chronic inflammatory diseases (CIDs) such as allergic and autoimmune diseases or to explain the emergence and re-emergence of infectious disease rise in CIDs; the concept of a third epidemiological transition is used to explain the increase in emerging and re-emerging infections. Building from the socio-ecological model recognizing that a broad array of systems and interrelated determinants of health operate either synergistically or antagonistically in modern epidemiology, they discuss the implications of their categorization for the understanding of the complex and multiple dimensions of health and disease over time as well as for clinical practice, global health policies, and future epidemiological research which can contribute to improving population health.

Since the International Conference on Population and Development in Cairo in 1994, the agenda of sexual and reproductive health has been brought to the forefront in women's health in LMICs. Yet, with rising aging populations and higher proportions of females than males surviving to old ages, non reproductive health conditions such as non-communicable diseases are becoming a fundamental public health concern in those countries. Using the example of depression and Type 2 diabetes co-morbidity in India, Mendenhall and Weaver establish that women are increasingly confronting these diseases within the complexities of the full spectrum of health concerns covering invariably communicable and non-communicable conditions (9). They call into question the existing paradigm of diseases categorization and propose a movement away from the traditional distinctions between 'chronic' and 'acute', 'communicable' and 'non-communicable' diseases; they make the case that in fact these conditions often occur together in most societies of LMICs. They echo the call for a move beyond diseased-focused model in public health to new public health paradigms rethought in light of the challenges of aging populations. Mendenhall and Weaver argue that women living in LMICs have distinctively unique experiences as they face social and health problems compared to women living in developed nations. They warn against bias of research from high-income nations in construing LMIC women's experiences and contributing to knowledge displaced from women's social experiences or policies and programs disconnected from the social, economic, and cultural factors surrounding women's mental and physical health problems in LMICs often due to socially-driven inequalities. They argue for a life course approach encompassing the role of social and economic determinants of health in women's lifetimes. Mendenhall and Weaver advocate an integrative approach that is *health*-focused as opposed to the disease-focused approach which dominates clinics and public health agendas as well as global health dialogues and funding structures, as the co-occurrence of mental and physical health problems gains recognition in the public health agenda, a more nuanced understanding of socio-cultural influences on women's lifetime health is crucial.

Two studies focus on India. Yadav and Arokiasamy assess the structural changes in the patterns of morbidity and mortality in India for understanding India's progress in epidemiological transition (10). They find structural changes in disease patterns concomitant with the transformation in the age pattern of morbidity and mortality. During the last four decades from 1970 to 2007, India moved quickly from the dominance of child and adult mortality to a progressive phase with the dominance of old age mortality. By the mid-1990s, the burden of communicable diseases increased considerably in adult and old ages. Using data from multiple sources, they suggest that all geographical regions of India have experienced a rise in morbidity accompanied with marked fall in mortality, despite notable heterogeneity among the states (e.g. highest morbidity rate of 255/1,000 persons in Kerala and lowest morbidity rate of 33/1,000 persons in poorer states of Jharkhand).

The secular decline of mortality in Kerala during the last century set this Indian state apart from the others and made Kerala a success story by most accounts. Thomas and James examine the pattern of mortality by cause of death and associated changes in human longevity in Kerala since the beginning of the 20th century, to see if changes in mortality rates and causes of death were characterized by a transition in mortality to the adult ages and if there was a shift in the patterns of causes of death from infectious to chronic, degenerative, life style diseases (11). They find that the major reduction in mortality occurred between 1951 and 1970 in Kerala. They also suggest that there is an ongoing epidemiological transition in the recent decades in Kerala whereby more deaths due to non-communicable diseases such as cardiovascular diseases, neoplasm, accidents and injuries are occurring, than from infectious diseases and maternal and child deaths.

Two studies are based on data from rural and urban settings of Africa. The relevance of the theoretical perspective of the epidemiological transition has been little assessed in urban Africa. Masquelier et al. use a rich database from monthly reports of deaths by cause (1900–1907), published estimates (1931–1951) and micro data from death registers (1976–2012) to summarize evidence on trends in mortality by cause of death in Antananarivo (Madagascar), to shed some light on the timing and pace of the mortality decline as well as on changes in cause-of-death patterns (12). They show that the onset of the secular mortality decline in Antananarivo

was ascribed to anti-parasitic and anti-microbial medicine and that the health care system has played a crucial role in mortality reduction in this urban setting despite recurrent political crises and limited public resources. From a theoretical perspective, Antananarivo has experienced mortality falls, reversals and stalls over time, with important setbacks particularly in the mid-1980s and the coexistence of infectious diseases and nutritional deficiencies with non-communicable diseases. It is only after 1990 that a sustained fall in mortality from infectious diseases has been observed. Most deaths have been captured in the vital registration system of this city as far back as the 1960s in Antananarivo, and trends in under-five mortality derived from death registers tend to be consistent with estimates from Demographic and Health Surveys (DHS) for the recent periods. Hence, it is feasible to set up civil registration of death in major African cities for monitoring changes in patterns of mortality by cause and responses of the health care system performance to health problems through health interventions.

Migration and urbanization per se were not the focus in classic formulations of theoretical perspectives of the demographic transition theory or the epidemiological transition theory. With over half of the world population now residing in urban settings where natural increase is playing an increasing role in population change and distribution, the topic of migration and health in the context of the epidemiological change has gained prominence in recent years. Collison et al. take advantage of the availability of longitudinal demographic and health data on temporary rural–urban migration of rural residents from the northeast of South Africa, to analyze trends in temporary migration and mortality and how they are related over time in this setting (13). Temporary migration is related to mortality from communicable diseases, but this association is inconsistent over time. For instance, there is a strong negative association between temporary migration of males and communicable disease mortality early during the observation period; in contrast, the association of temporary migration and mortality turns positive in the latter part of the observation period. However, in this study of the evolution of the relationship between temporary migration status and causes of death where the permanent residents who formed the baseline category were not necessarily a homogenous group during the study period, several selective processes including those directly related to health could not be ruled out.

Two papers consider health policy priorities in the context of epidemiological changes. The burden of mental, neurological, and substance use disorders in South Africa like in many LMICs has been increasing over time, and co-morbidities between these disorders and other diseases including HI/AIDS, diabetes, stroke, and epilepsy make them a public health concern. Specific challenges face South Africa's mental health system and there is limited evidence on economic assessments of mental health in sub-Saharan Africa. Jack et al. summarize current understandings and highlights key knowledge gaps on the direct and indirect costs of these disorders and the cost-effectiveness of their treatment interventions, and consider how mental health services can be scaled up toward universal health coverage in South Africa (14). Their review suggests that the most cost-effective interventions incorporate mental health care into primary care or community services without the use of specialized workers. Such interventions are appealing in South Africa given the high and increasing prevalence of these disorders and comorbid chronic conditions.

Metta et al. provide a narrative review of how the existing public policy environment, health system and community actions are dealing with non-communicable diseases in Tanzania (15). Like in many African countries, there is a lack of a policy for the rising burden of non-communicable diseases within the existing health care system in Tanzania. This hampers the development and implementation of effective strategies for the prevention and control of these diseases and their risk factors at the individual, family and community levels.

Finally, one paper deals with data needs for research on the epidemiological and demographic changes. Health information is notoriously deficient in the vast majority of developing countries, notably in Africa. To reliably document and appraise epidemiological changes for suitable health policy and planning in such settings, Byass, de Savigny and Lopez propose a practical and strategic approach to health information development. This approach focusses on a minimum dataset involving three interweaved components (16). The first component entails a continuous, reliable and unbiased documentation of age- and sex-specific mortality by major causes of deaths in the population using routine civil registration with vital statistics that are enhanced with mortality surveillance systems through verbal autopsy where necessary. They provide supporting evidence of a growing capacity development for producing and using cause-of-death data at country and sub-national levels through sentinel mortality surveillance systems such as the Health and Demographic Surveillance Site (HDSS) data from the INDEPTH Network with verbal autopsy or through sample registration with verbal autopsy (SAVVY). In the absence of accurate information on cause of death, verbal autopsy methods for LMICs are becoming increasingly standardized, adapted and simplified through machine coding of causes of death. The second component is a biennial documentation of exposure to the top 10 major risk factors for the leading causes of mortality by age and sex using population-based nationally representative surveys such as the

Citation: Glob Health Action 2014, **7**: 24804 - http://dx.doi.org/10.3402/gha.v7.24804

DHS being carried out since the late 1980s in these countries. The third component consists of an annual documentation of essential preventive and curative interventions for these major causes and risk factors at the district-level within national health information systems. These authors also discuss some critically important questions to be addressed, including: the cost-effective strategies for integrating these dataset within the capacity development of national health information systems; the methodological implications for upgrading national health information systems to reliably and timely capture epidemiological changes; and the ethical and political issues to ensure sustainable improvements in national health information.

Overall, this set of 11 articles provides a global and quite representative picture of a range of topics of interest to researchers, planners, policymakers and the international community. These articles shed light on trends in the changing disease and cause-of-death mortality patterns, how policymakers may use or have been using this information to make decisions about the priorities for the health sector. All these papers have emphasized the importance of collecting quality data on disease, mortality by cause of death, and risk factors that contribute to them. They have also highlighted the limitations of existing theoretical perspectives.

Inadequacies of the demographic, epidemiological, and health transitions for global health research

Omran's concept of the epidemiological transition is situated at the confluence of the concept of the demographic transition (17) – which preceded it – and the concept of health transition (18–20) – which followed it. The common feature of the three frameworks is that mortality transition is inherent to each of them. The demographic transition embodied the mortality transition and the fertility transition, migration and other demographic phenomena being generally treated as intervening variables. The epidemiological transition is based on the mortality transition of the demographic transition and expands the scope of this transition framework by incorporating the secular changes in disease patterns in tandem with secular changes in mortality; mortality decline is expected to trigger fertility decline generally with some time lag from the onset of the mortality decline. By and large and for all useful purposes, the health transition remains an ambiguous concept, which in empirical studies has been operationalized as an extension or a revision of the epidemiological transition in low-income countries (19), middle-income countries (18), and high-income countries (20).

Notwithstanding their merits for the description of demographic and epidemiological changes and disease patterns in populations of Europe and North America through the 1950s, the critiques over the years of these frameworks in dealing with the complexity of changes in the patterns of mortality and morbidity have revealed their limitations. In essence, they are descriptive models and not explanatory frameworks, and they cannot be used either as a theory of general validity or as a technical tool for health policy and planning, especially in most environments of the developing world. This is the case especially in sub-Saharan African countries (6). One criticism of these 'transition' frameworks is their inflexibility in stipulating a stage-wise linear approach, treating the population as an undifferentiated unit and in an oversimplification of the transition patterns, which do not fit neatly into either historical periods or geographic locations. This stage-wise approach to the demographic transition, the epidemiology transition and the health transition, has drawn the most criticism. The linear progression they suggest is in question given a number of variants describing at different spatial scales and population sub-groups, the complexity of demographic and epidemiological changes. Another criticism concerns the timetable, thresholds and number of stages in the 'transition' frameworks. As we show below, there is a need for different explanatory frameworks of complex changes in health, disease and mortality for guiding the development of data collection and methods of analysis in research for action in health promotion and health policy at the global, regional, national and local levels.

Why should the demographic transition theory be revisited?

The demographic transition is an interpretative description of historical changes in vital rates from high to low mortality and fertility and the associated trends in population growth in the process which began around 1,800 with declining mortality in Europe, in response to industrialization characterized by inherently different packages of social and economic factors. These secular changes are accepted as a definite succession of stages: during the pretransition stage, mortality and fertility are high; during the transition stage, first mortality and then fertility decline, causing a period of robust population growth followed by a deceleration to slow population growth, moving toward low fertility, long life and population aging (17, 21–23).

The demographic transition and the epidemiological transition (hence the health transition) stipulate mortality decline as a precondition for fertility decline (1, 17, 22, 24), thereby precluding the possibility that mortality declines may not be followed, with a lag of 50 years or more generally assumed, by fertility decline. But what will be the consequences if mortality declines and fertility does not? This situation has been shown to happen in several African countries and regions over the

past 60 years of historically unprecedented mortality reductions throughout the continent (6). The theoretical perspectives of the demographic, epidemiological and health transitions preclude this uncovered African situation. The data from Europe also showed that once marital fertility had dropped by as little as 10%, the decline spread rapidly whether or not infant mortality had already declined (25). In fact, the demographic transition in particular has not succeeded at predicting levels of mortality or fertility or the timing of the fertility decline in Africa. This is because the initial explanation for the demographic behavior during the transition tended to be ethnocentric, relying almost exclusively on the contention that what happened to the now-developed countries should happen to other countries in some predictable fashion. This transition is expected to spread to all parts of the world with a projected completion of 2,100 (22). The influential preconditions in African countries are considerably different from what they were when the now-industrialized countries began their transition. Demographic transition theory is notoriously inappropriate for predicting or explaining past and future trends in mortality, fertility and population growth, especially in African regions and countries (6) and is pretty much of no use for population policy. Prior to undergoing the transition, few of the now-developed countries had birth rates and death rates as high as those of most African countries over the last 60 years and currently for several countries. Internal economic development emerged as a sufficient though not a necessary cause of mortality and fertility reductions in industrialized countries; in contrast, prevailing conditions of mortality declines in Africa resulted from foreign aid coupled with public health measures and medical technology brought for disease prevention and control. Moreover, the two elements which affected the onset and sustainability of fertility reduction in developing countries included government policy intervention and new levels of communication in mass media.

The second critique has to do with the culture. There are regional patterns, along cultural and linguistic lines, in fertility and mortality patterns and life expectancy trajectories among African countries. We found that these African puzzling historical patterns occur in contiguous areas and countries that are culturally similar (same language, common ethnic background, similar lifestyles), even though the levels of urbanization and economic development are different. In particular, this applies to HIV/AIDS given its regional mapping in Africa. The health, disease and mortality patterns and population change more broadly over the last 60 years in Africa occurred in the context of widely differing political, social, economic, and demographic conditions which are quite distinct from those experienced by developed countries.

Finally, the demographic transition's end-point is still far from clear and remains debatable (26). The end-point of the demographic transition was supposed to be an older stationary and stable population corresponding with replacement fertility of 2.1 children on average, zero population growth, and life expectancies higher than 70 years. The expected stabilization of population and convergence in birth and death rates has yet to emerge (26), just as the three-stage demographic transition is far from starting in many African countries (6).

Why should the epidemiological transition theory be revisited?

The concept of the epidemiological transition was formulated as a model for integrating epidemiology with demographic changes in human populations (1, 2, 27). The sequence of events marking these changes represents an important trade-off between mortality and morbidity as a result of the interaction between epidemiological and demographic processes (28). There are several limitations to this framework.

First, for a number of scholars (18, 19, 26, 29–42), the epidemiological transition remains a conceptually weak concept in describing and explaining the epidemiology of population change. Paul Farmer has pointed out that the epidemiological transition is a deeply ambiguous framework when infectious diseases have remained so omnipresent in the global health context (43).

Second, the epidemiological transition concept is ill defined, and therefore cannot be put into operation without ambiguity, given the main problem with identifying the beginning of the epidemiological transition on the basis of changes in cause-of-death patterns (20, 44).

Third, Omran's epidemiological transition theory has been criticized for being overly focused on mortality and fertility at the expense of morbidity and its risk factors, including an insufficient account of the role of poverty in determining disease risk and mortality, especially in less developed countries (18, 20, 45, 46). This criticism is reminiscent of the Omran's concept drawing on Notestein's formulation of the demographic transition (17), just like other stage-wise formulations such as the health transition.

Fourth, there is an overemphasis on mortality rather than disease causality and morbidity, thereby failing to understand the correlation between the causes of death and the actual morbidity that people experience during their lives (35, 36). Thus, the model of epidemiological transition is compromised by the uncertain nature of the mechanisms that drive progress through the transition lives (47, 48). It has been argued that the epidemiological transition 'fails to grasp the global nature and the historical sequence of the mortality transition as it spread', and that it is 'insufficiently epidemiological in

Citation: Glob Health Action 2014, **7**: 24804 - http://dx.doi.org/10.3402/gha.v7.24804

that its focus was the changing causes of death rather than the changing causes of patterns of illness' (22: 160).

Fifth, the relative role and importance of infectious diseases (IDs) and non-communicable diseases remain unsettled (43, 47–51). On the one hand, Mackenbach (36) asserts that 'degenerative and man-made diseases' is a misleading term for conditions such as cancer and cardiovascular diseases which have complex etiologies. On the other hand, a recent debate has emerged on the epidemiological transition regarding what should be considered as 'infectious diseases' (52–55). Condrau and Worboys (41: 153–154) argue that the importance of infectious diseases as a cause of death in the 19th century in England and Wales has been overstated and conclude: 'If infections were not the major causes of death a century ago, then surely any major transition is a chimera'. French scholars have vividly argued that in France, the mortality increase hitherto ascribed to cardiovascular diseases may have been an artifact of cause-of-death misclassification of deaths due to ill-defined causes once deaths are properly distributed at least until 1925 (56). They based their contention on their reconstitution of historical series of deaths classified by cause of death on the basis of a constant definition above and beyond the various revisions of the International Classification of Diseases (56). Put differently, MartiInez and Gustavo (44: 543) raised the question 'Transition . . . towards what?'. Weisz and Olszynko-Gryn (49: 309) has noted: 'If Omran essentially ignored chronic disease in most of his work why did he bother including it in his theory? . . . Omran was a bricoleur who liked connecting everything he knew about a subject'. Not only do some infectious diseases have chronic disease characteristics, but infectious agents and related inflammatory processes are also important in the etiology of a number of chronic diseases and adverse outcomes (57) and preventive programs (58).

Sixth, the epidemiological transition has been criticized for failing to distinguish adequately the risk of dying from any given cause or set of causes from the relative contributions of the various causes of death to overall mortality (59). As patterns of disease and mortality change, there are changes in the relative contribution of different causes to overall mortality that may not reflect changes in actual risk (57). Heuveline et al. (60) have shown that the people in the poorest quintile suffer consistently higher mortality in all three of the major categories of disease (i.e. Group I – Communicable, maternal, perinatal and nutritional conditions; Group II – Non-communicable diseases; Group III – Injuries) used by the World Health Organization than those in the richest quintile, most of the excess mortality being primarily due to the higher risk of communicable diseases. In Africa, both the relative contributions and the actual risk of death from the major cause-of-death categories vary widely across countries, even between countries in the same region, as well as across population groups within a country (61–64). Thus, the epidemiological transition oversimplifies the patterns and relations among risk of mortality, mortality causes, and life expectancy. As substantiated by recent developments in epidemiological methods, the patterns are clearly more complex than simply declining mortality rates from infectious diseases and increasing rates of death from non-infectious diseases (57).

The seventh criticism is that the resurgence of old diseases and emergence of new diseases was not anticipated in the epidemiological transition. In fact, Omran (2) emphasized the fact that whether infectious diseases will ever be extinguished remains a question with a regrettable answer. Indeed, the idea of an 'epidemiological transition interrupted' has recently been proposed and discussed (37, 38), disproving Omran's epidemiological transition conceived as a three-stage transition (1). Our cross-examination of evidence from Africa (6) and empirical evidence both historical and contemporary from other studies (26, 37, 38, 43, 48, 51, 59) warns against the operation of a smooth and uninterrupted progression from stage one to stage three and beyond, which was ultimately proposed by Omran (2).

Why should the health transition be revisited?

Several criticisms of the epidemiological transition just identified apply to the health transition. Moreover, there is no agreed upon definition of health transition or its testable characterization in low-income countries, and the concept can hardly be put into operation without ambiguity (6). Omran (2: 99) stressed that 'all of the transitions involved in both the dependent and independent variables are subject of epidemiological study and, hence, are encompassed by the epidemiological transition. Epidemiology incorporates the scientific capacity to analyze social, economic, demographic, health care, technological and environmental changes as they relate to health outcomes. Classifying all the changes in these variables under the "health transition" would, however, be confusing. Health is a dependent variable of epidemiology, not vice-versa'. Building on the lessons from the historical experiences of developed countries, much of the decline in mortality in the late 18th century and throughout the 19th century in Western countries preceded the development of modern medicine. For Caldwell, the term 'health transition' refers to the driving forces (cultural, social and behavioral change) of improvements in health, with relevance to circumstances of poor economic growth found in many sub-Saharan African countries. It does not address the impact of economic growth and the introduction of modern medicine in such improvements. But, just as one cannot see changes in a society in a stage-wise perspective, so no model can assume that the poor economic growth

performance is homogenously invariant in Africa (e.g. Equatorial Guinea is now considered by the World Bank ranking as a high-income country). Chen et al. (65) also challenged the usefulness of the health transition for policy development.

All of these limitations of existing frameworks for investigating demographic and health changes in the epidemiological landscapes of countries around the world, call for a look anew at ways to better understand the mechanisms underlying the exposure to and occurrence of disease, illness, sickness, death and ensuing phenomena.

Beyond the demographic, epidemiological and health transitions: multilevel eco-epidemiological life course framework for the health, disease and mortality cross-continuum

Research traditions and general strategy

Building from prevailing research traditions in causality, modeling, causal inference and counterfactuals (66–74), health-related analytical frameworks (75–83), and theoretical perspectives just reviewed – demographic transition theory (17), epidemiological transition theory (1, 2) and health transition (18, 19) – we propose a multilevel eco-epidemiological life course framework for the health, disease and mortality cross-continuum for deepening understandings of the demographic and epidemiological changes and multilevel responses of the health care system to them, in the global health context. The framework is depicted in Fig. 1.

This framework links together the mechanisms and processes through which individual micro-level decisions and behaviors, household and family meso-level characteristics, and context-dependent macro-level environments that influence health and longevity aggregate to macro-level demographic and epidemiological profiles, trends and differentials in fertility, mortality, migration, burden of disease and population health patterns in human populations. Such a framework is designed to help construct a precise understanding of these mechanisms for the formulation of research design, planning of interventions and development of health policies which will contribute to the development of healthy societies and promote global health.

Among the many implications of population and health research within a cross-continuum perspective is the potential to improve understanding on how people during their life course stay healthily free of diseases, illness or sickness as well as how long people live over time and space as population ages.

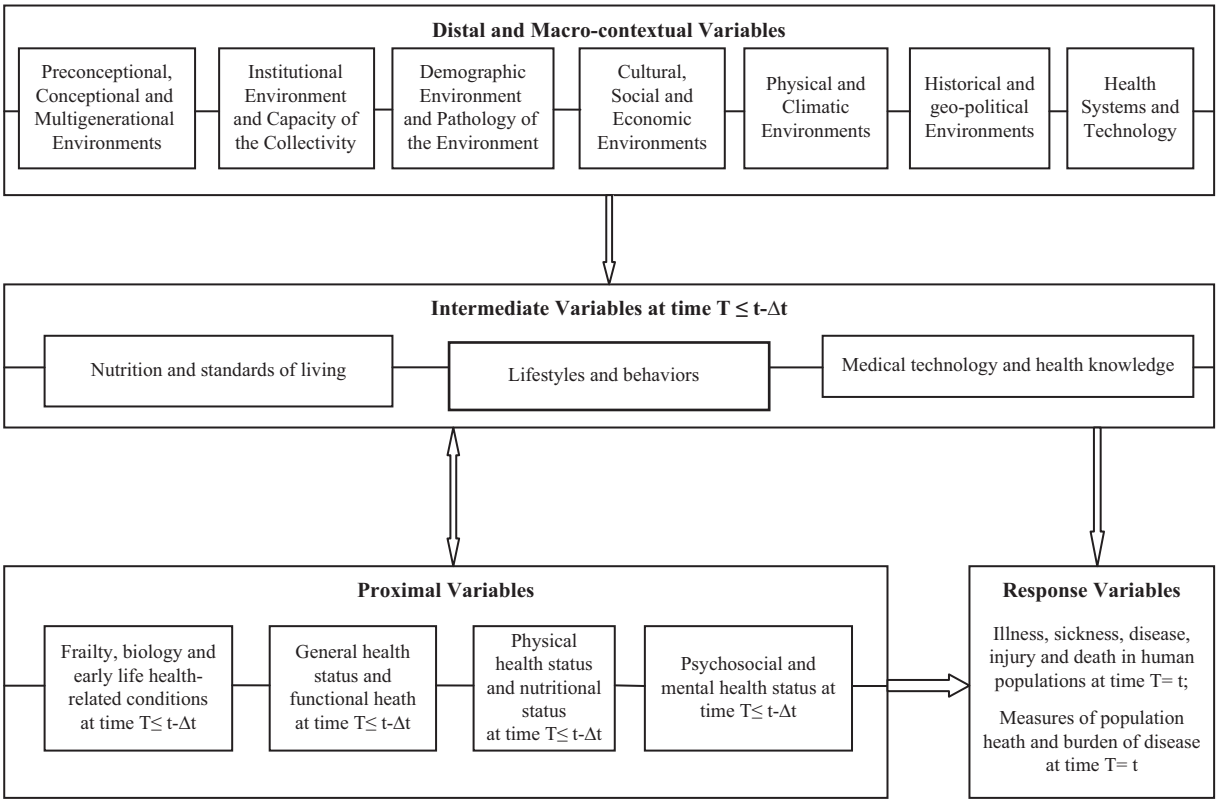

Fig. 1. Multilevel eco-epidemiological life course framework for the health, disease and mortality cross-continuum in human populations.

Citation: Glob Health Action 2014, **7**: 24804 - http://dx.doi.org/10.3402/gha.v7.24804

All indications are that the Millennium Development Goals designed by the United Nations to improve the health of the world's poorest will not be met by 2015 (84), most likely not even by the end of this century in most African countries if current trends continue (85).

In a global context of progresses in human longevity and life expectancy, changes in morbidity and disability trends and secular falls in mortality are expected. The idea that populations' health and disease profiles just as individuals' health and disease patterns change over time because common diseases have their roots in changes in geography, socioeconomic contexts, social norms, culture, socioeconomic, institutional settings, medical contexts, lifestyle factors, and multifaceted environments of human life, makes sense in preventive medicine (86, 87) and is supported by the new public health (88).

Cross-continuum of health, illness, sickness, disease, and death

Given the multiple dimensions of health (3) and the need for improving the accurate of measuring, reporting, and interpreting the health state of individuals and populations (86, 87), we propose a framework which articulates the nested structure of influential factors of individual health or population health across the health continuum. This health, disease and mortality cross-continuum framework embodies the continuously varying multilevel eco-epidemiological contexts of exposure to and occurrence of illnesses, sicknesses, diseases, and deaths of individuals nested in local or national populations. The time dimension is factored throughout the framework within the life course perspective (70–82).

The notion of continuum has been incorporated in many dimensions of population health and health care. For instance, the continuum of care is a concept involving an integrated system of care that guides and tracks patient over time through a comprehensive array of health services spanning all levels of intensity of care (77). The definition of the health/disease boundary being inevitably arbitrary (89), there needs to be measures for illness, sickness, disease along a continuum from conception to death in specific contexts, so as to capture the epidemiological profiles of populations. Progress toward preventing premature deaths across the lifespan and alleviating the burden of disease on populations of developing countries and those of sub-Saharan Africa in particular where such burden is greatest for preventable diseases, requires a better grasp on what are the illnesses, sicknesses, diseases and causes of death encountered by individuals in these populations. This is of paramount importance in many developing countries where culture and perceptions still play a major role in reporting and documenting diseases, disabilities and deaths, so that the burden of disease tends to be heavily underestimated by standard measures of health and disability.

The complexities of what constitutes a disease cannot be underestimated and discussing this requires careful distinction among related, but distinct concepts of illness, sickness, disease and death, on the birth-to-death continuum during the life course across space and time (90). The dictionary of epidemiology (91) defines them following Susser (90) who proposed some useful definitions of these concepts: disease is a physiological or psychological dysfunction; illness is a subjective state of the person who feels aware of not being well; sickness is a state of social dysfunction or a role that the individual assumes when ill. Disease progression is often taken to be reflected in the accumulation and severity of symptoms.

From these definitions, illness refers to the subjective sense of feeling unwell; illness does not define a specific pathology, but refers to a person's subjective experience of it, such as discomfort, tiredness, or general malaise (92). Cultural background influences the way a patient reports symptoms. Susser applied the term sickness to refer to socially and culturally held conceptions of health conditions (e.g. the dread of cancer, the stigma of mental illness, misconceptions about HIV/AIDS), which in turn influence how the patient relates to his or her health status and to the health care system. The social perceptions of a disease may affect the extent to which a condition is concealed, may modify the ways a patient perceives his symptoms, and may affect the likelihood of seeking care irrespective of its affordability in the traditional or modern health care systems. Cultural norms and practices likewise complicate the drawing of the boundary between disease and non-disease, since disease implies a focus on pathological processes that may or may not produce symptoms and that result in a patient's illness (92). In the 'biomedical model' of disease which focuses on pathological processes and on understanding, diagnosing, and treating the physical and biological aspects of disease, there is a great potential of attribution error and misspecification of disease based on the patient's symptoms. Moreover, there have been changes over time in cut-points for what is a normal value or 'normal' when measuring health of people to determine individuals who have a disease and those who do not have a disease. For instance, the cut-points for defining hypertension have changed over time. Prehypertension, a classification for cases where a person's blood pressure is elevated above normal but not to the level considered to be hypertension (high blood pressure) was redefined in 2003 to be blood pressure readings with a systolic pressure from 120 to 139 mm Hg or a diastolic pressure from 80 to 89 mm Hg. This implied revisions of past trends in prevalence of hypertension and increased the incidence and prevalence of hypertension and the associated treatment costs to individuals and the health care system, as most people eventually become hypertensive as they age. It is estimated that more than half of

people over age 60 and approximately three-quarters of people over age 70 have hypertension. Since most biological measurements of risk factors of non-communicable diseases produce a continuous range of values (e.g. blood pressure, body mass index), a cut-point on each of these scales has to be chosen to divide the 'normal' from the 'abnormal' results among a range of findings, which vary from definitely abnormal to definitely normal. Nutritionally, abnormality occurs at both ends of a continuum as being underweight and being overweight are both unhealthy.

The complexity of the relationship between health and illness has inspired important discussions of the nature of medicine and disease (92, 93). The World Health Organization provides a commonly used definition of health: 'the state of complete physical, mental, and social well-being and not merely the absence of disease' (94). This definition emphasizes the importance of physical, mental, and social health, suggesting that a breakdown or shift in any of these components may result in poor health but not necessarily disease. It encourages researchers to move away from simply equating health with the absence of clinical disease. This definition also allows health to be culturally defined and experienced by individuals within their cultural systems (95). In other words, illness categories and conditions resulting in poor health may vary between populations. The processes of demographic and epidemiological changes and their responses are also continuous through time, reflexive because a change in one component is eventually altered by the change it has induced in other components of these processes, and behavioral since these process involve human decisions in the pursuit of goals of living a long and healthy life as a fundamental aspect of human development, with varying means and conditions over time and space. The set of illnesses suffered by individuals and populations at a particular time and the process from individual good health to death or from population good health to mortality, are never the same at the next time while continuously connected. Abrupt changes may be more easily perceived than those that take place more slowly over a longer period of time, and these processes are essentially in a continuous state of flux. As a result, the subject has a frightening complexity and intricacies in developed and developing countries alike. In the former, despite massive and often accurate historical and contemporary data on these processes, a number of theoretical and substantive questions remain a matter of debate (51–59). In the latter, the lack of accurate data and studies has been a deterrent to attempting a good picture of the changing contexts of these processes, especially in the African context. The expectation that various developing countries including African countries go through stages and transitions in population structure and health shaped by changes in fertility and mortality by cause of death as they progress toward fuller industrialization as it was documented in now-developed countries of Europe and North America, has not been materialized in many African countries and regions over the last 60 years where infectious diseases remain the predominant causes of illness and death (61, 62, 6).

Distal, intermediate, proximal, and outcome variables

In the first half of the 20th century, scientific conceptions of causation underwent radical change (66), which has continued to date (67–76). In physics, the development of quantum mechanics led to an acceptance of indeterminism and a rejection of classical concepts of cause and effect; in biology, biostatistics, and the social sciences, there was increasing acknowledgement of the complexity of natural phenomena and the need for a broader concept of causation, which saw causes as multifactorial rather than as a single agent or event. There was also greater acknowledgement of the limitations of scientific measurement and the need to deal with uncertainty about causal mechanisms. Similarly, multilevel and multifactorial causes are involved in the production of health and diseases, and the idea of promoting better health involves tracking the courses of specific diseases. We argue that the continuing rise of chronic infectious diseases and non-communicable diseases usually occurs in a context of uninterrupted interactions between and among communicable and non-communicable diseases across the life course of individuals and communities in developing countries. A multilevel eco-epidemiological model within a life course perspective that captures the health and disease cross-continuum model is better suited for health research in environments typical of those of the vast majority of LMICs.

This perspective recognizes multiple and interrelated levels of causation, offering the possibility for models that are more integrated rather than fragmented. The implication of this framework for research and practice designed for health improvements has three pillars. The first is the life course perspective which requires thinking in terms of changes in causal pathways across the life span when considering health and disease patterns in human populations. Second, the causal models on which to rely must allow for multiple levels of determinants acting in complex and interrelated ways, often synergistically or with feedback loops or reciprocal lines of causality, given the interactions among diseases. Furthermore, we consider that higher-level determinants may have emergent properties above and beyond the aggregate of their constituent parts. Finally, when considering the multiple levels of this life course eco-epidemiological model, we rely on the understanding that disease occurs in individuals, but interventions can occur at any level,

Citation: Glob Health Action 2014, **7**: 24804 - http://dx.doi.org/10.3402/gha.v7.24804

including individual-level, family-level, community-level, resulting in healthy people in healthy communities.

This perspective sees the whole range of determinants as integral to individual, family, community and national health and well-being. Such life course eco-epidemiological model, akin to the modern ecological model of public health practice, stresses the multiple dimensions that constitute our lives, relationships, and environments, hence contributing to wellness or disease, disability and death along the health, disease and mortality cross-continuum in human populations.

Distal and macro-contextual variables (multilevel 4). A distal factor is a factor distant in time to the event-outcome. Distal-level variables include stable dispositional variables and environments that predate the intermediate and the (immediate) proximal contexts. It has been shown that environmental processes influencing the propensity to disease in adulthood operate during the preconceptional, conceptional/fetal, infant phases of life, and throughout the life course. Distal variables include the various environments which protect from or expose to various health-related outcomes: 1) preconceptional, conceptional, and multigenerational environments (80, 81, 96–99); 2) institutional environment and capacity of the collectivity (100, 101); 3) demographic environment and pathology of the environment (86, 87, 102); 4) cultural, social, and economic environments (86, 87); 5) physical and climatic environments (102); 6) historical and geopolitical environments (6); and 7) health systems and technology (103, 104). The idea that social conditions are root causes of disease and health of populations originated from McKeown (105, 106) over half a century ago, and the World Health Organization followed suit only recently by creating a Commission on Social Determinants of Health (107). McKeown's thesis states that the enormous increase in population and dramatic improvements in health that humans have experienced over the past two centuries owe more to changes in broad economic and social conditions than to specific medical advances or public health initiatives.

Intermediate variables (multilevel 3). An intermediate variable in a causal pathway is a variable that causes variation in the response variable and is itself caused to vary by the distal variables (108). Intermediate variables include: 1) nutrition and standards of living (109); 2) lifestyle and behaviors (102); and 3) medical technology and health knowledge (102–104).

The proximal-level variables (multilevel 2). They include the immediate settings, contexts, or conditions prevailing prior to the occurrence of the outcome of interest. A proximal factor is a factor close in time to the event or onset of the behavior of interest as outcome variable. Proximate variables include: 1) frailty, biology, and early life health-related conditions (104, 110, 111); 2) general health status and functional health (102, 111); 3) physical health status and nutritional status (102, 109); and 4) psychosocial and mental health status (112).

The response variables (multilevel 1). They may include both individual-based measures (illness, sickness, disease, injury, and death) as well as measures of population health and burden of disease.

The intermediate variables are determined by the groups of distal factors, including the demographic environment (e.g. population density, urbanization, rural–urban migration, population composition, population structure by age and sex); the capacity of the collectivity (e.g. agricultural and food security); the cultural and social environments (e.g. ethnicity, means of communication such as telephone and cellular phones, norms and practices that have bearing on health and survival in the life course); the economic environment (e.g. per capita gross domestic product, consumer price index, national macroeconomic and microeconomic foundations, fiscal policies); the physical and public health environments (e.g. improved measures of hygiene, sanitation, water access, preservation of the environment); and the health system and technology (e.g. health services, infrastructure and equipment, skilled personnel and quality of care).

Changes in intermediate variables (nutrition and standards of living, lifestyles and behaviors, medical technology and health knowledge including access to and quality of health services) are mostly responsible for the improvements in survival and longevity in contemporary societies just as they accounted for most of the reductions of mortality and increases in life expectancy in developed countries. The technological advances in prevention and treatment of communicable, maternal, perinatal, and nutritional conditions responsible for the majority of deaths have significantly impacted the role of these factors in mortality declines and health improvements over time.

Nonlinearity in the epidemiology of complex health and disease processes

Nonlinearity in the epidemiology of complex health and disease processes is well documented (113). Around the world, it is not infrequent to see during a visit at a health clinic that many patients will come with multiple conditions, both infectious and chronic or communicable and non-communicable. So, the concept of the cross-continuum of health, disease, and mortality proposed here is both conceptually and empirically appealing.

This cross-continuum is chronologically spanning the period from the beginning to the end (for human life) or for a societal health burden (since there is always in a society some degree of health problems, being communicable, non-communicable, or injuries) and with probability varying from 0 to 1. Therefore, it is possible to

develop standardized indicators along the health, disease, and mortality cross-continuum for each society so that they can allow international comparisons. Indeed, various health states are the outcomes ranging from complete absence of ill-health to demise as part of the continuum of health burden (from 0 to 1 on the probability scale) and result from multilevel determinants and consequences at the micro, meso, and macro levels of each society.

The discontinuity in health states is an empirical matter rather than a conceptual or theoretical issue. In particular, any measure of health state is at least in part a reflection of social, cultural, bio-behavioral, economic, physical, and medical influences within the broad context of risk, protective and resilience factors forming a continuum of multilevel states and renewal processes in the nonlinear and dynamic epidemiology of population change.

Data needs to meet the challenges of, and health sector responses to, the cross-continuum of health, disease, and mortality

The proposed framework provides a methodological foundation for designing data collection and conducting analysis of experiments and quasi-experiments as well as probabilistic samples for capturing the changing demographic and epidemiological profiles of national populations and responses from the health sector to these changes. This framework is multidisciplinary and cut across panoply of theories in fields including economics, environment, climatology, physical and environment sciences, social and biomedical sciences; it is transtheoretical. It also embraces the life course perspective given the fact that for most chronic conditions and diseases, their development until they reached the disease state proceeds over years or decades. It is also increasingly established the influences of early life conditions on later life health and survival, along this cross-continuum framework.

The variability in the availability and quality of data on mortality statistics, health and aging and cause of death, and the use of different frameworks for conducting research have invariably produced inconsistent and even contradictory perspectives on change and thus on the implications of health assessments for health policy.

There are growing numbers of datasets spanning multiple time periods in LMICs, and researchers are beginning to face the challenges of how to incorporate this longitudinal dimension into their studies in order to capture any significant change over time. Implementing simultaneously comparative and longitudinal models to such repeated cross-sectional data should provide insightful understanding on the spatial effects that operate at different levels to influence health outcomes and social change. By using the proposed framework for assessing demographic and epidemiological changes, it is possible to collect, analyze and use comparable and high quality data on communicable and non-communicable diseases, accidents and other disabilities. Such endeavor will have implications for the demand for health care, for the types of new health care delivery systems and human resources required for disease prevention and health promotion, and for their costs. Available evidence indicates that Africa in general and sub-Saharan Africa in particular constitutes the poorest and least developed regions in the world with the heaviest burden of disease irrespective of indicators used.

Recent patterns and trends of urban growth in developing countries indicate that over half of the world's total population now lives in urban settings. In many developing countries, the contribution of natural growth to the urban to rural growth differential is higher and is seriously outstripping the capacity of most cities to provide adequate services for their citizens and the deteriorating living conditions of the urban poor. The challenges of achieving sustainable urban development will be particularly formidable in Africa. In the coming decades, the world's rapid urbanization will be one of the greatest challenges to ensuring human welfare and a viable global environment. According to current estimates, cities occupy 4% or less of the world's terrestrial surface, yet they are home to almost half the global population, consume close to three-quarters of the world's natural resources, and generate three-quarters of its pollution and wastes (114). New data collection efforts within the framework proposed here should be designed so as to make progress at understanding the role of migration in population health patterns. Such progress which has been hampered by serious methodological issues, paramount among which are: the lagged effects of migration on health, disease, and mortality patterns and differentials between rural and urban settings; the interactions between demographic and epidemiological responses, including interactions of migration patterns with concurrent changes in the demographic and epidemiological profiles of populations; and the effects of political shocks and social crises which are historically adamant in most African countries and hamper health development by perpetuating vulnerabilities in populations and insubstantialities of institutions.

With the availability of most standard software suites (e.g. SPSS, SAS, and R), estimation of complex models using this framework is now possible in most research environments globally.

Conclusions

The epidemiological transition has been useful in laying out an overarching perspective on changing demographic and diseases patterns in developed countries at least through the 1950s. Its various criticisms suggest that it is

Citation: Glob Health Action 2014, **7**: 24804 - http://dx.doi.org/10.3402/gha.v7.24804

relevant as a way of describing and understanding to some extent the relation among disease and mortality patterns in the course of population change in Western societies until the 1950s, rather than as a universal description or prediction regarding population health patterns enlightening to the formulation of health policies in contemporary societies or in developing countries. The historical and contemporary demography and epidemiology of these countries are quite distinct from historical experiences of the Western societies. Moreover, they are faced with enormous and unprecedented disease burdens in tandem with ill-equipped, poorly funded and often dysfunctional health care system in social contexts where the family largely remains the sole source of social security and health insurance for the majority of people faced with disease and risks of premature death.

In his Inaugural Address on the 20th of January 1949, US President Harry S. Truman noted: '. . . More than half the people of the world are living in conditions approaching misery. Their food is inadequate. They are victims of disease. Their economic life is primitive and stagnant. Their poverty is a handicap and a threat both to them and to more prosperous areas. For the first time in history, humanity possesses the knowledge and skill to relieve the suffering of these people'. Over 60 years since this Address, few people will question the maxim that humanity has failed the vast majority of populations in Asia, Latin America and Africa. Why are poverty, malnutrition, disease and suffering, which are all avoidable given the proven interventions and measures which have created healthy environments for living, still rampant in so many societies of the developing world? The 11 articles in this Special Issue coupled with the editorial, have clarified important aspects of the answer to this question posed to the conscience of humanity and there are specific proposals for action made in each article: equipped with this new knowledge and the proposed framework in global health, the words of President Truman should not continue to largely fall on deaf ears of the international community.

Acknowledgements

A preliminary draft of the theoretical framework which is part of this article benefited from comments and suggestions from participants at meetings organized by the United States National Academy of Sciences' Committee on 'The Continuing Epidemiological Transition in Sub-Saharan Africa', held in Washington DC (USA), Accra (Ghana) and Johannesburg (South Africa) between 2009 and 2011. Professor Barthélémy Kuate Defo was chair of that Committee. I thank specifically Peter Byass, Barney Cohen, Alan Lopez, Stig Wall, Jacques Vallin and Richard Suzman for their insightful comments, suggestions, and discussions on the development of this new framework. This study was supported in part by the New Initiatives Grant Program from the Institut de Recherche en Santé Publique (IRSPUM) and the Global Health Competition Grant from the Direction des Relations Internationales (DRI) (Université de Montréal).

References

1. Omran AR. The epidemiologic transition: a theory of the epidemiology of population change. Milbank Q 1971; 49: 509–38.
2. Omran AR. The epidemiologic transition theory revisited thirty years later. World Health Stat Q 1998; 51: 99–119.
3. Chatterji S, Ustün BL, Sadana R, Salomon JA, Mathers CD, Murray CJL. The conceptual basis for measuring and reporting on health. Global Programme on Evidence for Health Policy Discussion Paper No. 45. Geneva: World Health Organization; 2002.
4. National Research Council (1993). The epidemiological transition: policy and planning implications for developing countries. Washington, DC: The National Academy Press.
5. National Research Council (2012). The continuing epidemiological transition in sub-Saharan Africa: a workshop summary. Washington, DC: The National Academies Press.
6. Kuate Defo B. Demographic, epidemiological, and health transitions: are they relevant to population health patterns in Africa? Glob Health Action 2014; 7: 22443. DOI: 10.3402/gha.v7.22443.
7. Santosa A, Wall S, Fottrell E, Högberg U, Byass P. The development and experience of epidemiological transition theory over four decades: a systematic review. Glob Health Action 2014; 7: 23574. DOI: 10.3402/gha.v7.23574.
8. Zuckerman MK, Harper KN, Barrett R, Armelagos GJ. The evolution of disease: anthropological perspectives on epidemiologic transitions. Glob Health Action 2014; 7: 23303. DOI: 10.3402/gha.v7.23303.
9. Mendenhall E, Weaver LJ. Reorienting women's health in low- and middle-income countries: the case of depression and Type 2 diabetes. Glob Health Action 2014; 7: 22803. DOI: 10.3402/gha.v7.22803.
10. Yadav S, Arokiasamy P. Understanding epidemiological transition in India. Glob Health Action 2014; 7: 23248. DOI: 10.3402/gha.v7.23248.
11. Thomas MB, James KS. Changes in mortality and human longevity in Kerala: are they leading to the advanced stage? Glob Health Action 2014; 7: 22938. DOI: 10.3402/gha.v7.22938.
12. Masquelier B, Waltisperger D, Ralijaona O, Pison G, Ravélo A. The epidemiological transition in Antananarivo, Madagascar: an assessment based on death registers (1900–2012). Glob Health Action 2014; 7: 23237. DOI: 10.3402/gha.v7.23237.
13. Collinson MA, White MJ, Bocquier P, McGarvey ST, Afolabi SA, Clark SJ, et al. Migration and the epidemiological transition: insights from the Agincourt sub-district of northeast South Africa. Glob Health Action 2014; 7: 23514. DOI: 10.3402/gha.v7.23514.
14. Jack H, Wagner RG, Petersen I, Thom R, Newton CR, Stein A, et al. Closing the mental health treatment gap in South Africa: a review of costs and cost-effectiveness. Glob Health Action 2014; 7: 23431. DOI: 10.3402/gha.v7.23431.
15. Metta E, Msambichaka B, Mwangome M, Nyato DJ, Dieleman M, Haisma H, et al. Public policy, health system, and community actions against illness as platforms for response to NCDs in Tanzania: a narrative review. Glob Health Action 2014; 7: 23439. DOI: 10.3402/gha.v7.23439.
16. Byass P, de Savigny D, Lopez AD. Essential evidence for guiding health system priorities and policies: anticipating epidemiological transition in Africa. Glob Health Action 2014; 7: 23359. DOI: 10.3402/gha.v7.23359.
17. Notestein FW. Population – the long view. In: Schultz TW, ed. Food for the world. Chicago, IL: University of Chicago Press; 1945, pp. 36–57.

18. Frenk J, Bobadilla J, Sepulveda J, Cervantes J. Health transition in middle-income countries: new challenges for health care. Health Policy Plan 1989; 4: 29–39.
19. Caldwell JC. Basic premises for health transition in developing countries. World Health Stat Q 1998; 51: 121–33.
20. Vallin J, Mesle F. Convergences and divergences in mortality: a new approach to health transition. Demographic Research 2004. DOI: 10.4054/DemRes.2004.S2.2.
21. Keyfitz N. On the momentum of population growth. Demography 1971; 8: 71–80.
22. Lee R. The demographic transition: three centuries of fundamental change. J Econ Perspect 2003; 17: 167–90.
23. Blue L, Espenshade TJ. Population momentum across the demographic transition. Popul Dev Rev 2011; 37: 721–47.
24. Mason KO. Explaining fertility transitions. Demography 1997; 34: 443–54.
25. Coale A, Watkins SC, eds. The decline of fertility in Europe. Princeton: Princeton University Press; 1986.
26. Vallin J. De la mondialisation de la transition au retour des incertitudes (1940–2000). In: Caselli G, Vallin J, Wunsch G, eds. Démographie: analyse et synthèse. Volume V, Histoire du peuplement et prévisions. Paris: Institut National d'Études Démographiques; 2004; pp. 117–70.
27. Omran AR. Population epidemiology: emerging field of inquiry for population and health students. Am J Public Health 1974; 64: 674–9.
28. Omran AR. The world population problem. In: Omran AR, ed. Community medicine in developing countries. New York: Springer; 1974, pp. 107–8.
29. Caldwell JC, Findley S, Caldwell P, Santow G, Cosford W, Braid J, et al., eds. What we know about the health transition: the cultural, social and behavioral determinants of health. Canberra: Australian National University; 1990.
30. Caldwell JC. Introductory thoughts on health transition. In: Caldwell JC, Findley S, Caldwell P, Santow G, Cosford W, Braid J, et al., eds. What we know about health transition: the cultural, social and behavioural determinants of health. Canberra: Australian National University; 1990, pp. xi–xiii.
31. Caldwell JC, Caldwell P. What have we learnt about the cultural, social and behavioral determinants of health? From selected readings to the first health transition workshop. Health Transit Rev 1991; 1: 3–20.
32. Caldwell JC. Health transition: the cultural, social and behavioural determinants of health in the third world. Soc Sci Med 1993; 36: 125–35.
33. Caldwell JC. Population health in transition. Bull World Health Organ 2001; 79: 159–60.
34. Caldwell JC. Demographers and the study of mortality. Scope, perspectives, and theory. In: Weinstein M, Hermalin AI, Stoto MA, eds. Population health and aging: strengthening the dialogue between epidemiology and demography. Ann N Y Acad Sci 2001; 954: 19–34.
35. Salomon J, Murray C. The epidemiologic transition revisited: compositional models for causes of death by age and sex. Popul Dev Rev 2002; 28: 205–28.
36. Small-Raynor M, Phillips D. Late stages of epidemiological transition: health status in the developed world. Health Place 1999; 5: 209–22.
37. Hill K, Vapattanawong P, Prasartkul P, Porapakkham Y, Lim SS, Lopez A. Epidemiologic transition interrupted: a reassessment of mortality trends in Thailand, 1980–2000. Int J Epidemiol 2007; 36: 374–84.
38. Vallin J. Commentary: 'epidemiologic transition' interrupted or sweep to the second stage of 'health transition'? Int J Epidemiol 2007; 36: 384–6.
39. Das V. What do we mean by health? In: Caldwell JC, Findley S, Caldwell P, Santow G, Cosford W, Braid J, et al., eds. What we know about health transition: the proceedings of an international workshop. Canberra: The Australian National University; 1990, pp. 27–46.
40. Palloni A. The meaning of the health transition. In: Caldwell JC, Findley S, Caldwell P, Santow G, Cosford W, Braid J, et al., eds. What we know about health transition: the proceedings of an international workshop. Canberra: The Australian National University; 1990, pp. xvi–xvii.
41. Ruzicka L, Kane P. Health transition: the course of morbidity and mortality. In: Caldwell JC, Findley S, Caldwell P, Santow G, Cosford W, Braid J, et al., eds. What we know about health transition: the proceedings of an international workshop. Canberra: The Australian National University; 1990, pp. 1–26.
42. Zimmet P, Serjeantson S, Dowse G, Finch C. Killed by the 'good life': the chronic disease epidemic adverse effects of lifestyle change in developing Pacific nations. In: Caldwell JC, Findley S, Caldwell P, Santow G, Cosford W, Braid J, et al., eds. What we know about health transition: the proceedings of an international workshop. Canberra: The Australian National University; 1990, pp. 275–83.
43. Farmer P. Social inequalities and emerging infectious diseases. Emerg Infect Dis 1996; 2: 259–69.
44. MartiInez SC, Gustavo LF. Epidemiological transition: model or illusion? A look at the problem of health in Mexico. Soc Sci Med 2003; 57: 539–50.
45. Frenk J, Bobadilla JL, Stern C, Frejka T, Lozano R. Elements for a theory of the health transition. Health Trans Rev 1991; 1: 21–38.
46. Caldwell J, Santow G, eds. Selected readings in the cultural, social and behavioural determinants of health. Canberra: Australian National University; 1989.
47. Mackenbach JP. The epidemiologic transition theory. J Epidemiol Community Health 1994; 48: 329–31.
48. Riley JC. Sickness, recovery and death: a history and forecast of ill health. Iowa: University of Iowa Press; 1989.
49. Weisz G, Olszynko-Gryn J. The theory of epidemiologic transition: the origins of a citation classic. J Hist Med Allied Sci 2010; 65: 287–326.
50. Robine JM. Life course, environmental change, and life span. Popul Dev Rev 2003; 29: 229–38.
51. Armelagos GJ, Brown PJ, Turner B. Evolutionary, historical and political economic perspectives on health and disease. Soc Sci Med 2005; 61: 755–65.
52. Condrau F, Worboys M. Second opinions: epidemics and infections in nineteenth-century Britain. Soc Hist Med 2007; 20: 147–58.
53. Noymer A, Jarosz B. Causes of death in nineteenth-century New England: the dominance of infectious disease. Soc Hist Med 2008; 21: 573–8.
54. Mooney G. Infectious diseases and epidemiologic transition in Victorian Britain? Definitely. Soc Hist Med 2008; 20: 595–606.
55. Condrau F, Worboys M. Second opinions: final response epidemics and infections in nineteenth-century Britain. Soc Hist Med 2009; 22: 165–71.
56. Mesle F, Vallin J. The health transition: trends and prospects. In: Caselli G, Vallin J, Wunsch G, eds. Demography, analysis and synthesis. A treatise in demography. New York: Elsevier; 2006, pp. 247–602.
57. McKeown RE. The epidemiologic transition: changing patterns of mortality and population dynamics. Am J Lifestyle Med 2009; 3: 19S. DOI: 10.1177/1559827609335350.
58. Blackburn H. Cardiovascular disease epidemiology. In: Holland WW, Olsen J, Florey C, eds. The development of

modern epidemiology: personal reports from those who were there. New York: Oxford University Press; 2007, pp. 71–92.
59. Gage TB. Are modern environments really bad for us? Revisiting the demographic and epidemiologic transitions. Am J Phys Anthropol 2005; 48: 96–117.
60. Heuveline P, Guillot M, Gwatkin DR. The uneven tides of the health transition. Soc Sci Med 2002; 55: 313–22.
61. Feachem RG, Jamison DT, eds. Disease and mortality in sub-Saharan Africa. New York: Oxford University Press; 1991.
62. Jamison DT, Feachem RG, Makgoba MW, Bos ER, Baingana FK, Hofman KJ, et al. Disease and mortality in Sub-Saharan Africa. 2nd ed. Washington, DC: The World Bank; 2006.
63. United Nations, Department of Economic and Social Affairs, Population Division (2012). Changing levels and trends in mortality: the role of patterns of death by cause (United Nations publication, ST/ESA/SER.A/318). New York: United Nations.
64. Kuate Defo B. Areal and socioeconomic differentials in infant and child mortality in Cameroon. Soc Sci Med 1996; 42: 399–420.
65. Chen LC, Macfarlane S, Jones DA. Health transition: from research to policy? World Health Stat Q 1998; 51: 137–43.
66. Parascandola M. The epidemiologic transition and changing concepts of causation and causal inference. Rev Hist Sci 2011; 64: 243–62.
67. Karhausen LR. Causation: the elusive grail of epidemiology. Med Health Care Philos 2000; 3: 59–67.
68. Morgan SL, Winship C. Counterfactuals and causal inference: methods and principles for social research. Cambridge: Cambridge University Press; 2007.
69. Pearl J. Causality: models, reasoning, and inference. 2nd ed. Cambridge: Cambridge University Press; 2009.
70. Spirtes P, Glymour C, Scheines R. Causation, prediction, and search. 2nd ed. Cambridge: The MIT Press; 2000.
71. Rubin DB. Causal inference using potential outcomes: design, modeling, decisions. J Am Stat Assoc 2005; 100: 322–31.
72. Shadish WR, Cook TD, Campbell DT. Experimental and quasi-experimental designs for generalized causal inference. Boston, MA: Houghton Mifflin; 2002.
73. Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. 2nd ed. Hoboken, NJ: Wiley; 2002.
74. Rizopoulos D, Lesaffre E. Introduction to the special issue on joint modelling techniques. Stat Methods Med Res 2014; 23: 3–10.
75. Koopman JS, Lynch JW. Individual causal models and population system models in epidemiology. Am J Public Health 1999; 89: 1170–4.
76. Krieger N. Epidemiology and the web of causation: has anyone seen the spider? Soc Sci Med 1994; 39: 887–903.
77. Evashwick C. Creating the continuum of care. Health Matrix 1989; 7: 30–9.
78. Cummins S, Curtis S, Diez-Roux AV, Macintyre S. Understanding and representing 'place' in health research: a relational approach. Soc Sci Med 2007; 65: 1825–38.
79. Hertzman C, Power C. Health and human development: understandings from life-course research. Dev Neuropsychol 2003; 24: 719–44.
80. Maklakov A, Lummaa V. Evolution of sex differences in lifespan and aging: causes and constraints. Bioessays 2013; 35: 717–24.
81. Misra DP, Guyer B., Allston A. Integrated perinatal health framework: a multiple determinants model with a life span approach. Am J Prev Med 2003; 25: 65–75.
82. Harris KM. An integrative approach to health. Demography 2010; 47: 1–22.
83. Mosley WH, Chen LC. An analytical framework for the study of child survival in developing countries. Popul Dev Rev 1984; 10: 25–45.
84. World Bank (2011). Global monitoring report 2011: improving the odds of achieving the MDGs. Washington DC: The World Bank.
85. Kuate Defo B. The importance for the MDG4 and MDG5 of addressing reproductive health issues during the second decade of life: review and analysis from time series data of 51 African countries. Afr J Reprod Health 2011; 15: 3–21.
86. Rose G. Sick individuals and sick populations. Int J Epidemiol 1985; 14: 32–8.
87. Rose G. Strategy of preventive medicine, with commentary by Kay-Tee Khaw and Michael Marmot. Oxford: Oxford University Press; 2008.
88. Baum F. The new public health. 3rd ed. Oxford: Oxford University Press; 2011.
89. Gluckman P, Hanson M. Living with the past: evolution, development, and patterns of disease. Science 2004; 305: 1733–6.
90. Susser MW. Causal thinking in the health sciences. New York: Oxford University Press; 1973.
91. Last JM, ed. A dictionary of epidemiology. New York: Oxford University Press; 1988.
92. Wikman A, Marklund S, Alexanderson C. Illness, disease, and sickness absence: an empirical test of differences between concepts of ill health. J Epidemiol Community Health 2005; 59: 450–4.
93. Hausman DM. Measuring or valuing population health: some conceptual problems. Public Health Ethics 2012; 5: 229–39.
94. World Health Organization. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19–22 June 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948. Available from: http://www.who.int/about/definition/en/
95. Hruschka DJ. Culture as an explanation in population health. Ann Hum Biol 2009; 36: 235–47.
96. Pongou R. Why is infant mortality higher in boys than in girls? A new hypothesis based on preconception environment and evidence from a large sample of twins. Demography 2013; 50: 421–44.
97. Bharadwaj P, Lakdawala L. Discrimination begins in the womb: evidence of sex-selective prenatal investments. J Hum Resour 2013; 48: 71–113.
98. Bianchi S. A demographic perspective on family change. J Fam Theory Rev 2014; 6: 35–44.
99. Mare RD. A multigenerational view of inequality. Demography 2011; 48: 1–23.
100. Siddiqi A, Hertzman C. Towards an epidemiological understanding of the effects of long-term institutional changes on population health: a case study of Canada versus the USA. Soc Sci Med 2007; 64: 589–603.
101. World Bank (2012). World development report 2012: gender equality and development. Washington, DC: The World Bank.
102. Kunitz SJ. The health of populations: general theories and particular realities. New York: Oxford University Press; 2007.
103. Inhorn MC, van Balen F, eds. Infertility around the world: new thinking on childlessness, gender, and reproductive technologies. Berkeley: The University of California Press; 2002.
104. Mirea L, Yang J, Paterson AD, Shah V, Bassil KL, Lee SK, et al. Relationship of mode of conception and sex concordance with mortality/morbidity in preterm twins. Twin Res Hum Genet 2013; 16: 985–93.

105. McKeown T. The origins of human disease. Oxford: Blackwell; 1988.
106. McKeown T. The role of medicine: dream, mirage or nemesis. Oxford: Blackwell; 1979.
107. Irwin A, Valentine N, Brown C, Loewenson R, Solar O, Brown H, et al. The commission on social determinants of health: tackling the social roots of health inequities. PLoS Med 2006; 3: e106. DOI: 10.1371.
108. Robins J. The control of confounding by intermediate variables. Stat Med 1989; 8: 679–701.
109. Fogel RW. New findings on secular trends in nutrition and mortality: some implications for population theory. In: Rosenzweig M, Stark O, eds. Handbook of population and family economics. Amsterdam: Elsevier; 1997, pp. 433–81.
110. Nobile A. Male excess mortality between biology and culture. In: Pinnelli A, Racioppi F, Rettaroli R, eds. Genders in the life course. New York: Springer Science + Business Media; 2007, pp. 249–81.
111. Fries JF. Compression of morbidity revisited: frailty, heart disease, and stroke – the compression of morbidity paradigm. Am J Prev Med 2005; 29: 164–8.
112. Prince M, Patel V, Saxena S, Maj M, Maselko J, Phillips MR, et al. No health without mental health. Lancet 2007; 370: 859–77.
113. Philippe P, Mansi O. Nonlinearity in the epidemiology of complex health and disease processes. Theor Med Bioeth 1998; 19: 591–607.
114. Redman CL, Jones NS. The environmental, social, and health dimensions of urban expansion. Popul Environ 2005; 26: 505–20.

Global Health Action

PART II

Demographic, epidemiological, and health transitions: are they relevant to population health patterns in Africa?

Barthélémy Kuate Defo*

Public Health Research Institute and Department of Demography, University of Montreal, Montreal, Quebec, Canada

Background: Studies of trends in population changes and epidemiological profiles in the developing world have overwhelmingly relied upon the concepts of demographic, epidemiological, and health transitions, even though their usefulness in describing and understanding population and health trends in developing countries has been repeatedly called into question. The issue is particularly relevant for the study of population health patterns in Africa and sub-Saharan Africa, as the history and experience there differs substantially from that of Western Europe and North America, for which these concepts were originally developed.
Objective: The aim of this study is two-fold: to review and clarify any distinction between the concepts of demographic transition, epidemiological transition and health transition and to identify summary indicators of population health to test how well these concepts apply in Africa.
Results: Notwithstanding the characteristically diverse African context, Africa is a continent of uncertainties and emergencies where discontinuities and interruptions of health, disease, and mortality trends reflect the enduring fragility and instability of countries and the vulnerabilities of individuals and populations in the continent. Africa as a whole remains the furthest behind the world's regions in terms of health improvements and longevity, as do its sub-Saharan African regions and societies specifically. This study documents: 1) theoretically and empirically the similarities and differences between the demographic transition, epidemiological transition, and health transition; 2) simple summary indicators that can be used to evaluate their descriptive and predictive features; 3) marked disparities in the onset and pace of variations and divergent trends in health, disease, and mortality patterns as well as fertility and life expectancy trajectories among African countries and regions over the past 60 years; 4) the rapid decline in infant mortality and gains in life expectancy from the 1950s through the 1990s in a context of preponderant communicable diseases in all African countries; 5) the salient role of adult mortality, mostly ascribed to HIV/AIDS and co-morbidities, since the 1990s in reversing trends in mortality decline, its interruption of life expectancy improvements, and its reversal of gender differences in life expectancies disadvantaging women in several countries with the highest prevalence of HIV/AIDS; 6) the huge impact of wars in reversing the trends in under-five mortality decline in sub-Saharan countries in the 1990s and beyond. These assessments of these transition frameworks and these phenomena were not well documented to date for all five regions and 57 countries of Africa.
Conclusion: Prevailing frameworks of demographic, epidemiological, and health transitions as descriptive and predictive models are incomplete or irrelevant for charting the population and health experiences and prospects of national populations in the African context.

Keywords: *systematic review*; *demographic transition*; *epidemiological transition*; *health transition*; *population health*; *epidemiology of population change*; *global health*; *Africa*; *sub-Saharan Africa*; *cross-national comparison*; *sex differentials*

Responsible Editor: Nawi Ng, Umeå University, Sweden.

*Correspondence to: Barthélémy Kuate Defo, Public Health Research Institute and Department of Demography, University of Montreal, C.P. 6128 Succursale Centre-ville, Montreal, Quebec, Canada H3C 3J7, Email: barthelemy.kuate.defo@umontreal.ca

This paper is part of the Special Issue: *Epidemiological Transitions – Beyond Omran's Theory*. More papers from this issue can be found at http://www.globalhealthaction.net

Received: 23 July 2013; Revised: 20 April 2014; Accepted: 20 April 2014; Published: 15 May 2014

Over the past six decades, life expectancy for the world's population increased from 47 years in 1950–1955 to 69 years in 2005–2010. By 2005–2010, life expectancy at birth in the most developed regions was 77 years, while it was 4 years shorter in Latin America and the Caribbean (73 years), 7 years shorter in Asia (70 years),

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

21 years shorter in Africa (56 years), and nearly 24 years shorter in sub-Saharan Africa (SSA) (53 years). These disparities are indicative of differences in the demographic and epidemiological changes that have taken place in various regions around the world. The increases in longevity have been accompanied by a historic shift in the cause-specific mortality risks in human populations, and researchers have developed a number of theoretical frameworks to describe and explain these patterns, including the demographic transition (1), the epidemiological transition (2–11) and the health transition (12–22). Over the past three decades or so, these frameworks have been used extensively by researchers from different disciplinary perspectives to study and compare the population and health experiences of various countries. These frameworks have been applied in studies of mortality, morbidity, health and development, population development, and health development in developed countries, in rapidly changing middle-income countries, and, to a lesser extent, in low-income countries (23–48).

For the vast majority of the past 140,000 years of human existence, birth and death rates were at very high levels – around 30–50 per 1,000 people. The transition to low birth and death rates was termed the 'demographic transition' by Notestein (1). This formulation offered a framework for describing and understanding population change, and during the 1960s through the 1980s it led to intense debates on the reasons for population growth and the means of population control (22, 45). The mortality component of this transition was first elaborated by Omran (2–11), who used the concept of the 'epidemiological transition' to describe and explain transformations in the patterns of disease occurrence and causes of death. By the 1990s, Omran's concept, which he had updated and elaborated over the years (2, 10, 11) had become a classic citation in demography, epidemiology, and public health, and it remains the most cited paper in population health (46). About that same time other authors had developed the concept of 'health transition' as either a revision or an extension of Omran's theory (12, 13, 29, 48) or as a new and distinct concept (19, 43, 49).

The extent to which these existing concepts provide an adequate framework for describing and understanding population health trends in developing in general and especially in Africa is unsettled (23, 28–45, 48–53), particularly in light of three types of health problems facing that continent: 1) the HIV/AIDS epidemic, along with re-emerging infectious diseases such as tuberculosis; 2) the widespread presence of illnesses caused by common infectious diseases (notably malaria and childhood diseases) and malnutrition; and 3) the emerging epidemic of chronic diseases, accidents, and mental disorders. This study is a systematic review and critical assessment of the relevance of these frameworks for describing and understanding health, disease, and mortality patterns in Africa over the past 60 years, which is a sufficiently long period for gauging changes within the historical demography and epidemiology of Africa.

Hypotheses developed to explain the health, disease, and mortality processes as they occurred in the last three centuries in Western Europe and North America have been tested in other developed countries as well as in some middle- and low-income developing countries since the 20th century. However, Africa's historical epidemiology and demography have posed a challenge to researchers and policymakers because of the scarcity of data and evidence-based written sources in large parts of the continent in general and in SSA in particular (31, 33, 39–41, 54–61). Health, disease and mortality processes have been taking place in African countries under various political, cultural, social, economic, demographic, structural transformations and institutional environments over the past 60 years and beyond.

The purpose of this study is two-fold. First, it reviews and clarifies any distinction between the concepts of demographic transition, epidemiological transition and health transition. When it comes to semantics in global health research, these concepts are often used synonymously in various contexts and from different disciplinary perspectives, without due caution to similarities and differences between them regarding their contours, their descriptive and explanatory dimensions, and their prognostic implications. Second, it identifies summary indicators of population health to test how well these concepts apply in Africa. There is no critical appraisal of these concepts in the context of all African regions and countries. As we will see, the relevant evidence points to the need for a new and different perspective for population and health changes in the African context. There are many examples between and within countries and over time where these transition frameworks obviously do not apply (62–70). Instead, the evidence indicates that the social, economic, political, cultural, and demographic contexts relate to the health, disease, and mortality patterns in Africa in ways that are quite different from the understanding derived from these perspectives.

Study methodology

Review on the demographic transition, epidemiological transition, and health transitions

For the review of the extant literature on the demographic transition, epidemiological transition, and health transitions, we searched OAJSE, Scopus, ScienceDirect, Scirus, PubMed, Google Scholar, SciCentral, MUSE, POPLINE, World Bank, World Health Organization (WHO), United Nations databases and publications for reviews and ascertained relevant publications from these and other sources, using subject headings or key words related to demographic transition; epidemiological transition; health

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

transition; population change; population development; population health; health demography; public health in Africa; epidemiology in Africa; African demography; health, disease and mortality in Africa; cause of death in Africa; infectious disease; chronic disease; communicable disease; noncommunicable disease; SSA; Africa. Web pages of organizations active in areas relevant to this study, such as WHO, United Nations, UNICEF, FAO, and regional offices of organizations based in Africa working on or around health, disease and mortality in Africa, were also screened for further pertinent publications. Relevant materials from these searches are used in this study.

Empirical data and methods of analysis

Data sources

For quantitative evidence spanning the last 60 years and involving all five regions and 57 countries forming Africa, data from the online latest estimates in the databases from the United Nations are used for assessing the levels, trends, and patterns in demographic changes as well as mortality statistics and characteristics in Africa from 1950 to 2010. We use the WHO mortality estimates by cause, age, and sex for the year 2008. WHO uses numerous data sources and epidemiological models to estimate the worldwide cause-of-death patterns. Such databases allow us to uncover patterns of demographic and epidemiological structures and processes and hypothesize their causal mechanisms and potential consequences on the health and disease patterns in Africa. National deviations from international patterns or those expected from theoretical perspectives will further enlighten our demographic and epidemiological understandings of the African epidemiological landscape and should pinpoint the specificities and similarities among and within countries given the social and cultural complexities of individual African countries. These databases provide a series of comparable national estimates from all African countries, and the completeness of their documentation reduces the risk of misuse. The reliability of these estimates for SSA in particular has been questioned (71). But they represent the best data available for all African countries since 1950. Although no sub-Saharan African country possesses reliable registration systems, model-based existing estimates of mortality statistics and characteristics for Africa indicate that SSA has the highest burden of disease in the world (64, 66, 72–74).

Methods of analysis

For each summary indicator selected for analysis, we will use a simple statistical approach for assessing changes over time. Let δ be the average yearly rate of change in indicators of interest between two consecutive years. Let μ_i be the proportion of events for a given year i. We have: $\mu_{i+1} = \mu_i (1 + \delta)$. To assess changes over time involving k ($k = 1, \ldots n$) time-dependent indicators, the equation is: $\mu_n = \mu_0 (1 + \delta)^n$, yielding $\delta = [(\mu_n / \mu_0)^{1/n} - 1]$. From this equation, the average yearly rate of change from say 1950 to 2010 is: $\mu_{2010} = \mu_{1950} (1 + \delta)^{60}$. The yearly rate of decline needed to reach the targeted indicator level between 1950 and 2010, according to expectations from demographic, epidemiological and health transitions is readily available. For instance, the MDG4 is to reduce under-five mortality (U5MR) by two-thirds between 1990 and 2015; thus, the targeted U5MR for 2015 is $\mu_{2015} = \mu_{1990}/3$. The MDG5 is to reduce by three-quarters, between 1990 and 2015, the maternal mortality ratio (MMR); hence, the targeted MMR for 2015 is $\mu_{2015} = \mu_{1990}/4$. The yearly rate of decline needed to reach the targeted U5MR rates (respectively MMRs) from 2008 to 2015 set by MDG4 (respectively MDG5) is $\delta = [(\mu_{2015} / \mu_{2008})^{1/7} - 1]$. The waiting times w from 2008, to reach the U5MR (respectively MMR) from 2008 to 2015 set by MDG4 (respectively MDG5), assuming the average yearly rate of change above is $w = [\log(\mu_{2015} / \mu_{2008})]/\log(1 + \delta)$.

Perspectives of the demographic, epidemiological, and health transitions

The general approach to a transition framework is the characterization and explanation of a set of long-run irreversible changes of substantial social, ecological, cultural, political, behavioral or health significance with enduring global generalization, which are experienced as a society is transformed from one state or structure to another. The concept of transition has been used to describe dynamics involved in human history. It has been approached from different disciplinary perspectives. The transition concept was first proposed by demographers to describe the differing patterns of the components of population dynamics between and within countries with varying levels of socioeconomic transformation in the course of population development. Subsequently, the concept of transition has been used to describe historical trends in demographic, epidemiological and nutritional changes in human populations in the context of economic, sociocultural and political changes. It has been conventional wisdom in all earlier formulations to use three phases for the demographic transition (1, 75, 76), the epidemiological transition (2, 10, 11), the nutrition transition (77–81) and the health transition (12–21). These perspectives are often used synonymously, and it is warranted to clarify any distinction between them, in order to assess their relevance to concurrent changes and their implied health and disease patterns.

Demographic transition

Definition

Demographic thinking in transition terms essentially began as early as 1929. It started as a description of demographic changes using a classification of populations into three groups according to different combinations of mortality and fertility levels resulting in three groups of countries with different population growth rates (75).

Independently, Landry (76) described the demographic revolution characterized by three demographic regimes: primitive regime associated with subsistence economies constraining mortality reduction, intermediate regime with fertility decline due to late marriage and celibacy, and modern regime where fertility is an object of conscious limitation. The classic formulation of the demographic transition is ascribed to Frank W. Notestein (1). A fundamental concept in modern demography, the demographic transition is the intermingling description with explanation of the decline from high to low rates of mortality and fertility as historically experienced by populations in high-income societies of Europe, North America and Australia, and conjectured to be a universal principle expected to occur nowadays in populations of middle- and low-income countries. As a descriptive concept, the demographic transition is the characterization of the long-run situation of population change in three transitional regimes: a pre-transition regime with a quasi-equilibrium of high and fluctuating mortality and high fertility allowing a modest population growth or decline; a transition regime characterized by a transitory disequilibrium of declining mortality followed by declining fertility triggering population growth; and a post-transition regime with quasi-equilibrium of low mortality, low and conceivably fluctuating fertility and declining population growth rate. Rises followed by falls in the population growth rates result in the inverted U-shaped demographic transition. From an explanatory perspective, technological change and industrialization as the modernization progress unfolds and paves the way to the urban transition or the progression from low to high urbanization rates, are viewed as the crucible of demographic transition.

Description

The two key elements of the demographic transition are the mortality transition and the fertility transition from long-term quasi-equilibrium high levels to long-term quasi-equilibrium low levels: during the transition process, mortality is expected to decline first, followed by fertility decline with a time lag of 50 or more years between the mortality and fertility declines. The fertility decline is viewed as an adjustment made necessary by the decline in mortality in response to the forces of change. The transition begins with sustained declines in mortality, especially infant and child mortality. It is predicted that all societies will experience such a transitional or lag period during which birth rates exceed death rates by a substantial margin that will trigger rapid population growth. The underlying assumption of the demographic transition is that a population would move from one long-term quasi-equilibrium to another (1, 76, 82–86). During the pre-transitional stage (before 1800 in Europe), life expectancy at birth was less than 40 years with women bearing on average five to eight births. During the transitional stage (from 1800 to 1950 in Europe), life expectancy at birth was between 40 and 65 years with women bearing on average 2.6–4.9 births. In Europe where fertility decline followed mortality decline, the rate of natural increase (birth rates minus death rates) during the transitional period from 1800 to 1950 ranged between 0.5 and 1% per year. During the post-transitional stage (after the 1950s in Europe), life expectancy at birth exceeded 65 years with women bearing on average 2.5 or fewer births.

Explanation

There are two chief driving forces of the demographic transition. On the one hand, mortality decline is viewed as initiating the transition process and is generally considered as the most influential factor of fertility decline. Modern mortality decline was unprecedented in human history. This decline has been explained by the reduction of epidemics through vaccination and better hygiene, improved diagnosis and treatment of disease, reduction of famines, fewer deaths from violence and civil wars, reductions in infant and child mortality, and improved standards of living. On the other hand, the demographic transition amounts to the dual transition of mortality and fertility in response to forces of change including urbanization, economic development and cultural change, individualism, secularism, self-fulfillment, and changes in attitudes about and values related to family life. Certain cultural changes such as an increase in the value placed on investment in child human capital, may trigger both mortality and fertility declines, given the growing evidence pointing to the influential role of conditions in early childhood in determining adult health and human capital.

Epidemiological transition

Definition

The epidemiological transition refers to the long-run shift in mortality and cause-of-death patterns inherent in the secular mortality decline from high to low levels embedded in a series of concurrent changes in population health. The concept was coined in the mid-1960s and published in 1971 in its first formulation by Abdel R. Omran (2). It provides an accurate characterization of the mortality transition from about 1750 through the 1950s ascribed mainly to the decline of infectious diseases, as experienced in Europe and North America. Omran's concept was initially influenced by theories of social evolution published by Riesman et al. (87) who divided social types into three historical stages based on Notestein's formulation of the demographic transition (1). By focusing on how concurrent changes affect the health of individuals within a population, a transition becomes a change in relationships between people and disease that occurs in a given time and space.

Description

The epidemiological transition is a descriptive and predictive model of change in health and disease patterns in

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

human populations. It is formulated in five propositions embodied in: a three-stage three-model in its first formulation (2), a three-stage four-model in the update of its first formulation (10), and a five-stage six-model in its ultimate formulation (11) by the time of Omran's death in 1999.

The first proposition in all formulations of the epidemiological transition, like the demographic transition for which mortality decline and fertility decline are the two key elements of population change (1), relates to the relative role of mortality and fertility in the epidemiological transition: mortality is a central force within the complex dynamics of this transition, notably during its early phases where fertility decline in Western populations was gradual and occurred 50–75 years after mortality declined (2, 10, 11).

The second proposition in its first formulation states that 'during the transition, a long-term shift occurs in mortality and disease patterns whereby pandemics of infection are gradually displaced by degenerative and man-made diseases as the chief form of morbidity and primary cause of death' (2: 516). This second proposition is unquestionably at the heart of Omran's concept of the epidemiological transition, and has served as the framework of reference in most empirical studies and reviews. In keeping with the stage-wise approach inherent in the demographic transition, Omran's ultimate formulation of the epidemiological transition (11) entails five stages, in response to critiques from several authors (12–21, 23–49, 88–91) who have proposed extensions of the epidemiological transition stages, by proposing a fourth stage called age of delayed degenerative diseases or hybristic stage (88, 89) and a fifth stage (of emerging infectious diseases) (90, 91). There is also some analogy here that can be made with the literature on the demographic transition which has recently suggested additional stages of the standard demographic transition, the 'second demographic transition' (92) and the 'third demographic transition' (93).

Omran's latest formulation of the second proposition encompasses a set of distinct transitions: the changing patterns of disease and health (he termed 'health transition'), the changing fertility and population age structure leading to ageing (he labeled 'partially the demographic transition'), the changing lifestyles (he called 'the lifestyle transition'), the changing health care patterns (he designated 'the health care transition'), the medical and technological evolutions (he called 'the technologic transition'), and the environmental and ecological changes (he termed 'the ecological transition') (11). Omran's earlier formulations of the epidemiological transition (2, 10) posited three evolutionary stages characteristically featuring all-cause mortality patterns and fertility patterns in the course of population change during the transition, and presumed that all societies were to go through these stages as did the now-developed countries of Western Europe and North America, just as conjectured in the demographic transition (1). Omran's ultimate formulation is modified to contrast the five successive stages in western societies with the three successive stages in non-western societies (11). By drawing such distinction between western and non-western societies, Omran has attempted to address some of the various assessments of the epidemiological transition since its inception (12–22, 43–50, 88–91). The distinction is based on mortality and fertility patterns, cause of death structure, health care, and socio-economic contexts. Mortality decline in non-western countries did not occur until the middle decades of the 20th century when the speed of such decline was unprecedented in comparison to the previous experience in the West. Mortality by cause of death in non-western countries was projected to shift in predominance from communicable diseases dropping from being responsible for 42.1% of all deaths in the less developed regions in 1970 to 19.4% in 2015, while neoplasm and circulatory disorders will rise from accounting for 21.6% of all deaths in 1970 to 48.9% in 2015, in tandem with life expectancy at birth which will rise from 57.5 to 68.5 years (11).

The first stage (age of pestilence and famine), is characterized by high and fluctuating mortality rates, variable life expectancy with low average life span of 20–40 years, and periods of population growth that are not sustained. The disease patterns that emerged are determined by increasing microbial exposures, dietary deficiencies, illnesses due to inadequate food storage, communicable and endemic diseases over wide geographic regions and increased mortality. In western societies and depending on the country, it entails the period from the Stone Age (or the extension of medieval or pre-industrial disease patterns) to the late 18th or early 19th centuries and was characterized by high and fluctuating mortality just as fertility was high, leading to slow and cyclic population growth. The average life expectancy at birth was short and variably swinging between 20 and 30 years and life expectancy for males was higher than or equal that of males, infant mortality exceeded 200 deaths per 1,000 live births, MMR was high (exceeding 1,000 per 100,000 live births), and infectious diseases, malnutrition and maternal complications claimed up to 75% of deaths while heart diseases and cancer claimed 6% of deaths. In non-western societies, the first stage is the longest length spanning from the medieval times to the middle decades of the 20th century: mortality is extremely high and fluctuating with peaks during the years of epidemics, famine, crop failures and wars, thus precluding sustained population growth; infant mortality exceeds 200–250 deaths per 1,000 live births, MMR exceeds 1,000–1,200 deaths per 100,000 live births, and life expectancy is short and oscillates between 20 and 35 years; disease pattern features a distinct predominance of communicable, maternal, perinatal and nutritional diseases; and fertility

is extremely high (7–10 births per woman) and childbearing starts at a young age and runs through the entire reproductive span.

The second stage (age of receding pandemics) is marked by a steeper decline in mortality as epidemics occur less frequently, and an increase in average life expectancy from around 40 to 55 years of age. This transition phase is characterized by a shift in patterns of disease and mortality from primarily infectious diseases to chronic and degenerative diseases. This shift is supposed to be accompanied by a shift in the population age distribution as early infectious disease deaths decline and deaths from chronic and degenerative disease increase. In western societies, this stage covered the period from the mid-18th century to approximately 1914. It saw the disappearance of the major acute infections as causes of death. The average life expectancy at birth increases steadily from about 30 to 50 years while infant mortality declines gradually to below 200 per 1,000 live births. Fertility remains at high levels and the widening gap between birth and death rates results in rapid population growth as well as a young age structure. During the latter part of this stage and usually after a time lag of 50 or more years from the onset of mortality decline, fertility decline takes place. In non-western societies, this second stage ended in the 1960s through to the 1990s and its beginning was delayed in some non-western countries until the 1940s and 1950s, when the onset of mortality decline occurred with the start of the recession of epidemics.

The third stage (age of degenerative, stress and man-made diseases in western countries; or age of triple health burden in non-western countries) features infectious disease pandemics that are replaced as major causes of death by degenerative diseases, with infectious agents as the major contributor to morbidity and mortality overtaken by anthropogenic causes. With declines in mortality rates, average life expectancy increases to exceed 70 years, fertility becomes more important to population growth, and the anthropogenic and biologic determinants of disease also change. In western societies, it ran from the second half of the 19th century to the late 1960s. This stage is a manifestation of the increasing prevalence of man-made diseases (e.g. radiation injury, occupational and health care hazards, chemical and biological warfare, environmental pollutants, motor vehicle and aviation accidents, carcinogens in the environment and in industry and food additives), stress-related diseases (e.g. depression and other mental illness, violence and drug dependency) as well as heart diseases, cerebrovascular accidents (strokes), cancer at various sites, diabetes, chronic obstructive pulmonary disease and metabolic disorders. These diseases have been progressively replacing infectious diseases and gross malnutrition which continue to be among the leading causes of morbidity and mortality without being the top causes of death. Mortality continues to decline and the average life expectancy at birth rises gradually from about 50 to 75 years or more, triggering population aging. It is during this stage that fertility becomes a crucial factor in population growth. In non-western societies, this third stage is termed the age of triple health burden: 1) unfinished old health problems including communicable diseases, perinatal and maternal morbidity and mortality, malnutrition, poor sanitation, rampant poverty, low literacy, overpopulation, limited access to health care and clean water; 2) rising new set of health problems including degenerative diseases (heart diseases, stroke, cancer and metabolic disorders), stress and depression, and man-made diseases as in western countries; and 3) lagging or ill-prepared health systems and medical training for the triple demands of quality services consisting of dealing with acute diseases generally resulting in short-term care, initiating prevention and care for chronic and non-communicable diseases (NCDs) usually requiring long-term medical or rehabilitation care, and handling problems of ageing. It started in the 1970s or later and its takeoff varies across countries given its transition model. During this stage, fluctuations more or less disappear and mortality continues to decline and eventually approaches stability at a relatively low level. The average life expectancy at birth rises gradually until it exceeds 70 years.

Omran (11:107) did not consider additional stages of the epidemiological transition for non-western countries, arguing that 'it is quite unlikely that these countries will enter the fourth stage in a manner similar to the one experienced by western countries'. Omran (11) noted six distinctive features of non-western transition. First, throughout the non-western countries, poverty, limited education, low status of women and slow pace of development have been major obstacles to well-timed and successful takeoff. Second, the timing of transition is crucial for the differentiation of the non-western transition. Third, since mortality decline was delayed until well into the 20th century, it was forcibly influenced by the availability of new health discoveries such as chemotherapy, antibiotics, insecticides, food quality and sanitary measures. Fourth, aging in non-western societies was delayed, but increased at a relatively fast rate in the later decades of the 20th century. Fifth, fertility decline in non-western countries was harder to initiate and required organized family planning campaigns to promote smaller family size norms and rise the age at marriage. Finally, the overlap of stages was much wider in the non-western transitions which Omran (11) classified further into three transition models.

The fourth stage is the age of declining cardiovascular mortality, aging, lifestyles modification, emerging and resurgent diseases. This new stage draws on later contributions (88–91). It begins with a plateau in epidemiological history where mortality reached at once equilibrium with

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

a life expectancy at birth in the 1970s at a value which was then believed to be close to the biological limit to the average length of human life. Olshansky and Ault (88) noted that a few years prior to Omran's seminal publication (2), the United States and other Western countries began to experience a rapid decline in cardiovascular mortality which occurred around 1970 in many developed countries; they proposed this cardiovascular revolution as a fourth stage of the transition labeled 'age of delayed degenerative diseases'. Characteristically, the age pattern of mortality by cause of death is unaltered from the third stage to this fourth stage, but the age distribution of deaths from degenerative causes shifts progressively to advanced ages with associated increases in the size of the population at advanced ages and on the health of the elderly especially the oldest old. There are further increases in life expectancy approaching 80 to 85 years or longer notably for women, increased disease chronicity, and aging accompanied by rising medical costs for individuals and the state.

The fifth stage is the age of aspired quality of life, equity, development and social justice for all. This stage is a futuristic stage with paradoxical longevity, chronic morbidity, disabling impairments, isolation and decline in social status, loss of independence, new morbidity, mounting medical, long-term care and nursing costs, persistent inequities, inadequacies and disparities between people, communities and countries because of the polarization of socioeconomic status within and between countries. At this stage of the epidemiological transition, much increased chronicity will be a common feature of degenerative, stress and man-made diseases which will be higher than the causes of morbidity, disability and mortality. Omran (11) conjectured that tobacco will be the most villainous disease risk factor in this stage.

Explanation

The following three propositions explain the differentials in epidemiological transition.

The third proposition is that there are age and sex differentials that come with the epidemiological transition. Regarding age differentials, the transition usually favors the young over the old: the most significant impact of the transition is the fall of infant and child mortality and maternal mortality and the rise of mortality at 50 years or older (10). Childhood survival gradually improved as pandemics receded in response to better living standards, improved nutrition and sanitation and modern public health measures in Europe. This was also apparent for the United States between 1880 and 1970, and for England, Wales and Chile (2, 10).

The fourth proposition is that the shifts in health and disease patterns that characterized the epidemiological transition prior to the 20th century have a closer association with rising standards of living and improved nutrition than with medical progress. In contrast, the 20th century transitions prevailing in the less developed countries are initiated by medical progress, organized health care, and disease control programs that are usually internationally assisted and financed, and thus largely independent of the socioeconomic level of the country. Further maturation of the transition, however, depends on a beneficial synergy of health care progress and socioeconomic development. Likewise, the fertility transition prior to the 20th century was largely determined by social and economic progress in countries that were more or less familiar with traditional methods of fertility control. These methods, as well as rising age at first marriage, were the major intermediate variables for fertility reduction. The more recent fertility decline in some developing countries, on the other hand, has been dependent on organized efforts of family planning in conjunction with socio-economic development.

The fifth proposition is that distinctive variations in the pattern, pace, determinants, and consequences of health, survival and population changes differentiate six transitional models of the epidemiological transition: the classical or Western transitional model (e.g. Western Europe, North America, Japan, and Australia), the semi-western/accelerated model (e.g. Argentina, Paraguay), the non-western rapid transition model (e.g. Chile, Costa Rica, Cuba, Jamaica, Martinique), the non-western upper intermediate transition model (e.g. Brazil, Mexico), the non-western lower intermediate transition model (e.g. Paraguay, Peru), and the non-western slow transition model (e.g. Bolivia, Haiti, Nicaragua) (2, 10, 11).

The western transition model describes the transition in western societies which have entered the fourth stage of the transition during the past three centuries. It is characterized by the shift of birth rates from 30 to 35 per 1,000 population to less than 15 per 1,000 population, the shift of death rates from 30 to 35 per 1,000 population to less than 10 per 1,000 population, the shift of short life expectancy of 30–40 years to unprecedented long life expectancy of 80–90 years and more. Mortality declined gradually and was due primarily to social, economic and environmental improvements, better nutrition and personal health habits. During the transition, pandemics and major epidemics receded and were gradually but not totally replaced by degenerative, stress-related and man-made diseases; these changes in disease patterns were associated with improvement in life expectancy notably for children, young adults, and females of reproductive age. Fertility decline in this Western model was also gradual and occurred 50–75 years after mortality declined, but this sequence of fertility decline following mortality decline was not universal because in France and some European parishes, fertility and mortality declines almost coincided.

The semi-western/accelerated model describes the experience in eastern Europe, the former USSR and Japan, and in European populations living outside Europe, North America or Australia at various historical epochs (e.g. Europeans living in Argentina, Israel, South Africa). In most of these countries, a slow process of modernization had begun prior to the drop in mortality in the 20th century, which was determined by general social improvements as well as by sanitary and medical advances. Wide use of abortion typically helped accelerate the fertility transition, like in Japan (10). By and large, countries in this model have not entered the fourth stage of the epidemiological transition, many of these countries exhibit a continuing rise in cardiovascular mortality, and some members of the former USSR have experienced lost years of life expectancy due to social and economic crises.

The other transition models apply to non-western countries, describing epidemiological changes over time in less developed countries where mortality decline took place no earlier than the 1930s through the 1950s and fertility decline was delayed further until after 1950. The sharp decline in mortality in tandem with high fertility led to unprecedented rates of population growth during this period, and Omran (11) distinguished these countries not only in terms of differences in the post-war patterns of mortality and life expectancy, but mostly given the speed, timing and magnitude of change in fertility decline.

The non-western rapid transition model describes the experience of rapidly industrializing or socially developing countries and territorial islands (e.g. Taiwan, Hong Kong, Singapore, South Korea, Sri Lanka, Mauritius, China, Chile, Cuba, Barbados, Costa Rica, and the French overseas departments such as Reunion). The onset of mortality decline to moderate levels in these countries was one or more decades before 1950 while fertility decline to less than five children per woman was delayed until sometime in the 1960s. These countries are still in the third stage of the epidemiological transition, but have been experienced also the triple disease burden.

The non-western (upper and lower) intermediate transition model embodies the experience middle- or low-income countries whose changes in mortality and disease patterns and fertility patterns are situated between the rapid and slow transition models. Countries in the upper non-western intermediate transition model include Indonesia, Colombia, Mexico, Brazil, Panama, Venezuela, Tunisia, Lebanon, and Thailand. Countries in the lower non-western intermediate transition model comprise India, Egypt, Morocco, Ecuador, Peru, Paraguay, and the Dominican Republic. These countries still face a dual burden of continuing communicable diseases and malnutrition and the rising prevalence of NCDs, often in combination with the emerging diseases such as HIV/AIDS or the resurgent diseases such as malaria.

The non-western slow transition model describes the experience of the least developed countries and some of the less developed countries in Latin America, Asia and Africa. In these countries, mortality started declining to moderate levels after 1950 while fertility remained at high levels until the 1990s. These countries also face the dual burden of the continuing preponderance of communicable diseases and malnutrition and of the rise in degenerative and man-made diseases, along with HIV/AIDS, malaria, tuberculosis, and other emerging or resurgent diseases; they also have been ill-prepared to handle any one health problem.

Health transition

Definition

The health transition refers to the various components in combination in a series of concurrent changes in population health during the development, and their implications for health and social policies and programs. For instance, the functional component represents change in functional health status (i.e. abilities and disabilities) of the population and the gerontological component denotes the increasing proportion of people at old and very old ages with their connected health problems. The health transition is the latest conceptual effort made by social and public health scientists to describe and explain secular transitions in population and health dynamics. However, as we show below, its conceptual clarity and contributions, empirical foundations and prognostic implications both regarding historical and contemporary populations, remain typically blurred and unmapped.

Description

Several appraisals of the epidemiological transition have suggested revisions to its earlier formulations (12, 13, 23, 26–32) or have rejected it altogether as not relevant to health and disease patterns of the low- and middle-income countries (18, 19–22). Undeniably the most successful attempt of such revisionism is the concept of the 'health transition' meant to be a wider framework that included both population change, health system, epidemiological profile of the population and ways in which societies respond (or not) to changing health situations as a result of cultural, social, and behavioral determinants (12–21). The 'health transition' gained considerable traction in the 1990s. While the health transition perspective seeks to provide some causation to the explanation of the secular changes in health and disease patterns in human populations, health transition represents different conceptual entities depending on the school of thought which have made progress towards a coherently crafted framework beyond Omran's epidemiological transition a daunting endeavor. The Caldwell's school of thought views health transition as rooted in the cultural, social and behavioral determinants of health (18–21, 49), while scholars associated with the Frenk's

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

school of thought (12, 13, 23, 26–29) view it as a revision or an extension of the Omran's epidemiological transition (2, 10).

Caldwell's perspective on health transition (19) departs from Omran's epidemiological transition (2, 10, 11) and Frenk's elements of health transition (12, 13). It builds instead on a paper by Simons (94) published in a volume edited by Caldwell and Santow (14) around the same time as the work of Frenk and colleagues (12). Caldwell's school of thought considers the health transition and improvement in health in circumstances of poor economic growth typical of developing countries such as those of SSA. What Caldwell and Santow (14) refer to as 'the health transition factor' is very influential in this perspective: societies with similar levels of income and provision of health services can exhibit very different levels of health and mortality. This can also apply to different cultures, families or households even within the same societies, indicating a behavioral or attitudinal effect at work that has to do with lifestyles. The main argument embodied in Caldwell's view of the health transition is that the driving forces of population and health changes in different epochs are cultural, social, and behavioral in nature in both western and non-western societies (18, 19, 20, 21). The evidence from changes in non-western societies suggests that a considerable part of the gap in causation is explained by individual and social change. At the individual level, new attitudes and behavioral patterns took advantage of socio-economic opportunity structures and medical advances to improve survival chances. First, death is abhorrent and must be avoided at all costs, and there must be an absolute commitment to ensuring the survival of all members of the community. Second, there must be a feeling in all individuals of a sense of active individual commitment to participating in ensuring survival of each other by avoiding death through the reduction of the risk of harm or death, rather than believing that external forces, resources, physical or moral entities will do it for individuals. The most powerful instrument that has been shown to accelerate those aspects of social change that enhance health improvement and survival has been found to be raising individual consciousness and sense of responsibility, particularly through formal education in general and female education notably (14, 15, 19). The routes to low mortality in such resource-disadvantaged settings is the establishment of a democratic health system, by implementing a network of functional health facilities that are reachable by the local populations with emphasis on accessibility (free or cheap) of the poor of rural and urban areas, as well as universal health insurance coverage of some form to them. Perhaps Caldwell's perspective on health transition is best styled in his following words:

> ... During the long contact with the important institution now known as the International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B), I have been astonished at the lack of interest in identifying necessary behavioural changes, and trying to supplement them. Most of the ICDDR, B's resources in the diarrhoeal area have been employed to identify diarrhoeal pathogens, and to search for methods of immunization... Except in the area of inducing parents to make and employ oral rehydration solutions, however, practically nothing on studying, or inducing individual or societal hygienic behavioural change has ever been attempted. The focus has always been on pathogens and therapies rather than people. (19: 125).

Frenk's views of the health transition were articulated in two publications in the late 1980s and early 1990s (12, 13). Frenk et al. (13) propose that the concept of epidemiological transition be replaced by the wider concept of health transition, which would include not only the development of epidemiological characteristics within the overall health situation, but also the ways in which societies respond to the health situation and vice versa. This perspective considers the health transition in the context of middle-income countries facing new challenges for health care, in order to account for response of the organized health system to long-term changes in the health conditions of populations. More recently, Vallin and Meslé (29) have attempted to broaden the scope of this concept to represent situations of convergence and divergence in mortality both in developed and developing countries. Frenk et al. (12) propose some corrections and additions to the original three-stage-three-model version of the epidemiological transition (2). In criticizing Omran's formulation, Frenk et al. (12) argue that between the countries in the advanced stage of the epidemiological transition and those in the initial stage lies a third group that is undergoing a new transition experience quite different from that of the developed nations. They describe such countries as belonging to the 'protracted and polarized' transition model. This model has four characteristic features. First, the transition stages are overlapping: the stages of the transition do not follow a sequential order but exhibit considerable overlap. Second, there are counter-transitions: the shift from high mortality to high morbidity occurs not only for degenerative diseases but also for infectious diseases as in some developing countries (e.g. case for HIV/AIDS where morbidity rates are increasingly higher than mortality rates as treatment become available to a greater proportion of the population). Third, the transition is protracted: the transition process is not clearly resolved as countries exhibit both infectious and chronic diseases. Fourth, there is an epidemiological polarization: a polarization is seen to occur between population subgroups with the poor and rural

populations succumbing to the pre-transitional pathology while the urban populations experience post-transitional pathology. Frenk and colleagues (12, 13) illustrate this perspective with evidence from the epidemiological transition in Mexico quite different from the one experienced in western and North American countries: in the 1950s, the 10 main causes of death already included many chronic and degenerative diseases such as heart conditions, tumors and cerebrovascular problems; infectious diseases caused more than 30% of all deaths. This is seen as the first description of the protracted epidemiological transition. The epidemiological polarization in Mexico is largely explained by the different pace of change in the causes of ill-health among the social classes.

Vallin and Meslé (29) suggest that health transition be characterized as an extension of the epidemiological transition. Within the concept of health transition tied to the healthcare system as proposed by Frenk (12, 13), Vallin and Meslé (29) consider the full process in term of divergence/convergence sequences inferred by successive major changes in health technologies and strategies. The entire health transition process thus breaks down into successive stages, each including a specific divergence-convergence sub-process. The first stage of the health transition is Omran's three-stage epidemiological transition in its classic formulation. The second stage of the health transition is the cardiovascular revolution. There is now overwhelming evidence indicating that Omran's third age of the epidemiological transition, the age of degenerative and man-made diseases (2, 10), was not the final one as also acknowledged by Omran (11). The third stage is the 'slowing the ageing process'. This stage relates to the most recent trends in life expectancy for Western industrialized countries. The conjecture hinges on the view that the new approaches toward the elderly developed since the 1980s may be creating a new stage of health transition, which a handful of countries may have already experienced. The impact of the decline in cardiovascular mortality mainly comes from the oldest age groups. In light of this increasing role of the oldest ages in the decline of cardiovascular mortality, increasing advances in the fight against these diseases are tied up with more general progress against ageing. The current cause-of-death classification is certainly not sufficiently ageing-related to allow a pinpointing of the impact of current strategies on such ageing process.

Measuring health and describing the course of health transition

Given the intricacies of measuring individual and population health over time and across cultures, Johanssson (95) recently proposed three strategies for measuring health in an attempt to reach conclusions about the course of the health transition and how such concept may be related to historical and contemporary trends in mortality. The first strategy is to continue the use of mortality data to assess health trends, as most researchers have been doing (12, 13, 23–29, 33, 34), assuming that rises in life expectancy signify improvements in health. This strategy is tantamount to suggesting that rising life expectancy is all that matters, so that comparing life expectancy levels between countries during their economic development can measure health inequalities. The second strategy rests on a medicalized life course perspective on health and entails using data that are relatively culture-free (e.g. blood pressure) to objectively determine individual health by measuring its diminishing from healthy state through sickness and disability; hence, the health transition becomes a 'doctor's transition', based on comparable age cohorts at different points in time during the rise of life expectancy. The third strategy is to measure the health transition in all its complexity in its multifaceted, multilevel and life-course dimensions relevant to health policy. These second and third strategies are not readily feasible and require a formidable amount of cooperation across disciplines, large-scale resources and skills which are not commonplace.

Comparing and contrasting the demographic, epidemiological and health transitions

Temporal coincidences between the demographic, epidemiological and health transitions

Four distinct historical periods have marked the mortality transition in the modern world described in the demographic transition (96) and which were used in subsequent efforts to describe and explain secular transitions in the epidemiological and health transitions. The first period runs through the first half of the 19th century. In Western Europe, mortality reduction is most clearly identified in the latter part of the 18th and first half of the 19th century. The second period covers the last third of the nineteenth century up to World War I. During this period, there was a revolution in medicine induced notably by the discoveries of Pasteur and Koch. The resulting reductions in child mortality and subsequently in infant mortality were responsible for much of the mortality decline particularly in mortality from diseases such as diarrhea and tuberculosis. During the inter-war period, sustained gains were achieved especially in medicine and health education, instilled by progress made during and after World War I. The third period involves the years during World War II and the following period through the 1960s. During this period, there was an explosion in the use of antibiotics, initiated by Fleming's discovery of penicillin and its synthesis in 1943. The cumulative effect of these developments has been a dramatic reduction in epidemic and communicable diseases. The fourth period covers the period since the 1970s when important gains have been achieved in reducing

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

mortality from cardiovascular diseases, and particularly in increasing the longevity of older adults. The adoption of modern lifestyles has not been responsible for increased mortality from degenerative diseases until recently, because of the high incidence of communicable diseases.

Similarities

A common feature of these three transition frameworks of investigation of changes involves a progression through a sequence of three stages in their first formulations. The idea of convergence is a general basis of these frameworks, not only for mortality and life expectancy trends but also for fertility (97). This notion of progression through the stages of transition is the subject of a substantial and ongoing debate; such stage-wise linear approach to change is a major shortcoming of these concepts. To the extent that mortality data track the health status of a population at all ages, the mortality transition may proxy any separate health transition (96), provided that more specialized measures of mortality are produced to highlight the importance of various components of concurrent changes in the production of health during development.

Frank Notestein's formulation of the demographic transition (1) is by and large the most influential framework since the 1950s. The epidemiological and health transition are tied to the demographic transition in at least two respects (the notions of transition stages and mortality transition), but also represent distinct constructs in their own right. The Notestein's approach to mortality analysis is most fully elaborated in Omran's epidemiology of the population change (2, 10, 11) recognizing that mortality transition involves more than simple quantitative reductions in mortality levels and their short-term fluctuations. Similarly, the concept of health transition is closely related to sustained progress at reducing mortality among infants and adults and to improvements in life expectancy by changing the patterns of disease and cause of death at older ages in the context of well-equipped and functioning healthcare system. The mortality transition refers to the secular decline in mortality and its associated changes in mortality patterns of age and cause of death. The epidemiological transition includes the mortality transition and additionally embraces a broader set of changes in morbidity and the disease environment. The health transition in Caldwell's sense includes the social and behavioral changes that parallel the epidemiological transition and may do much to propel it (16). The concept of health transition proposed by Frenk et al. (12, 13) is tied to the healthcare system.

Differences

The main aspect that separates the epidemiological transition (2, 10, 11) from the demographic transition (1) is the addition of a new element – a shift in cause-of-death patterns and the stage-wise characterization of the transition stages by the configurations of the causes of death as well as the influences on them. Despite his reliance on its main concepts, Omran explicitly rejected the demographic transition as a theoretical framework, and formulated the epidemiological transition in an attempt to provide a comprehensive approach to the dynamics of the mortality–fertility transition (4). In his view, the demographic transition ignored too many variables and too many historical developments that did not fit the theory (e.g. the postwar baby boom), as well as assuming a single path of population development based on socioeconomic progress. Omran consistently emphasized the importance of motivation rather than economic determinants in explaining why birth control was embraced enthusiastically in some places and rejected in others (2). He further believed that motivation could be promoted through health programs, directly by explaining the health benefits of family planning and indirectly by decreasing infant mortality and thus reducing the need for large families. For Omran (2, 10, 11), the key difference between epidemiological transition and demographic transition is that the former unlike the latter allows for multiple pathways to a low-mortality/low-fertility population regime.

A feature of the concept of the health transition that situates it apart from the demographic and the epidemiological transitions is the idea that there is more to health than death, especially at older ages around the world. For instance, it is known that the health of the elderly can deteriorate even if life expectancy is stable or rising at older ages. Continuing improvements in high life expectancies have made it imperious to develop measures of population health such as active life expectancy (ALE), disability-adjusted life expectancy (DALY) and quality-adjusted life years (QALY) (98) which are apt to track the health status of a population without relying on mortality data alone. However, since health research has generally relied on quantitative data and analysis, Johansson (95) anticipates that researchers will continue to use mortality data as a proxy for health trends and to treat the health transition just as another name for the mortality transition.

Changing contexts of health, disease and mortality patterns in Africa

Brief history and experience of Africa

Information on the situation in Africa before the independence years in the 1960s is sketchy and at best conjectural. Kuczynski's study on the Cameroons and Togoland, published in 1939, was the first critical investigation of information on the population of Africa and was largely negative concerning the possibility of

substantial knowledge on the basis of the information then available (99). By the 1960s, conditions were relatively favorable (100), but by the end of the 1970s there were clear signs that the post-colonial state was not only falling short of its ambitious designs, but facing a systematic crisis, and Africa was unable to point to any significant growth rate or satisfactory index of general well-being (101).

The 1980s were a decade of growing distress for much of Africa, and since the 1980s, while other parts of the world have significantly prospered, African economies have experienced numerous disruptions. For sub-Saharan countries, there was an average negative growth rate of 2.8% from 1980 to 1987 (102). All African countries but Botswana and Mauritius were overwhelmed by unsustainable debts, facing negative trends in many primary commodity markets. Shunned by much of international capital, African states had little option but to accept, at least formally, the structural adjustment programs proposed by the international financial institutions backed by the donor community as the condition for continuing assistance in the late 1980s and beyond (103). The budget reductions required by structural adjustment programs compelled cutbacks in social expenditures. These trends substantially weakened states for which studies already showed that educational and health expenditures had less success in improving health and educational attainment than comparable outlays in Asia and Latin America (104).

While at the beginning of the 1980s the external diagnosis of an emerging African crisis was essentially economic, by the end of the decade the exegesis of the post-colonial state condition led to a more far-reaching conclusion: the malady afflicting Africa was not simply economic but more fundamentally political and required democratic reforms. The kinds of political and economic arrangements that characterized the post-colonial state had become fundamentally unviable in a changing global environment.

The 1990s and beyond represented a fundamentally new era in African politics, with the rise of multiparty systems. A vision of a new type of Africa state was offered by the New Partnership for Africa's Development (NEPAD), which pledged an economically reformed, politically democratic, and governmentally accountable and transparent Africa. However, the resulting partially reformed states proved to be substantially weakened, and new patterns of civil conflict and internal warfare appeared in a number of countries. Governments were overthrown by peripheral insurgents eight times in the 1990s: Chad (1990), Liberia (1990), Ethiopia (1991), Somalia (1991), Rwanda (1994), and Sierra Leone and both Congo and the Democratic Republic of Congo (1997). The spread of civil conflict in Africa in the 1990s stood as a metaphor for weakened and unstable nations. More than a quarter of African states experienced armed internal conflict during the decade, and another quarter faced prolonged political crises and turbulence (105), some of which are on-going.

Conflict and instability

Over the last 50–60 years, African countries have experienced a variety of social, economic, and political crises that have affected population health and human longevity.

Politically, the 1950s and 1960s signaled the beginning of a new era for most of Africa (106–108), as the progress towards an orderly and peaceful transition of power from the colonial powers to the indigenous people began. Today, Africa contains 57 sovereign countries, most of which still have the borders drawn during the era of European colonialism. During the period from the early 1960s to the late 1980s, Africa had more than 70 coups and 13 presidential assassinations (109). Overall, 34 out of 52 African countries have experienced political instability (PI) as measured in 1996 or 2009, and 23 countries have witnessed a deteriorating political situation between 1996 and 2009. Since the 1960s, several African countries have experienced political coups, ethnic violence, oppressive dictators, and the region has endured many conflicts leading to humanitarian crises and mass migration of refugees.

The most devastating military conflict in modern independent Africa has been the Second Congo War, which by 2008 has claimed the lives of an estimated 5.4 million people. Since the 1960s, many western African countries have also been submerged under PI, with notable civil wars in Nigeria, Sierra Leone, Liberia, and Côte d'Ivoire, and a succession of military coups in Ghana and Burkina Faso. Southern Africa is distinct from the rest of Africa in that it has been exempt from wars.

Despite democratization trends since 1990, PI has been widespread in Africa. From 1956 to 2001, there have been 80 successful coups d'état, 108 failed coups d'état, and 139 reputed coup complots in 48 sub-Saharan African independent states (110). Elite PI remains widespread in SSA, in contrast to other regions of the developing world. Military-led PI has been shown to have deleterious effects on economic growth and human development in SSA and is a major cause of the enduring food, economic, and health crises in many African countries. This African context has huge implications on the capacity of the continent to set a path of continuous and sustainable social environment for human health development; unless Africa and Africans consciously change the course, their future may resemble their past.

Disease

Although countries in Africa share certain similarities, such as conflict-prone environments, they differ greatly

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

in many respects, including demography, epidemiology, culture, history, politics, and education. As such, it is not surprising to find that HIV epidemics across Africa range from highly concentrated to highly generalized, and the responses by African nations have been equally varied, ranging from intense interventions to complete denial of HIV disease.

Saying that HIV/AIDS has had a large impact in SSA is a gross understatement. Sixty-eight percent of people living with HIV worldwide live on the African continent, where every year 76% of all AIDS-related deaths in the world occur. In 2007 Africa accounted for 68% of new HIV infections in the world. At the same time, trends in the prevalence of HIV among adults aged 15–49 years shows great heterogeneity across countries. In 1990 over 10% of the adult population was HIV-positive in three countries (Zambia, Zimbabwe and Uganda); by 2000 the prevalence had risen to 24.8% in Zimbabwe and 14.4% in Zambia but had declined to 7.3% in Uganda. Botswana (26% and 24.8% in 2000 and 2009, respectively) and Lesotho (24.5 and 23.6% in 2000 and 2009, respectively) are the two countries with the speediest increase in HIV prevalence, having grown from 0.8% in Lesotho and 3.5% in Botswana in 1990.

Thanks to better HIV/AIDS treatments using antiretroviral therapies (ART), morbidity and mortality have decreased in HIV patients with advanced disease. However, as the number of patients on antiretrovirals increases, more and more people are living longer with HIV, raising new challenges. Furthermore, the majority of African countries have ART coverage under 45%, meaning that less than half of Africans who need antiretroviral therapy actually receive appropriate care, and current treatment guidelines often delay initiation of antiretroviral therapy during the early stages of disease.

The future of the epidemiological landscape of African countries will greatly depend on strategies to build capacity for prevention, treatment, and care of HIV/AIDS in Africa. If African states are to assume greater responsibility for responding to the HIV/AIDS epidemics, a major requirement will be to strengthen health-care systems in the region by building institutional and human resource capacity (111, 112).

Between 1990 and 2010, the prevalence of tuberculosis has increased in 25 African countries, and Africa is facing the worst tuberculosis epidemic since the advent of the antibiotic era (113). Driven by the HIV epidemic and compounded by weak health-care systems, inadequate laboratories, and conditions that promote transmission of infection, this devastating situation has steadily worsened, exacerbated by the emergence of drug-resistant strains of tuberculosis. The WHO estimates that the average incidence of tuberculosis in African countries more than doubled between 1990 and 2005, from 149 to 343 per 100,000 population – a stark contrast to the stable or declining rates in all other regions during this period (114). Although HIV is Africa's leading cause of death, tuberculosis is the most common coexisting condition in people who die from AIDS.

Economic factors

In 2009, gross domestic product (GDP) per capita – a proxy for economic development – ranged from 97 in the Democratic Republic of Congo to $8,011 in Equatorial Guinea. Although it has abundant natural resources, Africa remains the world's poorest and most underdeveloped continent. The informal sector is the largest contributor to GDP in the regions where agriculture is predominant, such as in SSA. Agriculture has always been and remains the dominant activity for the majority of Africans. Most agricultural activity is subsistence farming, which has made agricultural activity vulnerable to climate change and global warming. SSA is the region of the world with the largest estimates for the contribution of informal sector to GDP: nearly two-thirds including agriculture and one-third excluding agriculture.

Although over 50% of the adult population is employed in all African countries, labor force participation is dominated by the informal employment (115–119). The informal sector represents 80.4% of total employment in SSA, where 88.6% of women are more likely to be self-employed in the informal sector (118). This affects health because informal employment is usually characterized by an absence of social protection, in particular, health coverage (115, 118), which leaves large majorities of the continent's populations especially vulnerable to disease and other health issues.

Health-care systems

In the absence of vaccination, measles is estimated to infect virtually the entire population with the exception of isolated communities (120), and measles vaccination rates are sensitive indicators of functional public health systems (121, 122). The joint WHO/UNICEF (123) estimates indicate that measles-containing vaccine (MCV) coverage rates were lowest for the African region.

Vaccine rates indicate a great diversity in the functioning of the public health systems across African countries. In 2010 estimated coverage of MCV rates ranged from 46 to 94%. There were 12 African countries with MCV coverage of 65% or lower, indicative of the least functional public health system, while coverage was at least 90% in seven countries. The regional and local disparities in coverage result from limited resources, competing health priorities, poor management of health systems, and inadequate monitoring and supervision.

Many health systems in SSA currently lack the capacity to provide even basic health care to the population, let alone deal with the additional burden of comorbidities from communicable and NCDs. This is particularly obvious when examining the availability of physicians,

whose number per 100,000 inhabitants in 2004 was below the 1960 level in a number of countries (Tanzania, Zimbabwe, Liberia, Mozambique, and Sierra Leone). Elsewhere the number of physicians per 100,000 people stagnated over time or increased somewhat, but substantial increases between 1960 and 2004 were observed only in northern African countries and Mauritius.

The treatment gap in African countries has been mainly attributed to inadequately skilled personnel, cost of treatment, cultural beliefs, and unavailability of drugs, although lack of accessible health facilities has also been noted (50, 51, 91, 112, 114, 124–129). The lack of human resources and the difficulty retaining staff, especially in rural areas, is an important obstacle. A recent study found that alarming proportions of health care workers in African countries intend to migrate (129). The loss of health-care workers leaves crippling gaps in health-care systems of developing countries already struggling to combat mounting health crises, such as tuberculosis and AIDS (130, 131). Furthermore, while the bulk of social and health infrastructures and the best equipped hospitals are located in major cities in Africa, the African population is predominantly rural, with the urban population representing only 40 and 37%, respectively, of the total population in Africa and SSA in 2009.

Assessing the relevance of the transition frameworks in Africa

Testing the relevance of the demographic transition in Africa

To assess the extent of changes from 1950 to 2010 in mortality, fertility and population growth within the demographic transition perspective, we use two complementary approaches. First, graphical methods are used to assess trends in crude birth rates, crude death rates, infant and U5MR rates, total fertility rates, and life expectancy at birth from 1950 to 2010. We then evaluate whether the sequence of irreversible decline in fertility in response to mortality decline is occurring as expected from the demographic transition. Second, descriptive statistics are used to determine the ranges of variation (minimum, maximum) in estimates of crude birth rates, crude death rates, natural rates of population growth and total fertility rates. We then compare them to those expected when the demographic transition is/has taken place, as described above. These assessments are presented in Table 1.

Regional patterns of change in vital rates in Africa in comparative perspective between 1950 and 2010

One approach to understanding the changes that have taken place in African population over the past 60 years is to consider the differences in major demographic indicators between then and now, using the latest and best available estimates from the United Nations' online databases. Over the past six decades the population of Africa has witnessed a series of changes in population size and growth patterns due largely to variations in the two principal components of growth, that is, mortality and fertility. As during the demographic transition in Europe where fertility decline followed mortality decline in general, the rate of natural increase during the transitional period is expected to range between 0.5 and 1% per year in contemporary societies experiencing demographic transition (hypothesis 1). Trends in crude birth rates and crude death rates in Africa and across its regions between 1950–1955 and 2005–2010 are depicted in Fig. 1. Table 1 shows that the crude birth rates all remain very high and that with declining mortality though the levels are still very high, the range of variation in rates of natural increase has uniformly exceeded unity by a wide margin in all African regions and countries and over time (the one-time negative value for Rwanda is due to the mortality crisis during Rwandan genocide). Hence, we reject this hypothesis.

Between 1950–1955 and 2005–2010 crude birth rates and death rates both declined in Africa and its regions, although at a slower pace than in other regions of the world. Crude birth rates were higher in Northern Africa than SSA through the 1960s but have been higher in SSA than in Northern Africa ever since. Crude death rates, by contrast, have been consistently higher in SSA than in Northern Africa over time. Within SSA, differences across regions have varied over time; since the 1980s, crude birth rates and crude death rates have been highest in Middle Africa.

Over the past 50–60 years, major demographic changes have affected all regions and countries. The population growth rate in Africa rose from 2.11% in 1950–1955 (which was slower than that of Latin America, 2.72%, or of Oceania, 2.22%) to 2.30% in 2005–2010, which was the fastest rate in the world. Regional differences in population growth rate across Africa have been remarkable during the last 60 years. The growth rate declined from 2.34 to 1.74% in northern Africa, for example, but rose from 2.05 to 2.45% in SSA. Within SSA, southern Africa has experienced the sharpest decline in the rate of growth (from 2.30% in 1950–1955 to 1.02% in 2005–2010), while Middle Africa went from the lowest rate in 1950–1955 (1.93%) to the fastest rate in 2005–2010 (2.66%) in part due to a decline in infertility from controlling reproductive tract infections (132). Population growth has three major consequences for infectious disease: Growing populations in poor urban areas increase contact rates, facilitating epidemics; the sheer scale of cities provides more opportunities for a disease to persist with a ready supply of susceptible people; and the conditions associated with increases in travel and migration can turn epidemics into pandemics (133).

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

Table 1. Assessing the demographic, epidemiological and health transitions in Africa in comparative perspective: 1950 to 2010

Major area, African region and country	CBR range of variation* 1950–2010	CDR range of variation* 1950–2010	RNI range of variation* 1950–2010	TFR Range of variation* 1950–2010	Onset of sustained fertility rate since 1950	e(0) 1950–1955	e(0) 2005–2010	e(0) average yearly change 1950–1990 (a)	e(0) average yearly change 1990–2010 (b)
World	20.1; 37.0	8.1; 19.1	12.0; 20.6	2.5; 5.0	1955	46.91	68.72	0.43	0.19
More developed regions	11.1; 22.4	9.4; 10.6	0.8; 11.8	1.6; 2.8	<1950	64.67	76.90	0.23	0.12
Less developed regions	22.0; 43.7	7.7; 23.1	14.4; 25.6	2.7; 6.1	1955	41.62	66.96	0.50	0.21
Sub-Saharan Africa	39.6; 47.4	13.0; 27.7	19.7; 28.9	5.4; 6.8	1980	36.23	52.94	0.32	0.15
Africa	36.7; 48.1	11.8; 26.8	21.3; 28.8	4.9; 6.7	1975	37.38	55.55	0.36	0.15
Eastern Africa	39.4; 49.3	11.1; 27.0	22.3; 30.3	5.4; 7.1	1980	37.03	55.85	0.30	0.27
Burundi	42.5; 51.4	14.2; 25.2	24.2; 33.8	6.5; 7.6	1980	39.03	51.32	0.24	0.10
Comoros	36.8; 47.4	9.6; 23.6	21.4; 31.9	5.1; 7.1	1985	40.72	59.66	0.36	0.19
Djibouti	28.9; 44.9	9.7; 21.0	19.2; 30.8	3.8; 6.8	1975	41.04	59.05	0.38	0.12
Eritrea	39.4; 48.0	8.0; 27.8	19.3; 31.2	5.2; 7.0	1960	35.84	59.95	0.27	0.54
Ethiopia	36.4; 49.3	9.5; 29.9	19.3; 30.3	5.3; 7.4	1955	34.08	59.26	0.30	0.53
Kenya	38.2; 51.5	10.0; 23.6	27.7; 37.7	4.8; 8.1	1970	42.30	57.23	0.43	−0.09
Madagascar	36.3; 49.1	7.9; 27.7	21.3; 31.5	4.8; 7.3	1975	36.32	62.23	0.34	0.49
Malawi	41.8; 55.0	13.5; 27.8	19.7; 33.5	5.8; 7.6	1985	36.32	51.56	0.26	0.19
Mauritius	12.4; 43.8	5.9; 15.4	5.3; 35.1	1.6; 6.2	1965	50.20	72.76	0.46	0.17
Mayotte	33.7; 49.4	2.4; 26.9	14.3; 42.3	4.3; 7.9	1980	47.26	77.94	0.60	0.27
Mozambique	42.3; 49.4	15.8; 32.8	16.5; 28.0	5.6; 6.6	1975	31.29	48.35	0.29	0.21
Réunion	18.2; 49.6	5.1; 19.2	13.1; 31.7	2.4; 6.9	1965	47.86	78.20	0.60	0.26
Rwanda	38.1; 53.8	9.1; 25.0	−0.6; 36.3	5.1; 8.4	1985	40.02	59.84	0.14	0.56
Seychelles	18.8; 39.0	7.1; 15.5	11.5; 27.4	2.2; 5.9	1970	57.96	72.41	0.33	0.05
Somalia	45.6; 49.7	13.7; 29.8	19.1; 33.2	7.1; 7.4	1970	33.99	53.21	0.31	0.27
South Sudan	38.5; 51.4	13.6; 37.3	10.4; 27.3	5.4; 6.9	1980	27.88	52.12	0.35	0.41
Uganda	45.8; 51.3	11.4; 24.5	26.8; 34.4	6.4; 7.1	2000	40.00	55.17	0.22	0.26
Tanzania	41.6; 49.3	10.8; 22.4	26.6; 31.5	5.6; 6.8	1975	41.25	56.60	0.25	0.22
Zambia	43.4; 49.3	13.7; 21.8	24.9; 32.6	5.9; 7.4	1980	42.07	50.94	0.10	0.20
Zimbabwe	31.9; 48.3	8.8; 17.4	14.4; 36.3	3.9; 7.4	1975	48.54	47.33	0.31	−0.54
Middle Africa	44.3; 47.7	15.8; 27.4	18.6; 29.7	6.0; 6.9	2005	36.77	49.53	0.27	0.08
Angola	48.2; 54.4	15.8; 35.8	18.2; 32.4	6.5; 7.4	1970	30.00	49.63	0.28	0.34
Cameroon	39.6; 46.4	13.2; 24.7	17.4; 31.3	5.2; 6.7	1985	38.54	52.73	0.37	−0.02
Central African Republic	35.8; 43.8	17.3; 31.0	10.7; 25.2	4.8; 6.0	1985	33.44	46.45	0.35	−0.03
Chad	46.1; 51.3	16.4; 28.6	17.8; 32.8	6.1; 7.4	2000	36.06	48.68	0.26	0.10
Congo	38.0; 43.1	11.8; 20.9	21.4; 29.8	5.1; 6.3	1980	43.16	55.68	0.33	−0.03

Table 1 (*Continued*)

Major area, African region and country	CBR range of variation* 1950–2010	CDR range of variation* 1950–2010	RNI range of variation* 1950–2010	TFR Range of variation* 1950–2010	Onset of sustained fertility rate since 1950	e(0) 1950–1955	e(0) 2005–2010	e(0) average yearly change 1950–1990 (a)	e(0) average yearly change 1990–2010 (b)
DR Congo	45.1; 48.8	16.9; 25.3	20.8; 30.7	6.0; 7.2	2000	39.06	48.27	0.20	0.04
Equatorial Guinea	32.9; 47.4	15.2; 30.4	10.5; 26.8	5.4; 5.9	2000	34.48	50.09	0.28	0.18
Gabon	30.3; 38.0	10.2; 27.7	2.6; 26.1	4.0; 5.7	1985	36.99	61.32	0.59	0.03
Sao Tome and Principe	35.8; 47.7	7.4; 21.0	27.0; 30.6	4.5; 6.5	1980	46.40	65.48	0.38	0.16
Northern Africa	24.8; 50.7	6.7; 23.5	18.0; 29.2	3.1; 6.9	1965	42.36	68.32	0.51	0.22
Algeria	19.3; 52.5	5.3; 23.7	14.0; 32.8	2.4; 7.6	1975	42.89	70.26	0.57	0.18
Egypt	24.3; 50.8	6.9; 25.3	17.4; 28.0	3.0; 6.7	1960	41.14	69.92	0.56	0.25
Libya	22.6; 50.8	4.2; 30.3	18.1; 34.8	2.7; 7.9	1980	36.67	74.18	0.77	0.27
Morocco	20.1; 51.5	6.1; 20.2	13.9; 32.2	2.4; 7.2	1965	45.67	69.66	0.44	0.26
Sudan	36.1; 46.9	8.9; 20.3	26.3; 32.7	4.8; 6.9	1980	44.54	60.90	0.27	0.23
Tunisia	16.9; 51.2	5.2; 24.5	11.6; 28.9	2.0; 7.1	1965	38.78	74.56	0.71	0.30
Western Sahara	22.6; 52.0	5.9; 27.9	16.7; 27.4	2.6; 6.6	1975	35.52	65.95	0.52	0.39
Southern Africa	22.8; 43.8	8.0; 20.4	8.0; 25.2	2.6; 6.3	1955	44.74	51.89	0.40	−0.36
Botswana	24.9; 47.3	7.1; 18.8	7.8; 35.1	2.9; 6.7	1970	47.66	46.53	0.40	−0.68
Lesotho	28.2; 42.7	9.9; 22.7	10.9; 26.8	3.4; 5.9	1975	42.16	45.59	0.38	−0.47
Namibia	28.1; 44.2	8.7; 22.9	18.3; 31.6	3.4; 6.6	1980	41.75	60.05	0.47	−0.02
South Africa	22.1; 43.7	7.9; 20.2	7.3; 24.7	2.6; 6.3	1955	45.01	52.21	0.40	−0.35
Swaziland	31.4; 49.3	9.5; 22.6	16.0; 35.7	3.8; 6.9	1975	41.42	47.35	0.43	−0.45
Western Africa	41.2; 47.8	13.6; 30.1	16.8; 28.4	5.7; 6.9	1985	33.75	52.32	0.35	0.18
Benin	39.1; 47.2	10.3; 34.5	6.3; 32.0	5.3; 7.0	1985	33.71	58.20	0.45	0.26
Burkina Faso	43.8; 49.0	12.8; 31.9	15.2; 31.0	6.1; 7.2	1985	30.94	53.97	0.46	0.18
Cape Verde	21.9; 48.9	5.3; 21.8	16.7; 32.2	2.6; 7.0	1970	48.07	73.16	0.43	0.32
Côte d'Ivoire	36.1; 54.3	13.6; 32.0	20.0; 33.3	4.9; 7.9	1975	32.14	48.71	0.52	−0.16
Gambia	43.6; 51.6	10.5; 32.7	10.9; 33.2	5.3; 6.3	1985	30.25	57.45	0.53	0.24
Ghana	33.1; 47.3	9.6; 21.1	23.6; 30.6	4.2; 7.0	1970	42.17	60.02	0.33	0.19
Guinea	39.2; 47.4	12.7; 33.2	13.9; 29.4	5.4; 6.6	1990	33.08	54.47	0.37	0.26
Guinea-Bissau	39.4; 54.5	13.5; 25.4	15.1; 32.9	5.3; 7.4	1955	40.58	52.99	0.19	0.19
Liberia	38.6; 49.2	10.3; 30.5	16.5; 29.8	5.2; 7.0	1985	33.12	58.11	0.36	0.43
Mali	47.7; 49.4	14.9; 38.3	10.2; 33.1	6.5; 7.1	1990	26.96	52.71	0.46	0.30
Mauritania	36.0; 49.8	9.2; 23.8	24.9; 30.3	5.0; 6.8	1975	38.59	60.66	0.48	0.12
Niger	50.4; 56.6	12.7; 28.1	26.6; 37.7	6.9; 7.8	2000	34.97	55.64	0.19	0.52
Nigeria	42.2; 47.1	14.9; 29.6	16.5; 27.6	6.0; 6.8	1985	34.01	50.18	0.30	0.16

Table 1 (Continued)

Major area, African region and country	CBR range of variation* 1950–2010	CDR range of variation* 1950–2010	RNI range of variation* 1950–2010	TFR Range of variation* 1950–2010	Onset of sustained fertility rate since 1950	e(0) 1950–1955	e(0) 2005–2010	e(0) average yearly change 1950–1990 (a)	e(0) average yearly change 1990–2010 (b)
Senegal	38.8; 50.5	8.3; 29.0	20.7; 32.6	5.1; 7.5	1980	35.47	62.23	0.51	0.25
Sierra Leone	39.8; 49.3	18.7; 34.9	9.2; 24.0	5.2; 7.0	1985	28.81	43.97	0.26	0.19
Togo	38.0; 48.3	11.8; 30.0	17.5; 32.3	4.9; 7.3	1980	35.30	54.73	0.50	−0.03

Source: Author's derivations from calculations based on online datasets from United Nations, Department of Economic and Social Affairs, Population Division (2013). World Population Prospects: The 2012 Revision, DVD Edition. New York: United Nations.
Notes: CBR = crude birth rate (births per 1,000 population); CDR = crude death rate (deaths per 1,000 population); RNI = rate of natural increase (per 1,000 population); TFR = total fertility rate (number of children per woman); e(0) = life expectancy at birth.
*Range of variation = (Minimum value of the indicator over the past 60 years; Maximum value of the indicator over the past 60 years).
(a) = Average yearly gain (+) or loss (−) in life expectancy at birth over the period 1950–1955 to 1985–1990;
(b) = Average yearly gain (+) or loss (−) in life expectancy at birth over the period 1985–1990 to 2005–2010.

The world's population is shifting towards urban areas, with an estimated 51.6% living in urban areas in 2010 compared with 29.4% in 1950. By contrast, in 2010 only 39.2% of the population in SSA were urban dwellers. Due in large part to a diet linked with NCDs, urbanization is now linked to higher mortality rates for heart disease and chronic NCDs in poor countries (79–81, 134). Stuckler (135) found that each 1% increase in the population living in urban settings increased the long-term growth of chronic NCDs in the population by 3.2%.

Median age is a good summary of how young or old a population is. Half of the population in SSA is under 19 years, and half of the African population is under 20 years, indicating that these populations are extremely young. Population growth combined with the young populations typical of African countries and poorly resourced communities are influential factors of global health and human health in Africa (136).

The proportion of the population aged 60 years or older is lower in Africa than anywhere else in the world, ranging from 4.5% in middle Africa to 7.4% in northern Africa. Even if no changes were to occur in the age-specific risks of dying due to NCDs, population growth and aging in Africa would still produce large increases in the burden of mortality due to NCDs. The 2001 South African population census reported that 7.3% of the population was aged 60 years and older, and it is projected that even without changes in the risk factor profile or the mortality rates from cardiovascular disease, this demographic change will result in a doubling of the number of cardiovascular deaths in South Africa by 2040 (44).

It is conjectured that mortality decline will take place first, followed by fertility decline occurring 50 or more years from the onset of sustained fall in mortality. Levels of fertility and infant mortality are higher in Africa than other regions of the world, and within Africa the highest rates of fertility and infant mortality are in middle Africa. On average over the last 60 years, gains in life expectancy have also been lower in African regions (and lowest in middle Africa) than in the other regions of the world. During the transitional stage, the onset of mortality decline to moderate levels is expected to be followed after a lag time by fertility decline to less than five children per woman (hypothesis 2). It is unknown when mortality started declining in many African countries, since comparable data are available only since 1950.

Fertility rates in African countries and regions over the last 60 years in comparative perspective

Out of all African countries considered in this analysis using the latest and best available estimates from the United Nations' online databases, a total fertility rate of five children or more per woman was seen in 23 of them, concentrated in three regions of Africa: 11 of the 16

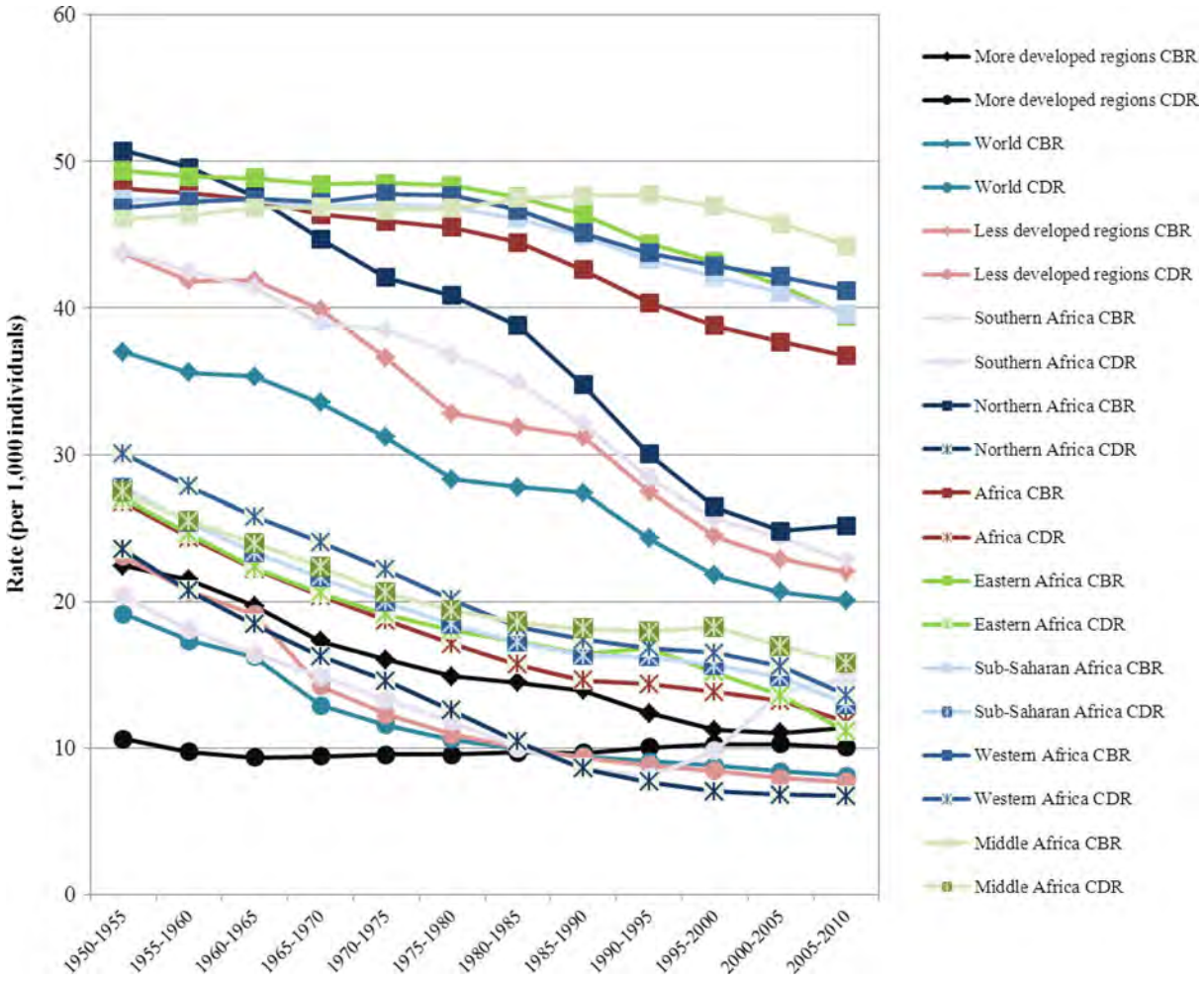

Fig. 1. Crude birth and death rates for Africa and its regions, and other regions of the World, 1950–2010.

countries of Western Africa (with Niger having the highest fertility rate, more than seven children per woman, in 2005–2010), eight of the 16 countries of Eastern Africa, and four of the nine countries of Middle Africa. For infant mortality, 11 countries (three in Western Africa, five in Middle Africa and three in Eastern Africa) had rates higher than 100 deaths per 1,000 live births in 2005–2010. Finally, 14 countries (three in Western Africa, two out of five in Southern Africa, five in Middle Africa, and four in Eastern Africa) had life expectancies at birth that were less than 50 years in 2005–2010.

Figure 2 displays the patterns of fertility change in Africa and African regions in comparison to other regions of the world. Between 1950–1960 and 1990, fertility levels changed little and by 1990, Africa was the continent with the highest levels of fertility. Sub-Saharan countries seem to remain more systematically outside any established and uniform process of fertility change. For Africa and Sub-Saharan Africa, total fertility rates stand respectively at 4.9 and 5.4 children per woman in 2005–2010, compared to 6.7 and 6.8 children per woman in 1950–1955. In 2005–2010, the corresponding figures are 2.3 for Asia and 2.3 for Latin America and the Caribbean. The timing of the onset and pace of any decline of fertility decline have been quite variable over the past six decades in Africa, and signs of any decline occurred as late as the 1990s in many countries of middle Africa and western Africa. Furthermore, the pace of decline has been slowest in the regions of Sub-Saharan Africa than elsewhere. The many causes of these lowest rates in Sub-Saharan Africa reflect the familiar cultural, social, and economic constraints that impact women's reproductive health negatively in general. The region of the world where access and choice of contraception is still problematic and will remain so for a long time is sub-Saharan Africa.

Despite substantial reductions in mortality rates and increases in life expectancy over the past 60 years, many African countries and regions have maintained high fertility rates. There are important regional differences within the continent, with Middle Africa having the slowest progress in health improvements and maintaining high

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

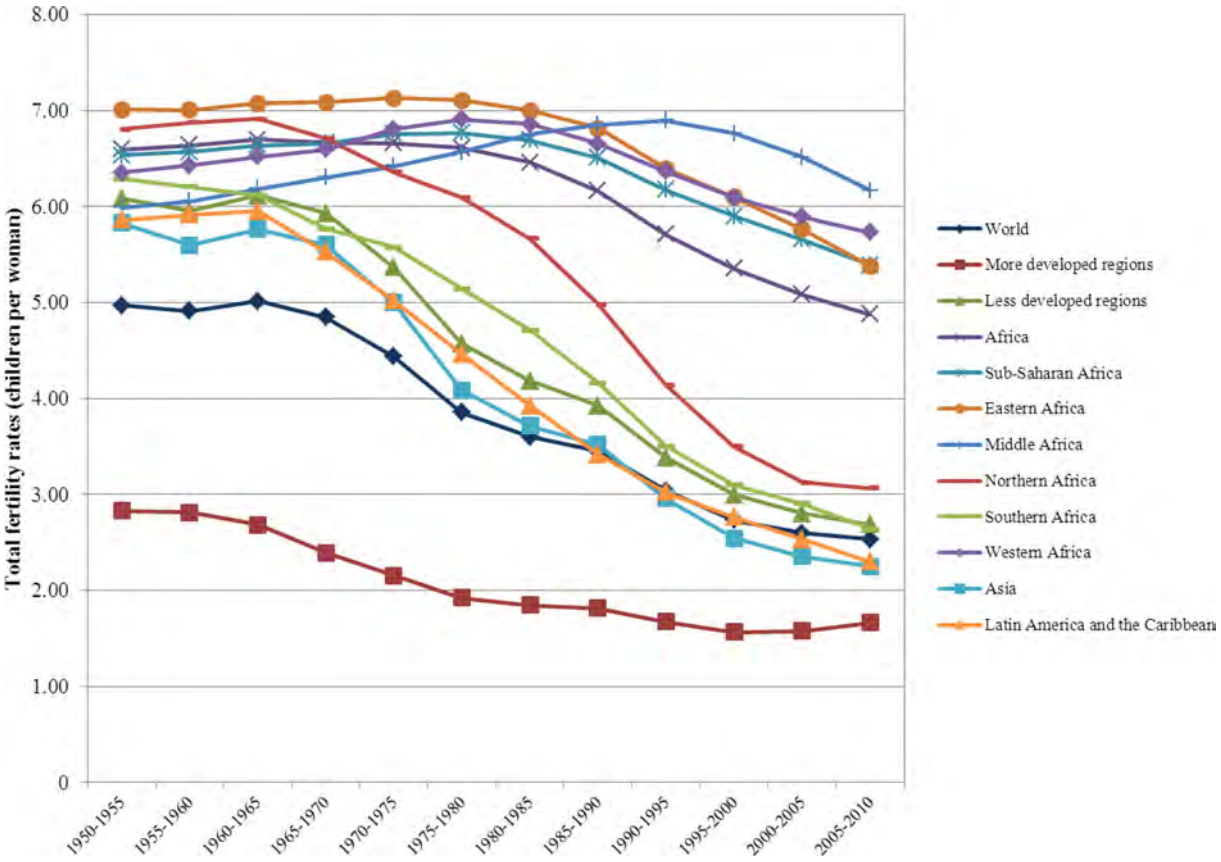

Fig. 2. Total fertility rates for Africa and its regions, and other regions of the World, 1950–2010.

and often increasing fertility levels. Across Africa, the pace at which changes have occurred and the factors influencing them has varied from one group of countries to another. With the exception of Mauritius, which shows a consistent pattern of changes over time, there is little evidence of a pattern in which the timing of changes in mortality can be linked to the timing of changes in fertility, and, indeed, most countries have achieved substantial reductions in mortality and gains in life expectancy at birth with virtually no variations in fertility. Despite what would be expected according to the frameworks of the demographic transition, the epidemiological transition, and the health transition, mortality declines in African nations were generally not followed by fertility declines, even after 60 years of scrutiny (2, 41, 137); these results substantiate arguments for what several scholars have termed 'African exceptionalism' (68–70), given Africa's departure from the prediction of the transition from high to low fertility in response to mortality decline inherent in demographic and epidemiological transitions. Northern and Southern Africa and the island countries (Mauritius, Reunion, Seychelles, Djibouti, and Cape Verde) are exceptions, but their transitions generally predated the 1960s. Although fertility decline started in SSA in 1980 and in Africa since 1975, fertility rates are still above the expected transition-level of five children per woman in many African regions (notably Middle, Eastern, and Western Africa). Hence, hypothesis 2 is not confirmed.

In short, the patterns of changes seen in Africa do not in general adhere to the predictions of the various transition frameworks, and the only places in SSA where these frameworks do apply are island countries such as Mauritius. The vast majority of the other African countries moved from a homogeneous pattern of populations with uniformly high levels of mortality and fertility prior to the independence years in the 1960s to a heterogeneous pattern of populations with widely varying rates of fertility, infant mortality, and life expectancy. This is in direct contrast to the expectations of the various transition frameworks, which conjecture populations that become more homogeneous over time with respect to patterns of mortality and fertility as well as life expectancy trajectories. We conclude that, with respect to mortality and fertility patterns, these frameworks do not perform well in accounting for the timing and sequencing of mortality decline and fertility decline since these are highly variable in Africa across African countries and regions.

We also compute the average yearly gains/losses in life expectancy between the 1950–1955 and 1985–1990 and between 1985–1990 and 2005–2010, to explore the impact of HIV/AIDS epidemic on trends in life expectancy

in African countries. We then compare these changes to those expected by the demographic transition. During the transitional stage (like from 1800 to 1950 in Europe), life expectancy at birth is expected to array between 40 and 65 years (hypothesis 3). Table 1 shows improvements in life expectancy for all countries through the 1980s. Since the 1990s, there are signs of stagnation or fall in life expectancy below the pre-1990 level. Such fall has been palpable in all countries from Southern Africa, in some countries of Middle Africa (Cameroon, Central African Republic, and Congo) or Eastern Africa (Kenya and Zimbabwe) and Western Africa (Côte d'Ivoire and Togo). Moreover, several countries have witnessed declines in life expectancy at birth between 1990 and 2010, falls which bring their life expectancy close to the 1950–1955 levels where some countries had life expectancy at birth under 40 years in 1950. Therefore, the pattern of change in life expectancy in Africa has been greatly influenced by the HIV/AIDS epidemic and among other influential factors, and hypothesis 3 is not accepted.

Overall, the evidence fails to confirm the conjectures of the demographic transition in Africa. These assessments also apply to shared features among the demographic, epidemiological and health transitions.

Testing the relevance of the epidemiological transition in Africa

We focus on the first three propositions which describe the epidemiological transition. We use reviewed evidence for the fourth and fifth propositions which are concerned with its determinants. The first proposition relating to the relative role of mortality and fertility in the epidemiological transition has been tested above in the context of the assessment of the demographic transition.

Distribution and group ratio of cause-specific deaths in Africa in comparative perspective

The second proposition is that during the transition, there is a long-term shift in cause-of-death patterns whereby NCDs predominate as primary cause of death. We compare the causes-of-death pattern in Africa with other parts of the world, using: 1) the ratio of group II causes-of-death (NCDs) to Group I causes-of-death (communicable, maternal, perinatal and nutritional conditions); and 2) the distribution of the top 20 causes of death to appraise the weight of mortality from communicable diseases and the expected changes in life expectancy. The ratio of Group II to Group I causes of deaths gives an indication of the process of the epidemiological transition: the higher the ratio, the more the epidemiological transition is advanced in a society. Tables 2 and 3 show the distribution of causes of death and ratios of causes of death.

According to the Omran's epidemiological transition (11), mortality by cause of death in non-western countries was projected to shift in predominance from communicable diseases dropping from being responsible for 42.1% of all deaths in the less developed regions in 1970 to 19.4% in 2015, resulting in life expectancy at birth rising from 57.5 to 68.5 years (hypothesis 4). Mortality and cause of death patterns indicates that across the world, 27% of deaths are ascribed to communicable, maternal, perinatal and nutritional conditions. In developed regions that number is less than 7%. Africa stands far apart from the rest of the world, with 65% of deaths caused by these conditions; infectious and parasitic diseases alone are responsible for 41% of all deaths in Africa, compared with only 15% for the world as a whole. In sharp contrast, 64% of all deaths worldwide are due to NCDs; that figure is 87% in the developed regions and 28% in Africa. In most developing regions, noncommunicable diseases (Group II causes of death) already exceeded Group I causes of death in 1990, whereas the ratio was only 0.4 in SSA (67). Table 2 indicates that almost 20 years later, that ratio is still at 0.4 for Africa, compared with 2.4 for the world, 12.4 in developed countries, 11.9 in Eastern Asia, 4.9 in Latin America and the Caribbean, and 1.4 in Southern Asia. Hence, hypothesis 4 is not verified in Africa.

Table 2. Percent distribution and group ratio of cause of deaths in the world, in the developed and developing regions, and in Africa: 2008

	World	Developed regions	Africa	Latin America and the Caribbean	Eastern Asia	Southern Asia
All causes	100	100	100	100	100	100
Group I. Communicable, maternal, perinatal, and nutritional conditions	27	7	65	15	7	38
Group II. Non-communicable diseases	64	87	28	73	83	52
Group III. Injuries	9	6	7	12	10	10
Ratio of Group II to Group I deaths	2.4	12.4	0.4	4.9	11.9	1.4

Source: Author's calculations based on online World Health Organization's regional cause-specific mortality estimates for the year 2008. (http://www.who.int/healthinfo/global_burden_disease/en/).

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

Table 3. Percent distribution of the top 20 cause of deaths in the world, in the developed world, in developing regions, and in Africa: 2008

Top 20 causes of death: world, 2008		Top 20 causes of death: developed world, 2008		Top 20 causes of death: Africa, 2008	
1. Ischemic heart disease	12.75	1. Ischemic heart disease	15.86	1. HIV/AIDS	12.87
2. Cerebrovascular disease	10.81	2. Other cardiovascular diseases	9.01	2. Lower respiratory infections	11.18
3. Lower respiratory infections	6.09	3. Cerebrovascular disease	8.58	3. Diarrheal diseases	9.08
4. Chronic obstructive pulmonary disease	5.76	4. Trachea, bronchus, lung cancers	5.97	4. Malaria	7.47
5. Diarrheal diseases	4.33	5. Alzheimer and other dementias	4.22	5. Cerebrovascular disease	4.43
6. Other cardiovascular diseases	3.77	6. Lower respiratory infections	3.83	6. Tuberculosis	3.93
7. HIV/AIDS	3.12	7. Chronic obstructive pulmonary disease	3.58	7. Ischemic heart disease	3.70
8. Trachea, bronchus, lung cancers	2.44	8. Other malignant neoplasms	3.39	8. Prematurity and low birth weight	3.31
9. Tuberculosis	2.36	9. Colon and rectum cancers	3.31	9. Birth asphyxia and birth trauma	2.93
10. Road traffic accidents	2.12	10. Other digestive diseases	2.65	10. Other cardiovascular diseases	2.74
11. Hypertensive heart disease	2.03	11. Hypertensive heart disease	2.14	11. Neonatal infections and other conditions	2.52
12. Other unintentional injuries	2.01	12. Other respiratory diseases	2.13	12. Road traffic accidents	1.66
13. Other malignant neoplasms	1.88	13. Breast cancer	1.97	13. Other digestive diseases	1.64
14. Other digestive diseases	1.82	14. Nephritis and nephrosis	1.67	14. Meningitis	1.62
15. Prematurity and low birth weight	1.75	15. Pancreas cancer	1.63	15. Violence	1.60
16. Cirrhosis of the liver	1.49	16. Stomach cancer	1.57	16. Childhood-cluster diseases	1.38
17. Neonatal infections and other conditions	1.46	17. Self-inflicted injuries	1.52	17. Other unintentional injuries	1.35
18. Malaria	1.45	18. Prostate cancer	1.51	18. Other respiratory diseases	1.33
19. Self-inflicted injuries	1.37	19. Lymphomas, multiple myeloma	1.41	19. Chronic obstructive pulmonary disease	1.30
20. Birth asphyxia and birth trauma	1.37	20. Cirrhosis of the liver	1.41	20. Protein-energy malnutrition	1.05
All the other causes	29.82	All the other causes	22.64	All the other causes	22.94
Top 20 causes of death: Latin America, 2008		**Top 20 causes of death: Eastern Asia, 2008**		**Top 20 causes of death: Southern Asia, 2008**	
1. Ischemic heart disease	11.47	1. Cerebrovascular disease	21.61	1. Ischemic heart disease	12.88
2. Cerebrovascular disease	7.96	2. Chronic obstructive pulmonary disease	12.88	2. Diarrheal diseases	9.54
3. Lower respiratory infections	5.16	3. Ischemic heart disease	10.79	3. Chronic obstructive pulmonary disease	8.21
4. Other cardiovascular diseases	5.07	4. Trachea, bronchus, lung cancers	4.81	4. Cerebrovascular disease	7.92
5. Violence	4.30	5. Liver cancer	3.93	5. Lower respiratory infections	7.82
6. Hypertensive heart disease	3.51	6. Stomach cancer	3.67	6. Tuberculosis	3.30
7. Other digestive diseases	3.44	7. Road traffic accidents	3.00	7. Neonatal infections and other conditions	2.91
8. Other malignant neoplasms	3.22	8. Other cardiovascular diseases	2.34	8. Prematurity and low birth weight	2.70
9. Chronic obstructive pulmonary disease	3.11	9. Hypertensive heart disease	2.24	9. Birth asphyxia and birth trauma	2.20
10. Road traffic accidents	3.06	10. Other unintentional injuries	2.13	10. Road traffic accidents	2.03
11. Other respiratory diseases	2.56	11. Esophagus cancer	2.13	11. Other unintentional injuries	2.01
12. Cirrhosis of the liver	2.46	12. Lower respiratory infections	2.03	12. Self-inflicted injuries	1.89
13. Nephritis and nephrosis	2.04	13. Self-inflicted injuries	1.81	13. Childhood-cluster diseases	1.87
14. Other unintentional injuries	1.99	14. Other malignant neoplasms	1.76	14. Cirrhosis of the liver	1.85

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

Table 3 (Continued)

Top 20 causes of death: Latin America, 2008		Top 20 causes of death: Eastern Asia, 2008		Top 20 causes of death: Southern Asia, 2008	
15. Trachea, bronchus, lung cancers	1.83	15. Tuberculosis	1.69	15. Hypertensive heart disease	1.67
16. Stomach cancer	1.67	16. Colon and rectum cancers	1.24	16. Nephritis and nephrosis	1.61
17. HIV/AIDS	1.42	17. Cirrhosis of the liver	1.22	17. Other malignant neoplasms	1.57
18. Prematurity and low birth weight	1.34	18. Falls	1.20	18. Other digestive diseases	1.48
19. Prostate cancer	1.25	19. Other digestive diseases	1.12	19. HIV/AIDS	1.47
20. Colon and rectum cancers	1.16	20. Other respiratory diseases	1.09	20. Falls	1.46
All the other causes	31.98	All the other causes	17.31	All the other causes	23.61

Source: Author's calculations based on online World Health Organization's regional cause-specific mortality estimates for the year 2008. (http://www.who.int/healthinfo/global_burden_disease/en/).

Table 3 presents the distribution of the 20 leading causes of death by region of the world. The list of the 20 leading causes of death in Africa highlights the double burden of increasing communicable and NCDs in Africa and contradicts the expectations of the epidemiological transition. Contrary to the expected shift between communicable and NCDs, with the former decreasing and the latter increasing, the emerging pattern of disease in the African context is one of increasing co-occurring morbidities from communicable and NCDs. The prevalence of old, emerging, or re-emerging communicable diseases such as tuberculosis, malaria, and HIV have been stable or increasing, while the levels of NCDs such as diabetes have been increasing (50, 51, 64, 91, 128, 138, 139).

The third proposition is that there are age and sex differentials resulting from the epidemiological transition, with the transition usually favoring young over old people (hypothesis 5) and mortality decline over time is expected to be more advantageous to females than males (hypothesis 6). To test this proposition, we generate age-sex-specific estimates by calculating male-to-female mortality ratios over the 1950–2010 periods, for infant, child and adult mortality. We then evaluate changes over time in these ratios. To test hypothesis 5, falling trends in age-specific mortality rates are graphed on the same scale in Fig. 3 (for infant mortality), Fig. 4 (for U5MR) and Fig. 5 (for adult mortality). To test hypothesis 6, ratios of male to female of infant, child and adult mortality are presented in Table 4.

Infant mortality in African countries and regions over the last 60 years in comparative perspective

Figure 3 shows the changing trends of infant mortality rates across regions in Africa since 1950, in comparison to other regions of Africa and the World. Infant mortality rates, often used as a key indicator of the well-being of a population, have been uniformly higher in Africa than any other continent of the world since World War II. Within the African continent, three patterns emerge. First, of the five regions of Africa, infant mortality was highest in Western Africa from 1950 to 1955 through 1985–1990; since then, middle Africa has had the highest infant mortality rates in the continent. Second, infant mortality rates were higher in northern Africa than SSA through the 1960s. Third, infant mortality rates were consistently lowest in southern Africa until the mid-1990s, when northern Africa became the region of the continent with the lowest infant mortality rates; AIDS-related deaths triggered by mother-to-child transmission of HIV caused infant mortality rates in southern Africa to rise between 1995–2000 and 2000–2005. There have also been huge differences in levels and trends in infant mortality rates across countries within Africa. South Africa had the lowest infant mortality rates in 1950–1955 (96 deaths per 1,000 live births) and Burkina Faso the highest

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

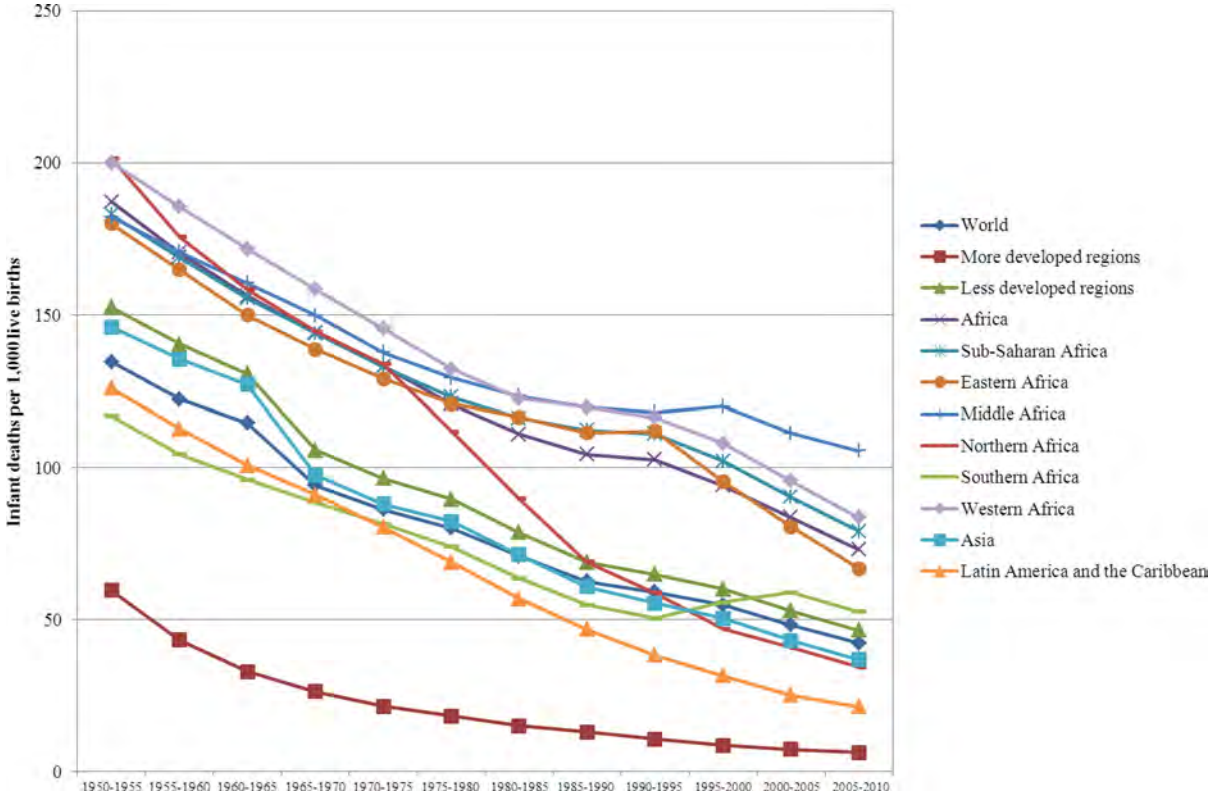

Fig. 3. Infant Mortality Rates for Africa and its regions, and other regions of the World, 1950–2010.

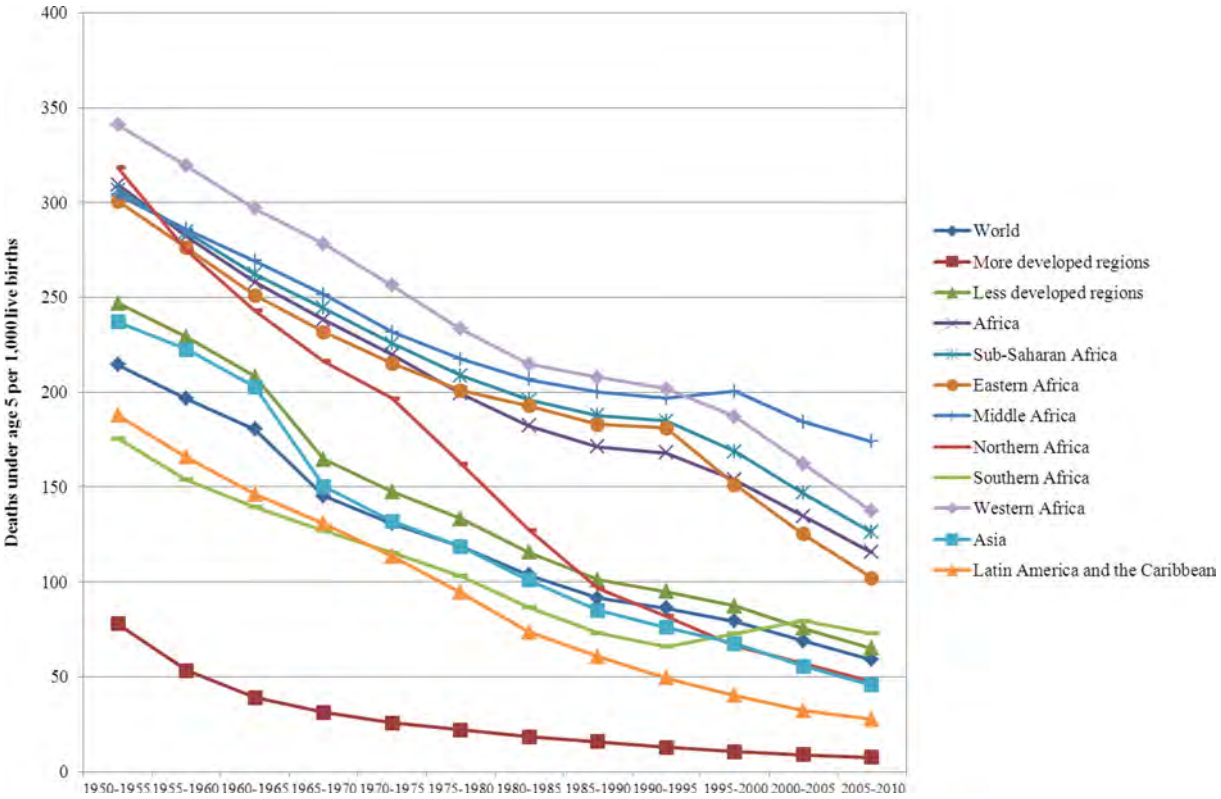

Fig. 4. Under Five Mortality Rates for Africa and its regions, and other regions of the World, 1950–2010.

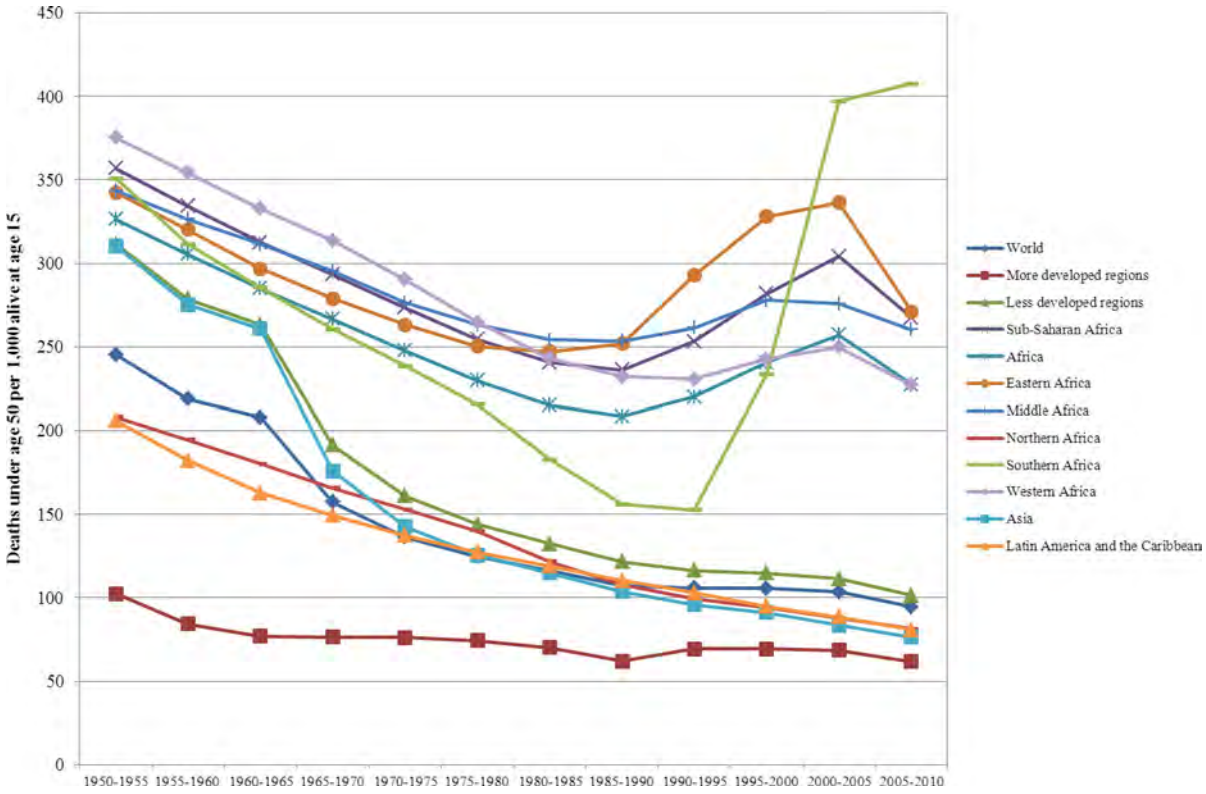

Fig. 5. Adult Mortality for Africa and its regions, and other regions of the World, 1950–2010.

(308 deaths per 1,000 live births). By 2005–2010 Mayotte and Reunion had the lowest rates, at six deaths per 1,000 live births, while Chad had the highest, with 131 deaths per 1,000 live births. Many African countries have maintained unacceptably high levels of infant mortality for over half a century and levels have stalled in a number of countries since the 1990s.

A major premise of the epidemiological transition is that mortality is the fundamental factor in population dynamics. However, between 1950–1960 and 1990, fertility levels in Africa changed little, and by 1990 Africa was the continent with the highest levels of fertility. In 2005–2010 total fertility rates in Africa were 4.88 children per woman, compared to 6.6 children per woman in 1950–1955; the corresponding estimates for SSA were 5.39 for 2005–2010 and 6.53 for 1950–1955, therefore situating sub-Saharan African fertility pattern using the latest and best estimates from the United Nations still at pre-transition levels (see Table 1). In 2005–2010, the corresponding figures were 2.3 for Asia and 2.3 for Latin America and the Caribbean. The timing of the onset and the pace of any fertility decline have been quite variable across Africa over the past 60 years, and in many countries of middle Africa and western Africa signs of a fertility decline did not start appearing until the 1990s. The pace of fertility decline has been slower in SSA than anywhere else in the world because of cultural, social, and economic constraints that impact women's reproductive health negatively in general (66, 138).

U5MR and adult mortality in African countries and regions in comparative perspective

Figures 4 and 5 show the changes in levels and pace of variations of U5MR and adult mortality in Africa and its individual regions in comparative perspective with other regions of the world. The situation in Africa is quite different from other regions of the world, where there is unquestionable evidence of regular mortality reductions over time in all regions (with the notable exception of eastern Europe, where adult mortality rate rose from 247 in 1995–2000 to 258 in 2000–2005, a divergent trend from the rest of Europe and developed world). In particular, as is the case with the fertility patterns discussed above, under-five and adult mortality patterns in sub-Saharan African countries do not adhere to the standard demographic, epidemiological, or health transitions. There are a variety of reasons for this. One is the widespread presence of wars and other forms of social and political unrest in the continent. A second is the emergence or re-emergence of such infections as HIV/AIDS and tuberculosis.

Perhaps the best way to observe the effect of wars and other conflicts on mortality rates across Africa is to examine those rates on a more local level. Indeed, local statistics in conflict areas indicate that mortality reductions

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

Table 4. Ratio of male to female (M/F) (per 100) of infant mortality rate (IMR), child mortality rate (CMR), and adult mortality rate (AMR) by major areas, African regions and countries: 1970–2010

	Ratio of M/F in the 1970s			Ratio of M/F in the 1980s			Ratio of M/F in the 1990s			Ratio of M/F in the 2000s		
Regions and African countries	IMR	CMR	AMR (a)	IMR	CMR	AMR (b)	IMR	CMR	AMR (c)	IMR	CMR	AMR (d)
World	111	90	128	110	89	138	108	87	144	107	86	142
More developed regions	129	123	230	129	122	232	128	124	255	126	124	244
Less developed regions	111	89	114	110	88	127	107	87	132	107	86	132
Sub-Saharan Africa	116	103	115	117	104	115	117	103	106	118	103	100
Eastern Africa			115			116			103			98
Burundi	115	99	115	115	99	117	115	99	106	115	98	107
Comoros	–	–	117	–	–	119	–	–	116	–	–	112
Djibouti	–	–	117	–	–	119	–	–	122	–	–	115
Eritrea	123	110	112	123	109	114	123	107	121	123	104	122
Ethiopia	122	109	114	124	107	113	125	103	103	126	97	100
Kenya	112	110	124	116	109	125	120	108	108	124	107	101
Madagascar	107	99	110	116	103	113	126	107	111	136	109	116
Malawi	112	104	111	112	107	115	111	110	102	110	115	86
Mauritius	126	98	207	131	114	212	133	116	234	131	115	232
Mayotte	–	–	186	–	–	210	–	–	235	–	–	263
Mozambique	117	98	114	114	102	114	111	106	113	108	110	102
Réunion	–	–	186	–	–	210	–	–	235	–	–	263
Rwanda	114	112	115	113	110	117	112	108	104	111	105	110
Seychelles	–	–	207	–	–	237	–	–	264	–	–	259
Somalia	116	108	114	116	108	115	116	108	117	116	107	117
South Sudan	–	–	112	–	–	113	–	–	111	–	–	104
Uganda	106	114	115	112	115	118	119	116	96	126	117	92
Tanzania	112	109	117	112	106	117	111	103	102	111	99	97
Zambia	112	91	119	114	100	109	116	108	95	118	117	97
Zimbabwe	134	94	122	124	102	127	115	114	101	105	129	90
Middle Africa			114			115			111			110
Angola	–	–	113	–	–	116	–	–	112	–	–	110
Cameroon	120	99	116	122	98	117	124	96	108	126	95	101
Central African Republic	113	99	119	114	101	123	114	103	109	115	105	104
Chad	115	108	110	116	105	112	118	102	106	119	99	103
Congo	119	95	118	119	94	119	119	95	105	119	95	105
DR Congo	106	112	114	106	112	114	106	112	114	106	112	116

Table 4 (*Continued*)

Regions and African countries	Ratio of M/F in the 1970s			Ratio of M/F in the 1980s			Ratio of M/F in the 1990s			Ratio of M/F in the 2000s		
	IMR	CMR	AMR (a)	IMR	CMR	AMR (b)	IMR	CMR	AMR (c)	IMR	CMR	AMR (d)
Equatorial Guinea	–	–	114	–	–	116	–	–	112	–	–	110
Gabon	137	109	118	137	107	122	137	103	114	137	103	102
Sao Tome and Principe	–	–	119	–	–	126	–	–	128	–	–	133
Northern Africa			125			132			137			139
Algeria	104	82	112	109	96	121	114	107	123	119	132	123
Egypt	103	81	157	105	81	159	109	84	160	122	116	161
Libya	114	82	126	114	78	129	114	69	128	114	60	129
Morocco	112	88	113	116	95	120	120	96	124	125	98	126
Sudan	120	101	117	118	101	118	115	102	132	113	102	135
Tunisia	109	78	106	119	94	124	128	112	147	137	141	154
Western Sahara	–	–	114	–	–	115	–	–	117	–	–	125
Southern Africa			133			151			115			94
Botswana	132	111	121	132	110	133	132	110	114	132	113	95
Lesotho	112	129	110	112	129	111	112	128	103	112	128	96
Namibia	110	110	114	118	111	129	126	113	115	134	117	128
South Africa	126	105	135	124	112	154	122	120	117	119	129	94
Swaziland	127	95	116	127	94	115	127	93	105	127	97	90
Western Africa			111			110			106			103
Benin	115	102	128	114	98	129	113	94	118	111	88	112
Burkina Faso	120	111	116	115	104	114	111	98	120	106	92	103
Cape Verde	–	–	106	–	–	132	–	–	189	–	–	207
Côte d'Ivoire	123	110	118	125	108	122	126	108	101	127	107	105
Gambia	–	–	112	–	–	113	–	–	114	–	–	114
Ghana	117	96	113	120	102	111	123	111	107	126	119	106
Guinea	111	96	111	113	98	111	116	99	97	118	100	108
Guinea-Bissau	–	–	117	–	–	122	–	–	113	–	–	109
Liberia	123	111	116	120	109	123	118	107	97	115	104	107
Mali	117	99	110	116	101	93	115	105	89	114	108	86
Mauritania	112	91	116	116	100	112	121	105	115	126	111	115
Niger	114	92	104	113	94	99	112	97	94	111	99	95
Nigeria	116	103	110	116	101	110	116	100	105	116	98	102
Senegal	117	97	112	117	101	120	118	106	118	119	113	117

Table 4 (Continued)

Regions and African countries	Ratio of M/F in the 1970s			Ratio of M/F in the 1980s			Ratio of M/F in the 1990s			Ratio of M/F in the 2000s		
	IMR	CMR	AMR (a)	IMR	CMR	AMR (b)	IMR	CMR	AMR (c)	IMR	CMR	AMR (d)
Sierra Leone	–	–	91	116	115	96	114	115	106	112	115	98
Togo	125	107	118	125	104	118	125	102	103	125	98	103

Source: Author's derivations from calculations based on online datasets from United Nations, Department of Economic and Social Affairs, Population Division (2013). World Population Prospects: The 2012 Revision, DVD Edition. New York: United Nations.

Notes: IMR = infant mortality rate; CMR = child mortality rate; AMR = adult mortality rates (probability of dying between exact ages 15 and 50). (a) = 1975–1980; (b) = 1985–1990; (c) = 1995–2000; (d) = 2005–2010.

would certainly have been deeper in the absence of wars in Africa.

The impact of wars is especially apparent in the data on U5MR in eastern African and middle African countries. In eastern Africa in the 1990–1995 period, five countries experienced reversals in mortality trends due to rampant wars: Rwanda, Burundi, Somalia, Zambia, and Uganda. In middle Africa wars have spanned several periods, and U5MR has increased for more than 10 years in several countries, notably Chad, Central African Republic, Congo, and Democratic Republic of Congo. In western Africa, wars reversed trends in U5MR rate in Sierra Leone, Liberia, and Côte d'Ivoire.

Similarly, AIDS has impacted and reversed trends in both U5MR and adult mortality in several countries in southern Africa, eastern Africa, middle Africa, and western Africa; in contrast, trends in mortality decline among children and adults have been robust in northern Africa (as well as the islands of Mauritius, Mayotte and Reunion, which by all accounts are apart in terms of population change and epidemiological landscapes from the rest of Africa). Since the period 1995–2000, every country in southern Africa (with the exception of Namibia which has singled itself out as the sole southern African country with a sustained decline in U5MR over time) has experienced increases in both U5MR and adult mortality due to the HIV/AIDS epidemic. These increases have been huge and sustained for adult mortality rates, with increases ranging from 25.9% in Namibia to 46.8% for Lesotho. South Africa has been particularly affected, and it is the only southern African country that has experienced continued increases in adult mortality rates since the 1990s, jumping from 369 in 1995–2000 to 490 in 2000–2005 and 533 in 2005–2010. U5MR rates for the whole of southern Africa increased substantially from 66 per 1,000 live births in 1990–1995 to 73 in 1995–2000 and 79 in 2000–2005.

In Eastern Africa, AIDS led to a reversal in the trends of decreasing mortality rates especially for adults. In middle Africa, adult mortality rates increased between 1995–2000 and 2000–2005 in all countries but Congo and Democratic Republic of Congo. The effects of AIDS on mortality rates are most obvious in Cameroon and Gabon, neither of which has been engaged in wars for decades: In Cameroon, for example, U5MR rate rose from 150 in 1985–1990 and 1990–1995 to 155 in 1995–2000 and 157 in 2000–2005 and remained at 152 in 2005–2010, and adult mortality rate increased from 370 in 1995–2000 to 412 in 2000–2005 and 413 in 2005–2010. In Gabon adult mortality rates increased from 270 in 1995–2000 to 307 in 2000–2005 and 286 in 2005–2010.

On the other hand, since 2005 Africa has been experiencing some of the largest decreases in U5MR ever seen anywhere. Many individual African countries have experienced declines in their U5MR rates, with most

of these decreases having been 4.4% or more yearly, which is the rate of decline needed to meet the Millennium Development Goal of cutting by two-thirds the child-mortality rate between 1990 and 2015. The striking thing about the decreasing rates is how widespread they have been, occurring in all countries and regions of the continent.

These mortality situations and patterns are sharply different from the historical situations experienced in developed countries and are also in disconnect with the patterns predicted by the standard demographic, epidemiological, and health transition frameworks. Therefore, hypothesis 5 is rejected.

In Table 4, we test the hypothesis that during the epidemiological transition, mortality decline over time is more advantageous to females than males (hypothesis 6). The mortality rates considered refer to the infant mortality rates (the probability of dying between birth and exact age 1 year), the child mortality (the probability of dying between exact ages 1 and 5 years), and the adult mortality (the probability of dying between exact ages 15 and 50 years). Despite falls in mortality, sex mortality differentials and male disadvantage has been noted whereby boys tend to have higher mortality rates than girls at early ages, leading to male-to-female mortality ratios expected to be greater than 100. The sex ratio of child mortality is generally lower than the sex ratio of infant mortality. During the epidemiological transition, it is expected that improvements in living conditions will make infectious diseases recede as a cause of death. On the one hand, as this occurs, perinatal and congenital causes form an increasing share of total mortality among infants; since newborn girls tend to be less vulnerable to such perinatal conditions (including birth trauma, intrauterine hypoxia and birth asphyxia, prematurity, respiratory distress syndrome and neonatal tetanus) and congenital anomalies, this will affect more boys than girls and male-to-female mortality ratio will rise because of this male disadvantage. On the other hand, external causes more typically affecting boys, form an increasing share of mortality for children between ages 1 and 5. Hence, as overall levels of mortality fall, female advantage in infant and child mortality would normally increase assuming no sex-specific changes in the treatment of children.

In Table 4, while there is little change over time in sex ratio of infant mortality or child mortality in most countries, excess female child mortality can be found notably concentrated at various periods of time in Northern Africa (Algeria, Egypt, Libya, Morocco and Tunisia), Western Africa (Benin, Burkina Faso, Ghana, Guinea, Mali, Niger and Senegal) and Middle Africa (Cameroon, Congo). Previous studies have signaled sex differential treatment in some of these countries but not in other (138). There is also a relative stability of male to female adult mortality ratios over time across countries, with notable exceptions of high-HIV countries of Eastern Africa (Malawi, Uganda, Tanzania, Zambia and Zimbabwe) and Southern Africa (Botswana, Lesotho, South Africa and Swaziland) where there is a consistent pattern of excess female mortality in recent periods most likely due to AIDS mortality. These results provide no evidence consistent with the conjectures from the epidemiological transition; hence, we reject hypothesis 6.

Testing the relevance of the health transition in Africa

All of the assessments above apply here. In the health transition view, formal education in general and female education notably is often cited as the most powerful instrument that has been shown to accelerate social changes enhancing health and improving life expectancies. Concerning sex differentials, mortality decline over time is expected to be more advantageous to females than males. In Western countries, female age-specific mortality risks shifted from being higher than those for the male in age groups under 45 to lower risks than the male in all age groups as life expectancy reached 60 or 70 years. Since improvement in survival is expected to favor females during the course of the epidemiological transition, sex differences favoring female populations in life expectancies at birth, at age 15 and at age 60 suggest the occurrence of health transition (hypothesis 7). These age-specific sex differentials in life expectancy are used to determine if there is any pattern of change consistent with these predictions from the health transition perspective.

Life expectancy in African countries and regions over the last 60 years in comparative perspective

There are four emerging patterns, as illustrated in Fig. 6 for both sexes combined to avoid clutter (separate figures for males and females are available upon request). First, life expectancy at birth has been consistently lower in SSA than in Northern Africa for both males and females. Second, among both sexes Western Africa had the lowest life expectancy in Africa from 1950 to 1955 though the mid-1980s, but since then Middle Africa has had the lowest. Third, while Southern Africa was the region of the African continent with the highest life expectancy at birth from 1950 to 1955 through 1985–1990, life expectancy at birth in this region declined by 6 years between 1990–1995 and 2005–2010; losses were more drastic for females (13 years) than for males (8 years), which resulted in a narrowing of the life expectancy gap between females and males from 7 years in 1990–1995 to only 2 years in 2005–2010. Fourth, Eastern Africa experienced a 1-year loss in life expectancy at birth between 1985–1990 and 1990–1995, which can be ascribed to a dramatic drop in life expectancy in one country: The Rwanda genocide led to a decline in life expectancy in Rwanda from 45 years in 1985–1990 to a historically low life expectancy of 24 years. This demonstrates how wars and civil conflicts – a

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

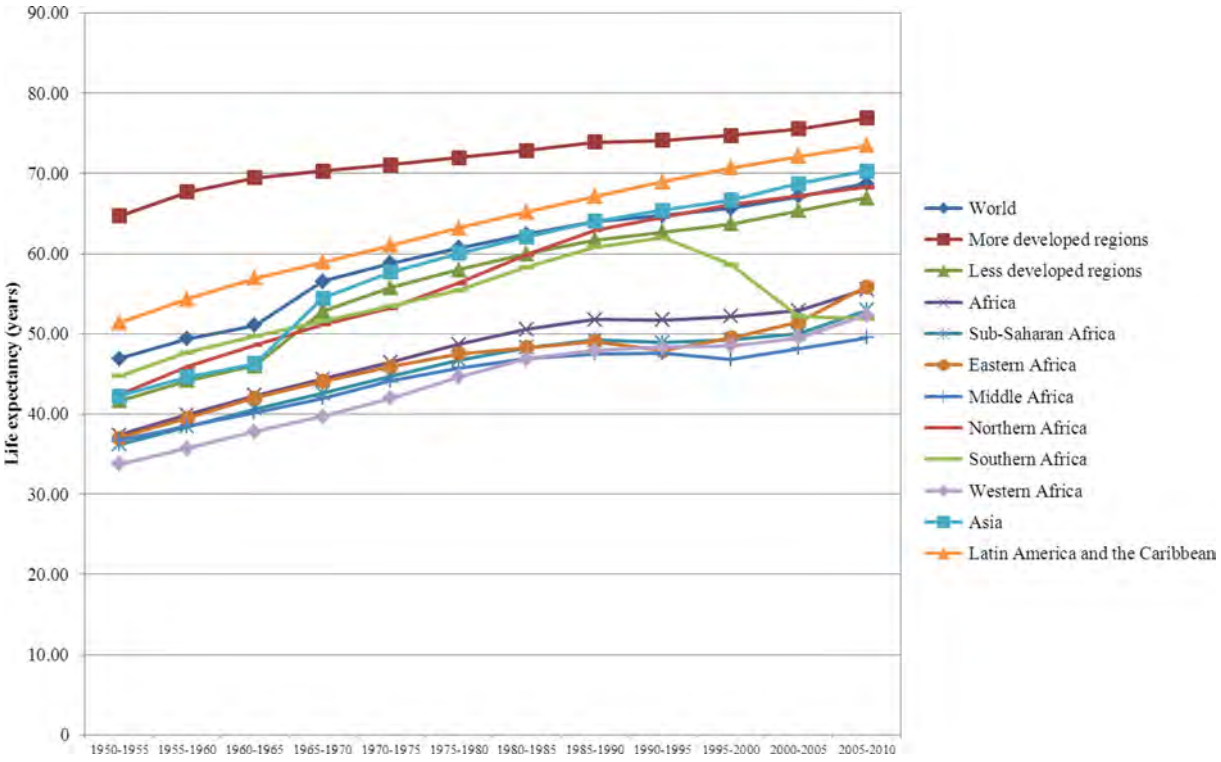

Fig. 6. Life Expectancy at Birth for Africa and its regions, and other regions of the World, 1950–2010.

fact of life in African countries since the independence years in the 1960s – play a noticeable role in mortality patterns and life expectancy prospects in the continent and in SSA in particular.

Table 5 present the sex differentials in life expectancies at birth, at age 15 and at age 60 from 1950–1955 to 2005–2010, at four distinct periods of time, to capture possible period effects in these differentials.

Just as the reasons for health improvements represent an important area of investigation in health transition research, so do sex differentials in life expectancies. The epidemiological and health transitions suppose that because infants and children of both genders, adolescents and young adult women were at greatest risk of premature death from infectious diseases during the pre-transitional period, reduction in the incidence of these conditions will result in an improvement in life expectancy, initially due to early reduction of infant and maternal mortality. A number of findings emerge from Table 5. There is female advantage in life expectancy at birth and at age 15 in all major regions considered and in all African countries with the exception of Malawi, Zimbabwe, Botswana, Swaziland and Mali in 2005–2010. One plausible explanation is the excess female AIDS-related mortality or harmful reproductive practices for female children and adolescent girls in these countries (66, 138). Deleterious reproductive health practices or differential treatment may also account for the female disadvantage in life expectancy at ages 15 and 60 for Benin, Mali and Sierra Leone in earlier periods than the 1990s and beyond when the possibility of AIDS-reducing life expectancy cannot be ruled out in some African countries with high HIV infection rates. Finally, there is robust female advantage in life expectancy at age 60 for all regions and countries over time, but Sierra Leone up to the 1980–1985 period. Overall, there are some inconsistent patterns for a handful of countries, but a general female advantage in life expectancy at birth, at age 15 and age 60 for the vast majority of countries.

In Table 6, the patterns of trends in total fertility and life expectancies between 1950–1955 and 2005–2010 for African regions (and SSA notably) are compared with the other world regions.

With respect to fertility trends, it is obvious that 'African exceptionalism' is undeniable in the sense that fertility decline has been slowest in SSA and its regions, compared to the rest of regions.

People around the world, including all its major regions, continents, and countries, have enjoyed significant increases in life expectancies over the last 60 years. This has largely been the result of medical advances directed against infectious and parasitic diseases in developing countries along with the cardiovascular revolution in the developed world, which together have been pushing forward the threshold conjectured by Omran's epidemiological transition. Table 6 demonstrates that the exception is Africa (and SSA in particular), which had a life expectancy at birth of less than 40 years in 1950–1955 and has ever since had the world's lowest life expectancy.

Table 5. Sex differential (female–male) in life expectancy at birth, at age 15 and at age 60 by major areas, African regions and countries: 1950–2010

Major areas, African regions and countries	F–M in 1950–1955			F–M in 1980–1985			F–M in the 1990–1995			F–M in 2005–2010		
	e(0)	e(15)	e(60)	e(0)	e(15)	e(60)	e(0)	e(15)	e(60)	e(0)	e(15)	e(60)
World	2.00	3.07	2.02	4.53	4.73	3.05	4.63	4.85	3.15	4.48	4.58	2.98
More developed regions	5.10	4.53	2.50	7.52	7.23	4.21	7.78	7.57	4.31	7.04	6.93	3.95
Less developed regions	0.76	1.83	1.43	3.22	3.32	2.05	3.40	3.56	2.23	3.57	3.63	2.29
Sub-Saharan Africa	2.54	2.09	0.96	3.03	2.57	1.24	2.88	2.37	1.29	1.78	1.22	1.33
Africa	2.34	2.26	1.05	3.16	2.79	1.46	3.07	2.63	1.53	2.15	1.68	1.57
Eastern Africa	2.80	2.31	1.01	3.13	2.60	1.13	2.82	2.21	1.19	1.91	1.07	1.37
Burundi	3.13	2.50	1.05	3.53	2.76	1.18	3.02	2.31	1.13	2.97	2.02	1.21
Comoros	2.50	2.07	1.23	4.00	2.93	1.66	3.67	2.81	1.65	2.67	2.35	1.56
Djibouti	2.66	2.27	0.97	3.09	2.57	1.12	3.25	2.83	1.15	2.94	2.40	1.19
Eritrea	3.96	2.60	2.42	3.99	2.77	2.44	4.47	3.65	2.75	4.62	4.37	3.30
Ethiopia	2.59	2.18	0.91	2.90	2.41	1.04	2.64	2.09	1.10	2.24	1.36	1.21
Kenya	3.63	2.78	1.18	3.75	2.93	1.25	3.61	2.80	1.35	2.43	1.39	1.41
Madagascar	2.03	1.47	0.53	2.07	2.12	0.97	2.60	2.35	1.05	2.85	2.27	1.22
Malawi	0.92	1.22	0.52	2.00	2.10	0.97	2.18	2.01	1.20	−0.03	−0.53	1.77
Mauritius	2.47	3.16	2.92	7.30	7.12	4.49	7.42	7.14	4.15	6.86	6.71	3.90
Mayotte	5.68	6.21	4.40	8.09	7.80	4.22	7.92	7.70	4.20	7.40	7.28	4.14
Mozambique	2.40	2.05	0.85	3.10	2.51	1.07	2.99	2.48	1.07	2.25	1.09	1.51
Réunion	5.68	6.21	4.40	8.09	7.80	4.22	7.92	7.70	4.20	7.40	7.28	4.14
Rwanda	3.13	2.51	1.06	3.28	2.87	1.13	3.37	2.29	1.07	2.59	2.02	1.16
Seychelles	4.43	4.18	1.62	7.90	7.38	3.92	10.04	9.62	5.16	9.22	8.81	4.81
Somalia	2.96	2.37	0.98	3.09	2.51	1.07	3.04	2.54	1.09	3.15	2.63	1.16
South Sudan	2.73	2.21	0.90	2.82	2.34	0.99	2.76	2.30	1.04	1.99	1.46	1.07
Uganda	3.13	2.51	1.06	3.13	2.66	1.11	2.40	1.81	1.10	1.06	0.06	1.16
Tanzania	3.34	2.63	1.11	3.24	2.66	1.15	2.87	2.14	1.26	1.62	0.83	1.29
Zambia	3.00	2.45	1.05	3.39	2.89	1.22	1.20	0.39	1.14	1.26	0.59	1.14
Zimbabwe	3.10	2.56	1.12	3.71	2.97	1.26	3.06	2.51	1.63	−1.03	−1.99	1.72
Middle Africa	3.12	2.53	1.07	2.92	2.45	1.05	3.07	2.50	1.12	2.79	2.11	1.15
Angola	2.88	2.35	0.97	2.93	2.40	1.02	3.86	2.89	1.24	2.83	2.16	1.15
Cameroon	2.66	2.25	0.96	3.07	2.57	1.12	2.93	2.42	1.12	1.86	1.15	1.12
Central African Republic	2.88	2.33	0.97	4.94	3.63	1.38	4.49	3.31	1.58	3.32	1.67	1.52
Chad	5.50	3.74	1.52	2.23	2.10	0.91	2.33	2.02	1.02	1.48	1.24	0.98
Congo	1.79	1.83	0.81	2.98	2.59	1.10	3.03	2.29	1.26	2.51	1.73	1.19
DR Congo	2.82	2.34	0.99	2.83	2.38	1.03	2.85	2.41	1.05	3.31	2.65	1.16

Table 5 (*Continued*)

Major areas, African regions and countries	F–M in 1950–1955			F–M in 1980–1985			F–M in the 1990–1995			F–M in 2005–2010		
	e(0)	e(15)	e(60)	e(0)	e(15)	e(60)	e(0)	e(15)	e(60)	e(0)	e(15)	e(60)
Equatorial Guinea	2.98	2.39	0.99	3.34	2.70	1.12	3.19	2.55	1.13	2.66	2.05	1.16
Gabon	3.08	2.48	1.04	3.19	2.63	1.14	3.00	2.44	1.16	2.10	1.46	1.20
Sao Tome and Principe	2.89	2.41	1.04	3.68	2.91	1.26	3.64	2.89	1.26	3.80	2.98	1.31
Northern Africa	1.88	3.44	1.75	3.75	3.69	2.17	4.03	3.76	2.23	3.95	3.66	2.25
Algeria	1.25	1.59	0.77	2.93	2.70	1.61	3.20	2.82	1.78	3.08	2.73	1.85
Egypt	1.04	5.25	2.80	4.53	4.97	2.84	4.79	4.74	2.72	4.65	4.44	2.68
Libya	2.25	0.92	0.26	3.63	3.11	1.76	3.36	3.00	1.86	3.81	3.64	2.82
Morocco	3.43	2.75	1.20	3.05	2.73	1.58	3.40	2.92	1.79	3.40	2.91	1.89
Sudan	2.84	2.39	1.03	2.93	2.48	1.08	3.13	2.78	1.15	3.46	3.16	1.16
Tunisia	2.13	1.31	0.31	3.15	2.61	1.85	4.59	4.20	3.07	4.74	4.72	3.47
Western Sahara	3.09	2.33	1.24	3.26	2.55	1.55	3.30	2.66	1.64	3.83	3.03	1.85
Southern Africa	2.23	1.75	2.13	6.56	5.61	3.43	6.82	6.00	3.90	2.40	1.39	4.07
Botswana	3.90	2.90	1.58	4.15	3.18	1.79	5.19	4.51	2.52	−0.31	−1.06	3.40
Lesotho	2.63	2.16	1.27	2.60	2.26	1.49	2.91	2.52	1.65	0.49	−0.47	1.71
Namibia	4.60	3.20	1.59	4.56	3.32	1.82	4.83	3.49	1.99	5.65	4.59	2.39
South Africa	2.00	1.56	2.21	6.90	5.89	3.68	7.21	6.35	4.19	2.58	1.55	4.32
Swaziland	4.00	2.88	1.49	3.88	2.92	1.69	2.94	2.33	1.67	−0.51	−1.56	1.79
Western Africa	2.13	1.75	0.56	2.26	1.87	0.87	2.17	1.79	0.86	1.01	0.93	0.62
Benin	0.30	−1.29	−0.49	6.42	4.38	1.71	4.58	2.46	1.28	2.74	1.97	1.11
Burkina Faso	1.98	2.74	0.85	2.36	2.32	1.11	2.36	2.28	1.07	1.16	1.02	0.72
Cape Verde	2.21	2.08	1.33	2.54	2.30	1.57	6.80	4.67	2.40	7.89	6.50	4.30
Côte d'Ivoire	1.84	1.88	0.80	4.01	3.01	1.27	3.86	2.94	1.25	1.54	1.07	0.52
Gambia	2.35	1.89	0.53	2.56	2.06	0.95	2.57	2.11	1.04	2.55	2.18	1.14
Ghana	0.49	0.88	0.41	2.34	1.99	1.00	2.32	2.04	1.11	1.71	1.41	0.91
Guinea	2.26	1.76	0.54	2.40	2.01	0.89	1.47	1.12	0.80	1.37	1.51	0.87
Guinea-Bissau	1.93	1.62	0.63	3.84	2.84	1.20	4.40	3.23	1.42	2.88	1.66	0.93
Liberia	5.17	3.49	1.12	3.70	2.68	1.18	4.01	3.09	1.35	1.74	1.45	0.89
Mali	1.41	1.88	0.48	1.72	1.59	0.76	−0.29	−0.80	0.08	−0.43	−0.88	0.09
Mauritania	0.18	0.69	0.27	2.66	2.13	1.15	2.73	2.09	1.20	2.97	2.20	1.29
Niger	−0.12	0.47	0.12	0.21	0.46	0.40	0.52	0.51	0.55	0.17	0.22	0.49
Nigeria	2.65	1.97	0.63	2.09	1.75	0.81	2.07	1.74	0.81	0.58	0.57	0.41
Senegal	1.30	1.26	0.40	2.73	2.38	1.15	3.20	2.71	1.41	2.82	2.37	1.33

Table 5 (Continued)

Major areas, African regions and countries	F–M in 1950–1955			F–M in 1980–1985			F–M in the 1990–1995			F–M in 2005–2010		
	e(0)	e(15)	e(60)	e(0)	e(15)	e(60)	e(0)	e(15)	e(60)	e(0)	e(15)	e(60)
Sierra Leone	3.49	−0.24	−0.19	0.43	−1.15	−0.35	1.48	0.57	0.10	0.30	−0.17	0.01
Togo	1.63	1.43	0.46	3.67	2.88	1.36	1.73	0.95	0.73	1.46	0.97	0.65

Source: Author's calculations based on online datasets from United Nations, Department of Economic and Social Affairs, Population Division (2013). World Population Prospects: The 2012 Revision, DVD Edition. New York: United Nations.
Notes: e(0) = life expectancy at birth (in years); e(15) = life expectancy at age 15 (in years); e(60) = life expectancy at age 60 (in years).

In fact, the health disparities in terms of life expectancy have remained or have widened consistently over the last 60 years between sub-Saharan African regions and all the other regions of the world, with respect to gap in life expectancy at birth, at age 15 and at age 60. Moreover, whereas all other of the world's regions have enjoyed uninterrupted increases in life expectancy, SSA and its regions and countries had a peak in life expectancy in the early 1990s and by 2010 some sub-Saharan African countries, such as Zimbabwe, were experiencing pre-1950 life expectancy levels. The stagnation in gains and reversals in trends in life expectancy for sub-Saharan African countries in the 1990s mainly reflect increasing mortality from HIV/AIDS. The consistent evidence from Table 6 points towards movements away from predicted trajectories of life expectancies and fertility rates inferred by the demographic, epidemiological or health transitions as predictive models for describing the historical experiences of population health change in African regions and countries over the past 60 years.

Discussion

The demographic transition and epidemiological transition are compelling frameworks which have provided an accurate description of historical experiences regarding the secular declines in mortality and fertility and the associated changes in patterns of disease and causes of death through the 1950s in today's developed countries. New forms of demographic change (88, 89) and epidemiological landscape (84, 85) have emerged in the developed world around the 1960s and 1970s, and new population and health developments have been occurring in less developed countries too (12–16, 29, 45, 58, 66–70, 74, 80, 140). The extent to which these historical experiences from developed countries and predictions derived from them could inform the demographic and epidemiological changes in developing countries have been called into question from the beginning, and remains an empirical question to which this study has provided a systematic answer.

This study has shown that the demographic, epidemiological, and health landscapes of the African continent are significantly different from those experienced by the now-developed countries of Europe and North America during their demographic and epidemiological transitions. The evidence clearly shows that, with the notable exception of island African countries such as Mauritius and probably some northern African nations, over the past 60 years African countries have not embarked on any evidence-based shift from one epidemiological regime to another and have not seen demographic changes and health improvements, as would be predicted from the demographic, epidemiological, and health transition frameworks or from the global trends when they are compared to other developing countries outside of Africa.

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

Table 6. Global context of patterns of trends in total fertility (TF) and life expectancy at birth (e(0)), at age 15 (e(15)) and age 60 (e(60)) in Africa

	TF: 1950–55		TF: 2005–10		e(0): 1950–55		e(0): 2005–10		e(15): 1950–55		e(15): 2005–10		e(60) 1950–55		e(60): 2005–10	
Major regions	(a)	(b)	(a)	(b)	(c)	(d)	(c)	(d)	(c)	(d)	(c)	(d)	(c)	(d)	(c)	(d)
World	4.97	1.57	2.53	2.86	46.91	−10.69	68.72	−15.78	46.94	−6.08	58.80	−11.31	14.16	−2.14	19.73	−4.27
More developed regions	2.83	3.70	1.66	3.73	64.67	−28.44	76.90	−23.96	55.65	−14.79	62.60	−15.11	16.84	−4.83	22.07	−6.61
Less developed regions	6.08	0.45	2.69	2.70	41.62	−5.40	66.96	−14.02	42.97	−2.11	57.43	−9.94	12.35	−0.34	18.54	−3.08
Sub − Saharan Africa	6.53	–	5.39	–	36.23	–	52.94	–	40.86	–	47.49	–	12.01	–	15.46	–
Africa	6.60	−0.06	4.88	0.51	37.38	−1.15	55.55	−2.61	42.43	−1.57	49.56	−2.08	12.52	−0.51	15.98	−0.52
Eastern Africa	7.01	−0.48	5.38	0.01	37.03	−0.80	55.85	−2.91	41.98	−1.12	48.80	−1.31	12.75	−0.74	17.02	−1.56
Middle Africa	5.99	0.54	6.17	−0.78	36.77	−0.54	49.53	3.42	41.96	−1.09	47.54	−0.05	12.79	−0.78	15.42	0.04
Northern Africa	6.81	−0.27	3.07	2.32	42.36	−6.13	68.32	−15.37	49.00	−8.14	57.32	−9.83	14.66	−2.65	17.56	−2.10
Southern Africa	6.28	0.25	2.64	2.75	44.74	−8.52	51.89	1.05	40.91	−0.05	42.06	5.43	11.75	0.26	15.26	0.20
Western Africa	6.35	0.18	5.73	−0.34	33.75	2.48	52.32	0.62	39.56	1.31	47.77	−0.28	11.19	0.82	14.06	1.40
Asia	5.83	0.71	2.25	3.14	42.24	−6.02	70.28	−17.34	42.96	−2.10	59.16	−11.67	12.27	−0.26	19.12	−3.66
Eastern Asia	5.60	0.93	1.61	3.78	46.21	−9.99	75.52	−22.58	44.19	−3.33	62.24	−14.75	11.62	0.39	20.65	−5.19
South-Central Asia	6.02	0.51	2.72	2.67	37.43	−1.21	65.53	−12.59	40.62	0.24	55.70	−8.21	12.68	−0.66	17.04	−1.58
Central Asia	5.23	1.30	2.67	2.72	54.51	−18.29	66.28	−13.34	51.45	−10.59	55.48	−7.99	15.28	−3.27	17.17	−1.71
Southern Asia	6.05	0.48	2.72	2.67	36.98	−0.75	65.51	−12.56	40.27	0.59	55.70	−8.21	12.55	−0.53	17.04	−1.58
South-Eastern Asia	5.92	0.62	2.35	3.04	43.98	−7.76	70.34	−17.39	45.01	−4.15	58.23	−10.74	13.95	−1.93	18.68	−3.22
Western Asia	6.32	0.22	2.92	2.47	42.61	−6.38	72.24	−19.30	47.20	−6.33	60.13	−12.65	14.10	−2.08	19.33	−3.87
Europe	2.67	3.87	1.54	3.85	63.59	−27.36	75.28	−22.34	55.71	−14.85	61.01	−13.52	16.79	−4.77	21.08	−5.62
Eastern Europe	2.91	3.63	1.41	3.98	60.33	−24.11	69.52	−16.58	54.46	−13.60	55.55	−8.06	16.54	−4.53	17.99	−2.53
Northern Europe	2.32	4.21	1.86	3.53	68.76	−32.54	79.11	−26.17	57.07	−16.21	64.62	−17.13	16.99	− 4.97	22.65	−7.19
Southern Europe	2.68	3.86	1.43	3.96	63.45	−27.22	79.92	−26.98	55.92	−15.06	65.48	−18.00	16.73	−4.72	23.15	−7.69
Western Europe	2.39	4.15	1.64	3.75	67.72	−31.50	80.21	−27.27	56.76	−15.90	65.66	−18.17	16.95	−4.94	23.46	−8.00
Latin America & Caribbean	5.86	0.67	2.30	3.09	51.37	−15.14	73.45	−20.51	49.62	−8.76	60.86	−13.37	15.36	−3.35	21.27	−5.81
Caribbean	5.27	1.27	2.37	3.02	52.05	−15.83	71.22	−18.28	50.21	−9.35	60.06	−12.57	15.57	−3.55	21.05	−5.59
Central America	6.73	−0.20	2.56	2.83	49.13	−12.91	75.26	−22.32	48.49	−7.63	62.42	−14.93	15.27	−3.26	22.00	−6.54
South America	5.66	0.87	2.19	3.20	52.10	−15.87	73.03	−20.09	49.92	−9.06	60.42	−12.93	15.37	−3.36	21.06	−5.60
Northern America	3.35	3.18	2.02	3.37	68.59	−32.36	78.36	−25.42	56.54	−15.68	64.08	−16.59	17.44	−5.42	22.88	−7.42
Oceania	3.84	2.70	2.47	2.92	60.44	−24.22	76.84	−23.90	52.49	−11.63	64.36	−16.87	16.35	−4.33	23.33	−7.87

Source: Author's calculations using online databases from United Nations, Department of Economic and Social Affairs, Population Division (2013). World Population Prospects: The 2012 Revision, DVD Edition. New York: United Nations.

Notes: SSA = sub-Saharan Africa; (a) = total fertility (TF); (b) = TF(SSA) − TF(Region); (c) = life expectancy (e); (d) = e(SSA) − e(Region).

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

Instead, the expected steady improvement in health and disease prevention has been short-circuited by PI, a lack of social consciousness, and health-care systems that are too ill-prepared and underfunded to deal effectively with the challenges of increasingly co-occurring communicable and NCDs in national populations. It seems certain that the epidemiological and demographic changes in Africa over the past six decades would have been completely different – and with a better outlook – if there had been 60 years of continuous social and political stability.

Africa is a continent of uncertainties, emergencies, and humanitarian crises against a backdrop of cultural, economic, political and geographic diversities, a situation that has created inertia in population and health development. As a consequence, the health and disease patterns in much of the continent remain characterized by the predominance of communicable diseases, with 65% of all deaths caused by diseases amenable to interventions. The 1990s saw losses in life expectancy for many national populations in Southern and Eastern Africa, and there is mounting evidence indicating that Middle Africa may undergo a similar course, triggered by the spread of the HIV/AIDS epidemic (141–145). The resulting interruptions in life expectancy improvements in many African countries are unprecedented in human history.

Before the 1970s, there were almost no examples of long-term reversals in mortality, with the obvious exceptions of those caused by war and famine. Reflecting this, many of the classic analyses of the 1970s that examined long-term demographic and epidemiological trends concluded that while further significant gains in longevity in countries with low mortality were unlikely, death rates in countries with the high mortality typical of African countries would fall, resulting in a worldwide convergence in mortality (2, 146). However, since the 1990s all five countries of Southern Africa, three countries in Middle Africa (Cameroon, Central African Republic and Chad), Burundi and Kenya in Eastern Africa, Côte d'Ivoire in West Africa, and several countries of Eastern Europe have experienced reversals in mortality trends and losses in life expectancy at birth. Most recently, from 1995 onwards, the mortality patterns has become much more diffuse, with several countries in Southern Africa, Eastern Africa, and Middle Africa experiencing stagnation or deceleration in improvements in life expectancy and declines in life expectancy at birth. These reversals have occurred in a context of pervasive civil wars and social unrest, PI, substandard public expenditures on health, precarious socioeconomic conditions for large percentages of the population who are mainly employed in the informal sector with no social security, the HIV/AIDS epidemic and its co-morbidities, an increasing prevalence of infectious diseases, and raging food insecurity in this historically agricultural continent.

By and large, the conjectured linkages between mortality, fertility, and population growth find little empirical support in much of Africa, calling into question the basic premise of the transition approach embodied in the demographic, epidemiological, and health transition models. Despite signs of an onset of fertility decline in a handful of African countries (39, 68, 147–149), the widening gap between fertility and mortality patterns within and across countries combined with the enduring prevalence of infectious diseases in the continent suggests that a new and different perspective is needed for understanding health and disease trends in Africa.

The lack of available and adequate historical epidemiological data on disease, morbidity and mortality by cause of death and the influences on them remain a major obstacle for analyzing health and disease patterns in Africa. The cause-of-death data in Africa remain very scarce, and there is an urgent need for collecting and analyzing such data. The relative weightings of the various sources of morbidity and mortality and the ways in which these change over time have great significance for the planning and provision of health and other human services. The lack of suitable longitudinal data on health and disease patterns is probably the major limiting factor for local and national studies for many African countries.

Our study indicates that, given that most populations in Africa live under continuous co-occurrence of infectious and non-infectious diseases origin with various degrees of chronicity, any useful study of the health and disease patterns in Africa will be theoretically, programmatically, and empirically limited unless morbidity and mortality are linked within a health and disease continuum framework. This is an area that deserves future research. In those African countries for which studies have been carried out, the disease and mortality patterns indicate that both morbidity and mortality have been rising over time for major diseases often occurring concurrently (44, 50, 51, 140). With an increasing dual burden of disease in SSA, our understanding of the associations between diseases will become of increased public health importance. Research is also needed to develop effective approaches to reducing the frequency and health impact of the co-morbidities.

Conclusion

Empirical evidence from 57 African countries substantiates that existing transition frameworks are not adequate for describing the patterns of demographic, epidemiological and health changes which have taken place in Africa over the last 60 years and beyond, nor are they predictive models of concurrent trajectories of population health resulting from historical and contemporary forces of change in the continent.

Prevailing frameworks of demographic, epidemiological, and health transitions as descriptive and explanatory

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

tools are beyond reasonable doubt incomplete or irrelevant in the African context on conceptual, methodological, empirical, and policy grounds. Charting the course of health and disease patterns in African countries and regions through functional systems of continuous and sustainable national and local data collection enterprises will significantly reduce the burden of disease in these countries and will undoubtedly narrow the health and economic gap between the developed and developing worlds as well as among African regions and countries.

Omran (10, 11) acknowledged that the dynamics of the Western transition were closely related to the unique historical and circumstantial experience of the industrial and social revolution in the West during the last three centuries, and may not automatically be transferable to less developed countries. Despite substantial uncertainty about the extent of changes in levels and differentials in mortality statistics and patterns of causes of death prevailing in various groups of regional or national populations studied, it is obvious that population and health changes in Africa have so far deviated from the demographic, epidemiological and health transitions.

Main findings

- For research, planning and policy purposes, empirical evidence spanning at least the past 60 years substantiates that the demographic, epidemiological and health transition frameworks are inadequate for establishing a comprehensive representation of past and current situations at the regional, national and sub-national levels in the African context.
- Regional and national patterns of changes in mortality, fertility, population growth, and cause of death structure in Africa and other regions of the world between 1950 and 2010, highlight that many African countries distinctively undergo uncharted paths compared to historical experiences of Europe and North America.
- The demographic transition, epidemiological transition and health transition are irrelevant for predicting how demographic and epidemiological changes will play out across and within African countries and regions where changing contexts of health, disease and mortality patterns are embedded in significant uncertainties.

Key messages for action

- Upon scrutiny of the concepts of demographic transition, epidemiological transition and health transition, these concepts may not be used synonymously. There are significant differences between them regarding their level of specificity and generalizability, their descriptive and explanatory dimensions, their prognostic implications and their empirical soundness in low- and middle-income countries.
- Model-based estimates of health, disease and mortality statistics for African countries are at best conjectural in the vast majority of African societies where vital events, health conditions, disabilities, accidents and their risk factors are undetected. Limited data available to health planners to better pinpoint and address these problems often emanate from small-scale rural sites or are poorly recorded or preserved. National governments, the World Health Organization and pertinent United Nations agencies, the World Bank, international foundations, funding agencies along with researchers, should move from ceaseless pious vows to concerted actions aligned to resources in order to collect nationally representative data on measures of health and disease burden as well as their risk factors in Africa for research, planning and policy.
- A new perspective embodying broadly suitable theoretical guidance is needed. Such perspective will be used to organize data collection, to develop models of demographic and epidemiological changes that can serve as a basis of hypotheses to be tested, and to inform a well-grounded understanding of how demographic and epidemiological profiles have changed or can be expected to change in the future. Such understanding will help the institutional capacity development as well as the planning and development of health systems and health policies in typically heterogeneous African societies.

Acknowledgements

A preliminary draft of this work was presented at the United States National Academy of Sciences' Workshop on 'The Continuing Epidemiological Transition in SSA', held in Johannesburg (South Africa) on October 20–22, 2011. Professor Barthélémy Kuate Defo was chair of the United States National Academy of Sciences' Committee on the Continuing Epidemiological Transition in SSA. I thank Peter Byass, Barney Cohen, Alan Lopez, Stig Wall, Jacques Vallin and Richard Suzman for their insightful comments, suggestions, and discussions on this work. The author, being the Guest Editor of the Special Issue for the journal, did not participate in the editorial decision-making process for his manuscript. The study was supported by the New Initiatives Grant Program from the Institut de Recherche en Santé Publique (IRSPUM) and the Global Health Competition Grant from the Direction des Relations Internationales (DRI) (Université de Montréal).

Conflict of interest and funding

The author declares that he has no conflict of interests.

References

1. Notestein FW. Population—the long view. In: Schultz TW, ed. Food for the world. Chicago: University of Chicago Press; 1945, pp. 36–57.
2. Omran AR. The epidemiologic transition: a theory of the epidemiology of population change. Milbank Mem Fund Q 1971; 49: 509–38.
3. Omran AR. Population epidemiology: emerging field of inquiry for population and health students. Am J Public Health 1974; 64: 674–79.
4. Omran AR. The world population problem. In: Omran AR, ed. Community medicine in developing countries. New York: Springer; 1974, pp. 107–8.
5. Omran AR. The epidemiologic transition in North Carolina during the last 50 to 90 years: I. The mortality transition. N C Med J 1975; 36: 23–8.
6. Omran AR. The epidemiologic transition in North Carolina during the last 50 to 90 years: II. Changing patterns of disease and causes of death. N C Med J 1975; 36: 83–8.
7. Omran AR. Epidemiologic transition in the United States: the health factor in population change. Popul Bull 1977; 32: 1–42.
8. Omran AR. A century of epidemiologic transition in the United States. Prev Med 1977; 6: 3–51.
9. Omran AR. Epidemiologic transition: theory. In: Ross JA, ed. International encyclopedia of population. New York: The Free Press; 1982, pp. 172–83.
10. Omran AR. The epidemiologic transition theory. A preliminary update. J Trop Pediatr 1983; 29: 305–16.
11. Omran AR. The epidemiologic transition theory revisited thirty years later. World Health Stat Q 1998; 51: 99–119.
12. Frenk J, Bobadilla J, Sepulveda J, Cervantes J. Health transition in middle-income countries: new challenges for health care. Health Policy Plan 1989; 4: 29–39.
13. Frenk J, Bobadilla JL, Stern C, Frejka T, Lozano R. Elements for a theory of the health transition. Health Transit Rev 1991; 1: 21–38.
14. Caldwell J, Santow G, eds. Selected readings in the cultural, social and behavioural determinants of health. Canberra: Australian National University; 1989.
15. Caldwell JC, Findley S, Caldwell P, Santow G, Cosford W, Braid J, et al., eds. What we know about the health transition: the cultural, social and behavioral determinants of health. Canberra: Australian National University; 1990.
16. Cleland J, Hill A, eds. The health transition: methods and measures. Canberra: Australian National University; 1991.
17. Caldwell JC, Caldwell P. What have we learnt about the cultural, social and behavioral determinants of health? From selected readings to the first health transition workshop. Health Transit Rev 1991; 1: 3–20.
18. Caldwell JC. Health transition: the cultural, social and behavioural determinants of health in the Third World. Soc Sci Med 1993; 36: 125–35.
19. Caldwell JC. Basic premises for health transition in developing countries. World Health Stat Q 1998; 51: 121–33.
20. Caldwell JC. Population health in transition. Bull World Health Organ 2001; 79: 159–60.
21. Caldwell JC. Demographers and the study of mortality, Scope, perspectives, and theory. Ann N Y Acad Sci 2001; 954: 19–34.
22. Caldwell JC. Toward a restatement of demographic transition theory. Popul Dev Rev 1976; 2: 321–66.
23. Salomon J, Murray CJL. The epidemiologic transition revisited: compositional models for causes of death by age and sex. Popul Dev Rev 2002; 28: 205–28.
24. Small-Raynor M, Phillips D. Late stages of epidemiological transition: health status in the developed world. Health Place 1999; 5: 209–22.
25. Hill K, Vapattanawong P, Prasartkul P, Porapakkham Y, Lim SS, Lopez A. Epidemiologic transition interrupted: a reassessment of mortality trends in Thailand, 1980–2000. Int J Epidemiol 2007; 36: 374–84.
26. Vallin J. Commentary: 'epidemiologic transition' interrupted or sweep to the second stage of 'health transition'? Int J Epidemiol 2007; 36: 384–6.
27. Vallin J. Diseases, deaths, and life expectancy. Genus 2005; 61: 279–96.
28. Mesle F, Vallin J. The health transition: trends and prospects. In: Caselli G, Vallin J, Wunsch G, eds. Demography, analysis and synthesis. A treatise in demography. New York: Elsevier; 2006, pp. 247–602.
29. Vallin J, Mesle F. Convergences and divergences in mortality: a new approach to health transition. Demographic Research 2004. DOI: 10.4054/DemRes.2004.S2.2.
30. Stevens G, Dias RH, Thomas KJA, Rivera JA, Carvalho N, et al. Characterizing the epidemiological transition in Mexico: national and subnational burden of diseases, injuries, and risk factors. PLoS Med 2008; 5: e125. DOI: 10.1371/journal.pmed.0050125.
31. Mackenbach JP. The epidemiologic transition theory. J Epidemiol Community Health 1994; 48: 329–31.
32. McKeown RE. The epidemiologic transition: changing patterns of mortality and population dynamics. Am J Lifestyle Med 2009; 3: 19S. DOI: 10.1177/1559827609335350.
33. Maher D, Sekajugo J. Research on health transition in Africa: time for action. Health Res Policy Syst 2011; 9: 5.
34. Maher D, Smeeth L, Sekajugo J. Health transition in Africa: practical policy proposals for primary care. Bull World Health Organ 2010; 88: 943–8.
35. Dhai A. HIV and AIDS in Africa: social, political and economic realities. Theor Med Bioeth 2008; 29: 293–6.
36. O'Brien S, Broom A. The rise and fall of HIV prevalence in Zimbabwe: the social, political and economic context. Afr J AIDS Res 2011; 10: 281–90.
37. Reher DS. The demographic transition revisited as global process. Popul Space Place 2004; 10: 19–41.
38. Mogford L. Structural determinants of child mortality in Sub-Saharan Africa: cross-national study of economic and social influences from 1970 to 1997. Biodemography Soc Biol 2004; 51: 94–120.
39. Tabutin D, Schoumaker B. La démographie de l'Afrique au sud du Sahara des années 1950 aux années 2000. Synthèse des changements et bilan statistique. Population 2004; 59: 521–621.
40. Zuberi T, Sibanda A, Bawah A, Noumbissi A. Population and African society. Ann Rev Sociol 2003; 29: 465–86.
41. Bah SM. Indirect estimation of cause of death structure in Africa and contemporary theories of mortality. Biodemography Soc Biol 1995; 42: 247–55.
42. Bah SM. Quantitative approaches to detect the fourth stage of the epidemiologic transition. Soc Biol 1995; 42: 143–8.
43. Gaylin DS, Kates J. Refocusing the lens: epidemiologic transition theory, mortality differentials, and the AIDS pandemic. Soc Sci Med 1997; 44: 609–21.
44. Mayosi BM, Flisher AJ, Lalloo UG, Sitas U, Tollman SM, Bradshaw D. The burden of non-communicable diseases in South Africa. Lancet 2009; 374: 934–47.
45. Teitelbaum M. Relevance of demographic transition theory for developing countries. Science 1975; 88: 420–5.

Citation: Glob Health Action 2014, **7**: 22443 - http://dx.doi.org/10.3402/gha.v7.22443

46. Weisz G, Olszynko-Gryn J. The theory of epidemiologic transition: the origins of a citation classic. J Hist Med Allied Sci 2010; 65: 287–326.
47. McKeown T. The modern rise of population. London: Edward Arnold; 1976.
48. Caselli G, Meslé F, Vallin J. Epidemiologic transition theory exceptions. Genus 2002; 58: 9–52.
49. MartiInez SC, Gustavo LF. Epidemiological transition: model or illusion? A look at the problem of health in Mexico. Soc Sci Med 2003; 57: 539–50.
50. Feachem RG, Jamison DT, eds. Disease and mortality in sub-Saharan Africa. Washington, DC: The World Bank; 1991.
51. Jamison DT, Feachem RG, Makgoba MW, Bos ER, Baingana FK, Hofman KJ, et al., eds. Disease and mortality in Sub-Saharan Africa. Washington, DC: The World Bank; 2006.
52. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 2006; 3: e442. DOI: 10. 1371/journal.pmed.0030442.
53. Holmes MD, Dalal S, Volmink J, Adebamowo CA, Njelekela M, et al. Non-communicable diseases in Sub-Saharan Africa: the case for cohort studies. PLoS Med 2010; 7: e1000244. DOI: 10.1371/journal.pmed.1000244.
54. Fetter B, ed. Demography from scanty evidence: Central Africa in the colonial era. Boulder: Lynne Reinner; 1990.
55. Martin CJ. Some estimates of the general age distribution, fertility and rate of natural increase of the African population of British East Africa. Popul Stud 1953; 7: 181–99.
56. Shillington K. History of Africa. New York: Palgrave Macmillan; 2005.
57. Iliffe J. Africans: the history of a continent, 2nd ed. New York: Cambridge University Press; 2007.
58. Collins RO, Burns JM. A history of Sub-Saharan Africa. New York: Cambridge University Press; 2007.
59. Page WF. Encyclopedia of African history and culture: from conquest to colonization (1500–1850). New York: Learning Source Books; 2001.
60. Ehret C. The civilizations of Africa. Charlottesville: University of Virginia; 2002.
61. Martin PM, O'Meara P. Africa, 3rd ed. Bloomington: Indiana University Press; 1995.
62. World Bank (2011). Global monitoring report 2011 – improving the odds of achieving the MDGs. Washington, DC: The World Bank.
63. United Nations, Department of Economic and Social Affairs, Population Division (2012). Changing levels and trends in mortality: the role of patterns of death by cause (United Nations publication, ST/ESA/SER.A/318). New York: United Nations.
64. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012; 380: 2095–128.
65. World Health Organization, Regional Office for Africa (AFRO) (2010). Towards reaching the health-related millennium development goals: progress report and the way forward: Report of the Regional Director. Brazaville: World Health Organization, Regional Office for Africa.
66. Kuate Defo B. The importance for the MDG4 and MDG5 of addressing reproductive health issues during the second decade of life: review and analysis from time series data of 51 African countries. Afr J Reprod Health 2011; 15: 3–21.
67. World Health Organization (2008). The world health report 2008: primary health care now more than ever. Geneva: World Health Organization.
68. Bongaarts J, Casterline J. Fertility transition: is Sub-Saharan Africa different? Popul Devel Rev 2012; 38: 153–68.
69. Bledsoe C, Banja F, Hill A. Reproductive mishaps and western contraceptive: an African challenge to fertility theory. Popul Dev Rev 1998; 24: 15–57.
70. Caldwell J, Orubuloye O, Caldwell P. Fertility decline in Africa: a new type of transition. Popul Dev Rev 1992; 18: 211–42.
71. Cooper R, Osotimehin B, Kaufman J, Forrester T. Disease burden in Sub-Saharan Africa: what should we conclude in the absence of data? Lancet 1998; 351: 208–10.
72. 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.
73. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL. Global burden of disease and risk factors. New York: Oxford University Press; 2006.
74. Salomon JA, Wang H, Freeman MK, Vos T, Flaxman AD, Lopez AD, et al. Healthy life expectancy for 187 countries, 1990–2010: a systematic analysis for the Global Burden Disease Study 2010. Lancet 2012; 380: 2144–62.
75. Thompson W. Population. Am J Sociol 1929; 34: 959–75.
76. Landry A. La révolution démographique. Paris: Sirey; 1934.
77. Popkin BM. The nutrition transition in low-income countries: an emerging crisis. Nutr Rev. 1994; 52: 285–98.
78. Popkin BM. The nutrition transition and its health implications in lower-income countries. Public Health Nutr 1998; 1: 5–21.
79. Popkin BM. Urbanization, lifestyle changes and the nutrition transition. World Dev 1999; 27: 1905–16.
80. Popkin BM, Gordon-Larsen P. The nutrition transition: worldwide obesity dynamics and their determinants. Int J Obes Relat Metab Disord 2004; 28(Suppl 3): S2–9.
81. Popkin BM. Global nutrition dynamics: the world is shifting rapidly toward a diet linked with noncommunicable diseases. Am J Clin Nutr 2006; 84: 289–98.
82. Davis K. The world demographic transition. Ann Am Acad Polit Soc Sci 1945; 237: 1–11.
83. Crenshaw EM, Christenson M, Oakey DR. Demographic transition in ecological focus. Am Soc Rev 2000; 65: 371–91.
84. Dyson T. Population and development: the demographic transition. London: Zed Books; 2010.
85. Casterline J. Demographic transition. In: Demeny P, McNicoll G, eds. The encyclopedia of population. New York: Thompson and Gale; 2003, pp. 210–16.
86. Lam D, Marteleto L. Stages of the demographic transition from a child's perspective: family size, cohort size, and children's resources. Popul Dev Rev 2008; 34: 225–52.
87. Riesman D, Glazer N, Denney R. The Lonely Crowd: a study of changing American character. New Haven, CT: Yale University; 1950.
88. Olshansky J, Ault A. The fourth stage of the epidemiologic transition: the age of delayed degenerative diseases. Milbank Q 1986; 64: 355–91.
89. Rogers RG, Hackenberg R. Extending epidemiologic transition theory: a new stage. Biodemography Soc Biol 1987; 34: 234–43.
90. Olshansky SJ, Carnes BA, Rogers RG, Smith L. Emerging infectious diseases: the fifth stage of the epidemiologic transition? World Health Stat Q 1998; 51: 207–17.
91. Olshansky SJ, Carnes BA, Rogers RG, Smith L. Infectious diseases: new and ancient threats to world health. Popul Bull 1997; 52: 1–58.
92. Lesthaeghe R. The unfolding story of the second demographic transition. Popul Dev Rev 2010; 36: 211–51.

93. Coleman D. Immigration and ethnic change in low-fertility countries: a third demographic transition. Popul Dev Rev 2006; 32: 401–46.
94. Simons J. Cultural dimensions of the mother's contribution to child survival. In: Caldwell J, Santow G, eds. Selected readings in the cultural, social and behavioural determinants of health. Canberra: Australian National University; 1989, pp. 132–45.
95. Johansson SR. Health transition. In: Demeny P, McNicoll G, eds. The encyclopedia of population. New York: Thompson and Gale; 2003, pp. 479–83.
96. Kirk D. Demographic transition theory. Popul Stud 1996; 50: 361–87.
97. Coleman D. Populations of the industrial world. A convergent demographic community. Int J Popul Geogr 2002; 8: 319–44.
98. Murray CJL, Salomon J, Mathers CD, Lopez AD. Summary measures of population health: concepts, ethics, measurement and applications. Geneva: World Health Organization; 2002.
99. Kuczynski RR. The Cameroons and Togoland. Oxford: Oxford University Press; 1939.
100. McGillivray M, Feeny S, Hermes N, Lensink R. Controversies over the impact of development aid: it works; it doesn't; it can, buy that depends. J Int Dev 2006; 18: 1031–50.
101. Young C. The end of the post-colonial state in Africa? Reflections on changing African political dynamics. Afr Aff 2004; 103: 23–49.
102. World Bank (1989). Sub-Saharan Africa: from crisis to sustainable growth. Washington, DC: World Bank.
103. Easterly W. What did structural adjustment adjust? The association of policies and growth with repeated IMF and World Bank adjustment loans. J Dev Econ 2005; 76: 1–22.
104. Van de Walle N. African economies and the politics of permanent crisis 1979–1999. Cambridge: Cambridge University Press; 2001.
105. Newbury C. States at war: confronting conflict in Africa. Afr Stud Rev 2002; 45: 1–20.
106. Sandbrook R. The politics of Africa's economic stagnation. Cambridge: Cambridge University Press; 1985.
107. Asante M. The history of Africa. London: Routledge; 2007.
108. Gordon AA, Gordon DL. Understanding contemporary Africa. Boulder: Lynne Rienner; 1996.
109. Henderson EA. The impact of culture on African coups d'État, 1960–1997. World Aff 1998; 161: 10–21.
110. McGowan P. African Military coups d'état, 1956–2001: frequency, trends and distribution. J Mod Afr Stud 2003; 41: 339–70.
111. Hruschka DJ. Culture as an explanation in population health. Ann Hum Biol 2009; 36: 235–47.
112. Institute of Medicine (2011). Preparing for the future of HIV/AIDS in Africa: a shared responsibility. Washington, DC: The National Academies Press.
113. Chaisson RE. Tuberculosis in Africa—combating an HIV-driven crisis. N Engl J Med 2008; 358: 1089–92.
114. World Health Organization (2007). Global tuberculosis control: surveillance, planning, financing. Geneva: World Health Organization.
115. Charmes J. The informal economy worldwide: trends and characteristics. J Appl Econ Res 2012; 6: 103. DOI: 10.1177/097380101200600202.
116. Hart K. Informal income opportunities and urban employment in Ghana. J Mod Afr Stud 1971; 11: 61–89.
117. International Labour Organization (1972). Employment, incomes and equality: a strategy for increasing productive employment in Kenya. Geneva: International Labour Organization.
118. International Labour Organization (2011). Statistical update on employment in the informal economy. Geneva: ILO Department of Statistics.
119. International Monetary Fund (2010). Regional economic outlook: Sub-Saharan Africa, resilience and risks. Economic and Financial Surveys. Washington, DC: International Monetary Fund.
120. Black FL. Measles. In: Evans AS, ed. Viral infections of humans: epidemiology and control, 3rd ed. New York/London: Plenum Medical Book Company; 1989, pp. 451–69.
121. Miller MA, Sentz JT. Vaccine-preventable diseases. In: Jamison DT, Feachem RG, Makgoba MW, Bos ER, Baingana FK, Hofman KJ, et al., eds. Disease and mortality in Sub-Saharan Africa. Washington, DC: The World Bank; 2006, pp. 163–77.
122. World Health Organization. Global routine vaccination coverage, 2010. Wkly Epidemiol Rec 2011; 46: 509–20.
123. WHO/UNICEF (2012). Immunization summary. Geneva: World Health Organization.
124. Prince M, Patel V, Saxena S, Maj M, Maselko J, Phillips MR, et al. No health without mental health. Lancet 2007; 370: 859–77.
125. National Research Council (1993). The epidemiological transition: policy implications for developing countries. Washington, DC: National Academy Press.
126. World Health Organization (2006). Working together for health. The World Health Report 2006. Geneva: World Health Organization.
127. Institute of Medicine (2010). Mental, neurological, and substance use disorders in Sub-Saharan Africa: reducing the treatment gap, increasing quality of care: workshop summary. Washington, DC: The National Academies Press.
128. Institute of Medicine (2010). Promoting cardiovascular health in the developing world: a critical challenge to achieve global health. Washington, DC: The National Academies Press.
129. Awases M, Gbary A, Nyoni J, Chatora R. Migration of health professionals in six countries: a synthesis report. Brazzaville: World Health Organization; 2004.
130. Schubert C. Nurses disappearing from developing nations. Nat Med 2003; 9: 979.
131. Dovlo D. Migration of nurses from Sub-Saharan Africa: a review of issues and challenges. Health Serv Res 2007; 42: 3. DOI: 10.1111/j.1475-6773.2007.00712.x.
132. Inhorn MC, van Balen F, eds. Infertility around the world: new thinking on childlessness, gender, and reproductive technologies. Berkeley: The University of California Press; 2002.
133. Caraël M, Glynn JR, eds. HIV, resurgent infections and population change in Africa. New York: Springer; 2008.
134. Stuckler D, Basu S, McKee M. Drivers of inequality in millennium development goal progress: a statistical analysis. PLoS Med 2010; 7: e1000241. DOI: 10.1371/journal.pmed.1000241.
135. Stuckler D. Population causes and consequences of leading chronic diseases: a comparative analysis of prevailing explanations. Milbank Q 2008; 86: 273–326.
136. Blum RW. Youth in Sub-Saharan Africa. J Adolesc Health 2007; 41: 230–8.
137. Omran AR, Roudi F. The Middle East population puzzle. Popul Bull 1993; 48: 1–40.
138. Institute of Medicine (1996). In her lifetime: female morbidity and mortality in Sub-Saharan Africa. Washington, DC: National Academy Press.
139. Rao C, Lopez AD, Hemed Y. Causes of death. In: Jamison DT, Feachem RG, Makgoba MW, Bos ER, Baingana FK, Hofman KJ, et al., eds. Disease and mortality in Sub-Saharan Africa. Washington, DC: The World Bank; 2006, pp. 43–58.

140. Young F, Critchley JA, Johnstone LK, Unwin NC. A review of co-morbidity between infectious and chronic disease in Sub Saharan Africa: TB and diabetes mellitus, HIV and metabolic syndrome, and the impact of globalization. Global Health 2009; 5: 9. DOI: 10.1186/1744-8603-5-9.
141. Menéndez C, Romagosa C, Ismail MR, Carrilho C, Saute F, Osman FN, et al. An autopsy study of maternal mortality in Mozambique: the contribution of infectious diseases. PLoS Med 2008; 5: e44. DOI: 10.1371/journal.pmed.0050044.
142. Magruder JR. Marital shopping and epidemic AIDS. Demography 2011; 48: 1401–28.
143. Clark S, Poulin M, Kohler HP. Marital aspirations, sexual behaviors, and HIV/AIDS in rural Malawi. J Marriage Fam 2009; 71: 396–416.
144. Pictet G, Le Coeur S., M'Pele P, Brouard N, Lallemant M. Contribution of AIDS to the general mortality in central Africa: evidence from a morgue-based study in Brazzaville, Congo. AIDS. 1998; 12: 2217–23.
145. McIntyre J. Maternal health and HIV. Reprod Health Matters 2005; 13: 129–35.
146. Preston SH. Mortality patterns in national populations. New York: Academic Press; 1976.
147. Vallin J. De la mondialisation de la transition au retour des incertitudes (1940–2000). In: Caselli G, Vallin J, Wunsch G, eds. Démographie: analyse et synthèse. Volume V, Histoire du peuplement et prévisions. Paris: Institut National d'Études Démographiques; 2004, pp. 117–70.
148. Cleland J. Mortality–fertility relationships. In: Demeny P, McNicoll G, eds. The encyclopedia of population (Vol. 2). New York: Macmillan Reference USA; 2003, pp. 668–72.
149. Gaisie SK. Demographic transition: the predicament of Sub-Saharan Africa. In: Douglas RM, Jones G, D'Souza RM, eds. The shaping of fertility and mortality declines: the contemporary demographic transition. Health Transit Rev 1996; 6(Suppl): 345–69.

Global Health Action

PART II

The development and experience of epidemiological transition theory over four decades: a systematic review

Ailiana Santosa[1]*, Stig Wall[1], Edward Fottrell[2], Ulf Högberg[3] and Peter Byass[1,4]

[1]Division of Epidemiology and Global Health, Department of Public Health and Clinical Medicine, Umeå Centre for Global Health Research, Umeå University, Umeå, Sweden; [2]Institute for Global Health, University College London, London, UK; [3]Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden; [4]MRC/Wits Rural Public Health and Health Transitions Research Unit, School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Background: Epidemiological transition (ET) theory, first postulated in 1971, has developed alongside changes in population structures over time. However, understandings of mortality transitions and associated epidemiological changes remain poorly defined for public health practitioners. Here, we review the concept and development of ET theory, contextualising this in empirical evidence, which variously supports and contradicts the original theoretical propositions.

Design: A Medline literature search covering publications over four decades, from 1971 to 2013, was conducted. Studies were included if they assessed human populations, were original articles, focused on mortality and health or demographic or ET and were in English. The reference lists of the selected articles were checked for additional sources.

Results: We found that there were changes in emphasis in the research field over the four decades. There was an increasing tendency to study wide-ranging aspects of the determinants of mortality, including risk factors, lifestyle changes, socio-economics, and macro factors such as climate change. Research on ET has focused increasingly on low- and middle-income countries rather than industrialised countries, despite its origins in industrialised countries. Countries have experienced different levels of progress in ET in terms of time, pace, and underlying mechanisms. Elements of ET are described for many countries, but observed transitions have not always followed pathways described in the original theory.

Conclusions: The classic ET theory largely neglected the critical role of social determinants, being largely a theoretical generalisation of mortality experience in some countries. This review shows increasing interest in ET all over the world but only partial concordance between established theory and empirical evidence. Empirical evidence suggests that some unconsidered aspects of social determinants contributed to deviations from classic theoretical pathways. A better-constructed, revised ET theory, with a stronger basis in evidence, is needed.

Keywords: *epidemiological transition*; *demographic transition*; *mortality*; *social determinants*

Responsible Editor: Barthélémy Kuate Defo, University of Montreal, Canada.

*Correspondence to: Ailiana Santosa, Division of Epidemiology and Global Health, Department of Public Health and Clinical Medicine, Umeå University, SE-901 87 Umeå, Sweden, Email: ailiana.santosa@epiph.umu.se

This paper is part of the Special Issue: *Epidemiological Transitions – Beyond Omran's Theory*. More papers from this issue can be found at http://www.globalhealthaction.net

Received: 13 December 2013; Revised: 12 March 2014; Accepted: 29 March 2014; Published: 15 May 2014

Changes in mortality patterns and increases in life expectancy, with subsequent impacts on population, have been documented in industrialised countries since the 19th century. Early studies on population change over time were overviewed from a demographic rather than epidemiological perspective (1), including Thompson's early work observing changes in fertility and mortality rates in populations (2). Landry introduced the term 'demographic transition' in describing secular changes in fertility and mortality in 1934, later reprinted in English (3). This idea was further developed in association with socio-economic development (4). In 1971, Omran proposed a theory of 'epidemiological transition (ET)', which grew out of the demographic transition model and incorporated more detailed consideration of particular diseases as causes of death. He particularly based this on mortality changes in England, Wales, Japan, and Sweden during the 19th century (5).

Citation: Glob Health Action 2014, **7**: 23574 - http://dx.doi.org/10.3402/gha.v7.23574

This ET theory, in five propositions, describes changing population patterns in terms of fertility, life expectancy, mortality, and leading causes of death (5). The first proposition states that mortality is an important aspect of population dynamics. The second proposition describes changes in disease and mortality patterns, as 'pandemics of infectious disease are gradually shifted by degenerative and man-made diseases as the leading cause of morbidity and main cause of death'. The third proposition explains that children and young women experience the most profound impacts of ET, resulting in declining infant and maternal mortality and reduced fertility rates. The fourth proposition links long-term population changes in health and disease patterns to demographic, economic, and social determinants and mortality changes. The final proposition outlines three basic variants of ET that are functions of 'peculiar variations in the pattern, the pace, the determinants and the consequences of population change'.

Omran proposed three stages of transition as underlying the changes in patterns of mortality and morbidity. The first stage, 'the age of pestilence and famine', is characterised by high and fluctuating mortality due to epidemics, famines and war, and poor living conditions. In this stage, a combination of high crude death rate, high fertility rate, and low life expectancy at birth (between 20 and 40 years) results in slow population growth. The most common causes of death are infectious and parasitic diseases, especially among children and women of childbearing age. The second stage, 'the age of receding pandemics', witnesses declining mortality rates, initially high but later decreasing fertility, and life expectancy at birth increasing to around 55 years. The major driving forces in this stage of transition are sanitation improvements, control of major outbreaks of infectious diseases, and medical breakthroughs (including contraception). While infectious diseases remain as major causes of death, non-communicable diseases (NCDs) start to increase steadily. The third stage, 'the age of degenerative and man-made disease', is characterised by decreasing and relatively stable low mortality and increasing life expectancy at birth to over 70 years, manifesting in a population that is ageing. In this stage, NCDs dominate causes of death, with many deaths attributable to cardiac and cerebrovascular ailments, chronic lung and metabolic diseases, cancers, injuries, and stress-related disorders (5).

In 1983, Omran recognised the need to update his theory to incorporate a more extended description of the transition, as emerging analyses of transition patterns based on historical data did not fit the original model (6). Omran later acknowledged the presence of one and possibly two additional stages to his original theory of ET. He added the fourth stage as 'the age of declining cerebrovascular mortality, ageing, lifestyle modifications and resurgent diseases', during which life expectancy continues to increase (up to 80–85 years), and the mortality attributed to cardiovascular diseases declines and stabilises as a result of improved medical care and lifestyle modifications. Omran's fifth stage was characterised by the emergence of new diseases (HIV/AIDS, hepatitis) and re-emergence of old diseases (cholera, malaria, diphtheria, tuberculosis, plague) (7), which were already being described by others (8, 9). In his original fifth proposition, Omran proposed three basic variants of transition, but later added an additional model, similar to the classic model but starting several decades later and passing faster through the different stages of the transition (7).

Critiques on Omran

The applicability and universality of ET theory across various places and contexts remain contentious (10–22). Criticisms of ET theory as over-simplistic peaked in the 1990s on the basis that it failed to understand the comprehensive nature and historical sequence of mortality transitions (10, 14–17, 22). A major critique of Omran's theory is the assumption that all countries will experience similar linear progression of transitions with respect to onset and speed. However, not all countries necessarily encounter ET in the same way. Omran treated entire populations as undifferentiated units; his conclusions drawn from the mortality statistics of Sweden, England, and Wales have been considered contestable; that the theory 'fails to grasp the global nature and the historical sequence of the mortality transition as it spread', and that it is 'insufficiently epidemiological in that its focus was the changing causes of death rather than the changing causes of patterns of illness' (23). Mackenbach argued that the ET theory is ambiguous because it was developed based on Western data and it is difficult to ascertain the beginning and end of the transition (20). In addition, Frenk et al. (18) and Smallman-Raynor and Phillips (19) challenged the assumptions of ET theory's unidirectional structure and continuous development, introducing the concepts of counter transition and epidemiological polarisation. It was also suggested that ET was part of a broader effort to reorient American and international health institutions towards the pervasive population control agenda of the 1960s and 1970s rather than focusing on the increasing burden of chronic disease (13).

The generalisability of ET theory has been doubted, based on the great variations in mortality trends among population subgroups (14, 21). Ruzicka and Kane examined inequalities in mortality and concluded that mortality patterns vary widely by race, sex, economic indicators, and class, resulting in substantial 'heterogeneity within social class and within any other socio-cultural, demographic or economic category'. They criticised the ET theory for the assumption that communities will gradually progress to the point where they have virtually eliminated infectious diseases as a major health threat (15).

Gaylin and Kates determined the importance of morbidity and mortality differences between population subgroups, using the HIV/AIDS pandemic as a case study, and showed important inconsistencies with the optimistic, equitable trends implied by ET theory, suggesting that the modern picture may be more complex than the original theory could predict (21).

Carolina and Gustavo examined the validity of ET theory as an effective model in the interpretation of mortality and morbidity changes, with reference to Mexico and to low- and middle-income countries (LMICs) in general. They found that the main theoretical problem in using the ET theory related to a preference for phenomenological descriptions rather than theoretical explanations of the causal patterns of death and disease and their links with the changes experienced in societies. Interpretations using the ET theory were inevitably based on scientific and social perspectives frozen at the time of its use (17). The paradox was that researchers frequently misunderstood ET theory as representing more or less the same concept as Omran's description of demographic transition, casting it as a theory about changing disease conditions progressing everywhere in a uniform and linear manner. Despite his reliance on its main concepts, Omran explicitly rejected demographic transition as a theoretical framework, postulating that ET was formulated in an attempt to provide a more comprehensive approach to the dynamics of the mortality–fertility transition. In his view, recent mortality declines in the developing world depended not on economic development but on national and international programs of health service provision and environmental control. However, this viewpoint contrasted with the thesis of McKeown et al. that broad population shifts in disease occurrence during transitions to industrialised societies, from declines in infectious disease to increases in NCDs, were due to improved nutrition and increased exposure to 'conditions for which we are genetically ill-equipped' (24). Preston pointed out that mortality is not associated with economic growth and the theory becomes weaker at very low levels of income (25).

In addressing the conceptual drawbacks of the theory, the term 'health transition' was coined to incorporate new elements into Omran's theoretical assumptions (10, 12). Health transition was described as 'a dynamic process whereby the health and disease patterns of a society evolve in diverse ways as a response to broader demographic, socio-economic, technological, political, cultural and biological changes', and divided into ET (changes in health patterns) and health care transition (the organised response to health conditions). Omran, however, argued that the health transition is part of the ET, not vice versa (6). The concept of 'health transition' was proposed as a wider framework that included not only epidemiological characteristics but also the ways in which societies responded to changing health situations as a result of cultural, social, and behavioural determinants (22, 23, 26, 27).

Clear understanding of mortality transition and its implications is still hampered due to a lack of evidence from LMICs (28). The lack of quality mortality data in many parts of the world makes it difficult to understand the generalisability of the theory globally, as well as its interpretation. Changes in disease classifications over time also limit the comparability of available data for assembling a comprehensive pattern of mortality transition (29). The theory itself received relatively little attention before the global incidence of NCDs increased in the 1990s. An overview of how the ET theory has been applied since its conception, and the identification of gaps where it fails, are warranted.

Objectives

This paper intends to synthesise published evidence on mortality transition, and, if possible, assess how ET theory has been applied in understanding the transition in specific contexts. More specifically, this paper aims to answer the following questions: 1) What evidence on mortality transition is available, who are the beneficiaries during the transitions, and what are the social–economic determinants that coexist with the ET?; 2) What existing evidence is available to illustrate changing patterns of causes of death?; and 3) What deviations from the classic ET theory have been observed, and do these reveal emerging patterns?

Methods

We conducted a systematic review in PubMed using the keywords 'epidemiological transition (s) or epidemiologic transition(s) or demographic transition(s) or health transition(s)' and 'mortality'. We selected these keywords to cover a wide range of transitions, including health and demographic transitions that are related to ET. We included only articles on human research, and published in English between 1 January 1971 and 31 December 2013.

We obtained 547 articles, which were later screened by reading their titles and abstracts. From this step, a total of 324 articles were excluded (210 irrelevant papers and 114 reviews/commentaries/editorials). Review papers including Omran's and Caldwell's conceptual papers were excluded but used to provide framework in the 'Discussion section'. The full texts of remaining 223 articles were searched and read through. From this step, we obtained 16 additional articles, not found in the original search, from the reference lists of relevant studies and review articles. We could not obtain full text for 29 papers. Therefore, only a total of 210 articles were included for full text review in the second step. All citations were saved in the PubMed database and imported into the EndNote X6 database. All full-text articles were reviewed by two of the authors for inclusion in the study. Uncertainties over

Citation: Glob Health Action 2014, **7**: 23574 - http://dx.doi.org/10.3402/gha.v7.23574

study inclusion were discussed between the researchers and resolved through consensus. Another 74 papers were further excluded after reading the full text, mainly because of insufficient mortality data ($n = 23$), observational studies assessing risk factors and outcomes in a defined population ($n = 17$), and papers not directly relevant to this study ($n = 34$). A final 136 articles were included in the review. Details of this literature search are summarised in Fig. 1.

Results

Of the 136 articles, 112 articles were performed in single countries, and the other 24 articles were conducted in multiple countries in which some of multi-countries studies used the Global Burden of Disease models. Seventy-five out of the 136 articles (55%) used individual-level data, and the remaining 45% used ecological data. Fifty papers (37%) used historical data before Omran's theory was postulated in 1970, and the remaining 63% focused more on contemporary society after 1970. Seventy-nine of the articles (58%) reported time trends in outcome indicators, while the remaining 42% reported cross-sectional observations.

The main outcome indicator used in most studies was mortality rate (either total, sex specific, age specific, and/or cause specific), which was used in 117 articles (87%). The remaining 13% articles used absolute number or proportion of deaths (9%), DALY (3%), life expectancy and standardised mortality ratio (1%). On disease outcomes reported, 49% of the papers examined both all and cause-specific deaths, 27% only focused on all-cause deaths, and 24% reported cause-specific deaths.

With regard to the research questions, 50 papers (37%) assessed trends in outcome indicators, with 37 additional papers (27%) analysing both the trends and their determinants. Twenty-three papers (17%) reported the prevalence of the outcome indicators, and 21 additional papers (15%) reported both the prevalence and their determinants. Five papers (4%) reported both the prevalence and the trends of the outcome indicators. Twenty-nine papers (21%) reported results from studies in rural areas, 1% in urban areas, 16 papers (12%) in both urban and rural areas, and 90 papers (66%) did not characterise the study area. A total of 112 papers (82%) reported results from studies in general populations, 16 papers (12%) on children under 15 years, and eight papers (6%) on adults aged 15 years and above (Table 1).

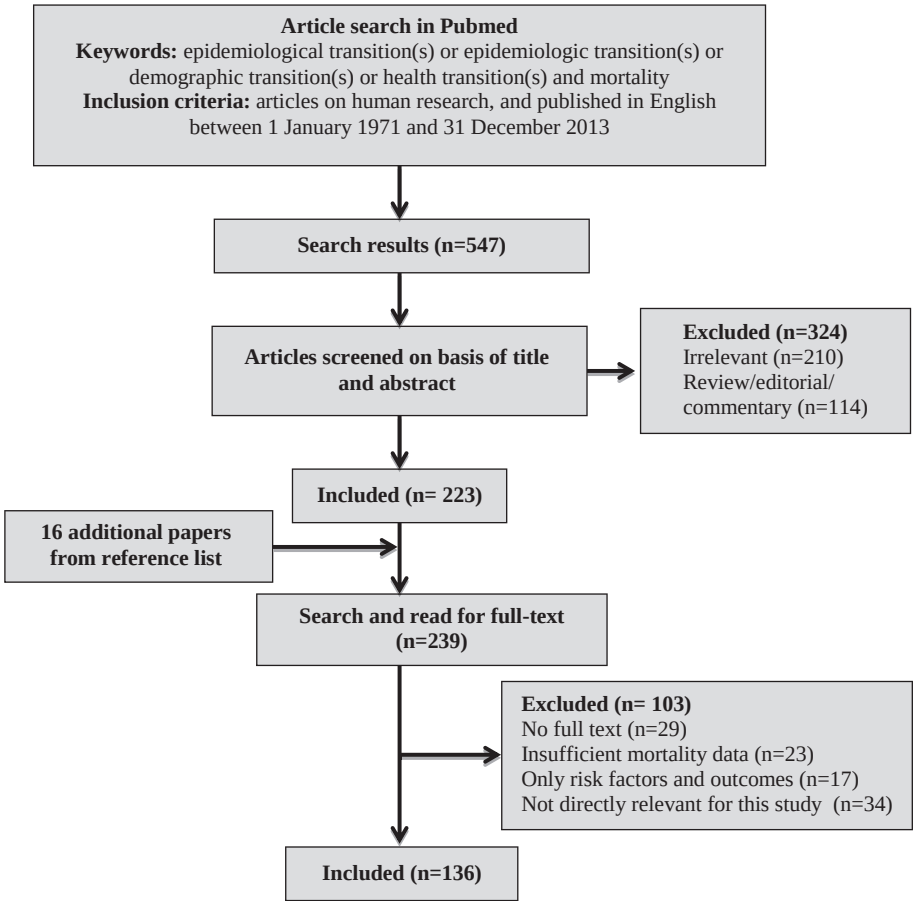

Fig. 1. Flow diagram of the literature search process.

Table 1. Regions and countries from where the papers included in this review originated ($n = 136$ papers)

	Papers
Africa	Ethiopia (30, 31), Ghana (32), Malawi (33), Mauritius (34), Morocco (35), Seychelles (36, 37), South Africa (38–44), Sub-Saharan Africa (45), Tanzania (46)
America – North	Canada (47–52), Cuba (53, 54), Costa Rica (55), Haiti (56), Jamaica (57), Mexico (17, 58–63), Trinidad and Tobago (64, 65), USA (66–68)
America – South	Brazil (69–74), Chile (75), Peru (76, 77), Colombia (78)
Asia	China (79–83), Hong Kong (84), India (25, 85–89), Israel (90), Japan (91, 92), Korea (93–95), Palestine (96), Russian (97), Singapore (98), Sri Lanka (99), Thailand (100, 101), Turkey (102), Vietnam (103, 104)
Europe	Bolivia (105), Bulgaria (106), Czech Republic (107), France (108), Hungary (109), Lithuania (110, 111), The Netherlands (112, 113), Spain (114–119), Sweden (120–123), United Kingdom (124, 125)
Pacific	Australia (126, 127), Fiji (128), French Polynesia (129, 130), Nauru (131, 132), New Zealand (133, 134)
Multi-countries	WHO databases, mainly from World Health Statistics and Global Burden of Disease database (135–150), national data (151–156), Human mortality database (157–159), Demographic Health Surveillance System in INDEPTH Network (160, 161), other databases (147, 162–164)

Discussion

Omran's theoretical concept of ET has broadly been used in public health because it offers explanations for changing cause-of-death patterns corresponding to different stages of transition. Most studies reviewed in this paper described the theory in relation to specific situations in particular countries, rather than considering its generalisability. Our review attempts to overview published evidence over four decades on mortality transition in relation to the ET theory. We structure the discussion in several sections to discuss how the empirical research articles illustrate the overall and cause-specific mortality transitions, and whether deviation from the ET theory and emerging patterns have been observed. A comprehensive exploration of health and ET in each country included in this review, as well as a thorough review of historical demographic research that led to Omran's theory, is beyond the scope of this paper.

What evidence on overall mortality transition exists?

Many studies have confirmed Omran's first proposition on 'mortality as a fundamental factor in population dynamics'. In many countries, mortality in the population decreased significantly in the past century, alongside decreasing fertility, increasing life expectancy, and a growing number of older population. In this section, we review papers, which discuss how transition of mortality has occurred and how it influenced the population structure over a period of time.

Mortality decline has been observed with different speeds of transition in the population among all age groups (17, 32, 34, 53, 55, 61, 65, 67, 75, 86, 91, 98, 106, 114, 116, 123, 124, 145) and across specific age groups, including neonates and infants (50, 99, 117), children (35, 56), adolescents (142), and adults (73, 85, 127). In some countries – such as Spain, Japan, Mexico, and Sri Lanka – good quality vital registration allows researchers to study mortality transition over long time periods. The mortality rate in Spain increased during the 19th century and peaked at the end of the century, after which the rate decreased at a rate of 2.3 deaths per 1,000 population per decade during the 20th century (114, 116). During the 20th century, many countries experienced quite similar reductions in death rates with a reduction between 2 and 3 deaths per 1,000 population in a decade, such as in Japan (91) and Mexico (17). These data illustrate how the speed of mortality transition has become faster over time.

Decreasing fertility and birth rates led to an increase in life expectancy and a change in the population structures (35, 75, 114). While the mortality rate in Spain started decreasing by the end of the 19th century, the birth rate did not start declining until 1930. Since then, it has decreased by two-thirds and reached the level of 9.75 per 1,000 population by the end of the 1980s. The proportion of population dying at 70 years or older changed from 25% in the beginning of the 19th century to 74% by the end of the 20th century (114). Chile experienced a faster transition, with the proportion of deaths among older people over 65 years doubling from 35 to 60% during a 22-year period (1970–1992) (75). In Monaco, the fertility rate halved within a 30-year period (1980–2011), with 29% of the population being children under 15 years and 7% older people aged 65 + (35).

In contrast to the general pattern of decreasing mortality during the 20th century, populations in several settings experienced stagnation of mortality reductions, such as Eastern European countries (147), and even reversal of mortality, such as in Nauru (131), Philadelphia in the USA (66), Russia (152), South Africa (38–40), Tanzania (46), and Thailand (101). A high rate of tuberculosis mortality among the adult population, which slowed down the increase in life expectancy, was observed in Philadelphia during the beginning of the 20th century (66). In Nauru, the age-standardised mortality rate among adult men doubled during 1960–1981, mainly due to the high

Citation: Glob Health Action 2014, **7**: 23574 - http://dx.doi.org/10.3402/gha.v7.23574

burden of chronic NCDs (131). In rural South Africa, Kahn et al. demonstrated significant increases in mortality for men and women, mainly due to the HIV epidemic, during the last decade of the 20th century, with life expectancy decreasing rapidly by 12 years in females and 14 years in males in a 10-year period (38). The overall mortality rate increased by 87% from 1992 to 2005 (39). In Thailand, the generalised HIV epidemic since the 1990s has led to an increase in male mortality rate and a decreased life expectancy by 4 years during 1990–2000 (101).

Some studies have been conducted in indigenous populations in Bolivia (105), Canada (50, 52), Mexico (61), New Zealand (133), and Sweden (120, 121). Indigenous populations tend to have higher birth rates, mortality rates, and lower life expectancy than the general population, as has been observed among sub-Arctic Indians in Canada (50) and the aboriginal Maori population in New Zealand (133). The Zapotec-speaking genetic isolate in the Valley of Oaxaca, southern Mexico, experienced mortality crises during 1930–1950 with the measles epidemic, resulting in deaths exceeding births. After 1955, birth and death rates diverged in a pattern typical of rapid population growth in early Stage II of the demographic transition (61). Recent evidence has shown that health among indigenous populations has improved over time, as has been observed among the Sami population in Sweden by the end of 19th century (120), and among the Tsimane Amerindian population in Bolivia during the second half of the 20th century (105). In Sweden, the gaps in infant mortality rate (IMR) between the Sami and non-Sami populations became narrower over time. Despite these improvements, health inequality persisted between the south Sami and the north Sami (121).

Evidence of transition among children and young women

In his third proposition, Omran stated that 'During the epidemiologic transition the most profound changes in health and disease patterns obtain among children and young women'. Evidence from many countries in this review supports this claim (17, 32, 35, 53, 56, 58, 64, 66, 67, 72, 75, 80, 86, 95, 102, 107, 114, 116, 124, 128, 142, 146, 158). For example, the IMR and perinatal mortality rate decreased steadily in Chile during 1970–1992, while the neonatal mortality rate has been stable since the 1970s (75). During 1992–2011, the mortality rate among children in Morocco decreased by more than 40% over a 20-year period (a decrease of 44% in neonatal mortality rate, 54% in IMR, 64% in under-five mortality rate, and 66% in postnatal mortality rate) (35). Sastry reported significant decrease in under-five mortality from 117 per 1,000 live births in 1970 to less than 50 by 1991 in Brazil (72).

In some countries, however, sustainable health improvements among children and young women were not observed during the transition. Infant mortality among Tsimane population in Bolivia did not decrease significantly concomitantly with large reductions in adult and old age mortality during 1950–2000 (105). Child mortality in Nauru fluctuated and increased over time from 6 to 10.8 per 1,000 population during 1960–2007 (131). Unlike other countries in South East Asia that experienced a decline in infant mortality during 1960–1980, Thailand still struggled with high infant mortality of 40–45 per 1,000 live births in the 1980s (154). Kahn et al. showed that children (0–4 years) and young adults (20–49 years) in rural South Africa experienced rapid increases (two and five times, respectively) in mortality rates over a 10-year period due to the HIV/AIDS epidemic (38, 40). In South Africa, the HIV/AIDS epidemic led to the disease becoming the leading cause of death among infants and young adults by 2000. Women suffered more from HIV/AIDS and chronic diseases than men (42).

What evidence is available to illustrate changing patterns of causes of death?

In his second proposition, Omran postulated 'During the transition, a long-term shift occurs in mortality and disease patterns'. Omran suggested a shift in the causes of death from predominating infectious diseases to NCDs. Therefore, we tried to identify how countries included in this review stood in terms of ET stages over time and later discuss specific transitions that did not follow the patterns proposed by Omran.

Table 2 shows a summary of countries at various ET stages over time. Industrialised countries generally started ET earlier but proceeded slowly. Western European countries such as Sweden, United Kingdom, and France took most of the 19th century to shift from Stage 2 to Stage 3 and then a further half-century from Stage 3 to Stage 4 (108, 122, 124, 165). Cuba, USA, and Australia entered Stage 2 later than many European countries (53, 54, 66, 67, 126) but reached Stage 3 almost at the same time. The Netherlands (113) faced Stage 1, 'the age of famine and pestilence', when USA and Australia were encountering Stage 2. Hungary moved rapidly from Stage 2 to Stage 3 at the end of the 20th century, after the rapid political changes, and is now in Stage 4 of ET with declining cardiovascular mortality and increasing life expectancy (109). The same stage, with declining cardiovascular mortality, was observed in Canada during 1978–1996 (51). In the former Soviet Union, ET did not progress homogeneously across different geographical regions in the country. The male population in the northern part experienced higher mortality due to chronic diseases, such as cardiovascular diseases, injuries, and lung cancer (166).

Spain entered Stage 1 at the beginning of the 19th century (116, 117). The number of deaths have fluctuated and have slightly increased since the beginning of the 19th century in Spain, with several severe mortality crises

Table 2. Timeline of selected countries reported to be at various stages of epidemiological transition

	Stage 1	Stage 2	Stage 3	Stage 4
1870–1874	United Kingdom (153, 157), Sweden (153, 157), Finland (153, 157), Iceland (153, 157), USA (66), Germany (167)			
1875–1879	The Netherlands (113)			
1880–1884				
1885–1889				
1890–1894				
1895–1899				
1900–1904	Spain (114, 116, 117)			
1905–1909				
1910–1914		United Kingdom (124, 153, 157), Sweden (153, 157), Finland (153, 157), Iceland (153, 157), The Netherlands (113)		
1915–1919				
1920–1924	Mexico (17)	Spain (114, 168)		
1925–1929				
1930–1934		Germany (167)		
1935–1939	Japan (91)			
1940–1944				
1945–1949		Costa Rica (55)		
1950–1954	Hong Kong (154), Trinidad and Tobago (64, 65)	Japan (91, 158), Ghana (32), Canada (47, 169)	The Netherlands (113), United Kingdom (124, 153, 157), Sweden (153, 157), Iceland (153, 157)	
1955–1959		China (158)		
1960–1964			Cuba (140), Puerto Rica (140)	
1965–1969		Hong Kong (154), Singapore (154)	Spain (114, 116, 117), Costa Rica (55)	
1970–1974	Peru (77)	Trinidad and Tobago (64, 65)	Japan (158), Mauritius (34), Canada (169)	Netherland (113)
1975–1979		Korea (95)		
1980–1984				
1985–1989		Mexico (59, 140)	Hong Kong (154) Malaysia (154)	Costa Rica (55), Canada (47), Spain (168)
1990–1994		Peru (77)	Seychelles (37), Ghana (32), Cuba (54)	Japan (158), France (108), Australia (126, 127), Trinidad and Tobago (65)
1995–1999			Mexico (17), Korea (93)	China (158)
2000–2004			India (85, 86), Thailand (100, 101)	Mexico (60), Peru (76), Hungary (109), Brazil (69, 73, 118)
2005–2009				Vietnam (104), Mexico (170)

including the cholera epidemic in 1855, the influenza pandemic in 1918, and the nutritional deficiency problem during the Spanish Civil War in 1941. At the beginning of the 20th century, circulatory mortality replaced infectious diseases as the leading cause of death (114). By the end of the century, circulatory disease was responsible for 75% of total deaths and less than 5% deaths were related to infectious disease, a characteristic of Stage 3 (114).

Citation: Glob Health Action 2014, **7**: 23574 - http://dx.doi.org/10.3402/gha.v7.23574

The fourth stage, which is characterised by declining circulatory disease, increasing degenerative diseases, such as malignant neoplasms and metabolic disease, and a stable, low proportion of infectious disease, was observed in Spain at the beginning of the 21st century (119).

Up to the mid-20th century, average life expectancies in LMICs were much lower than those of industrialised countries, largely driven down by premature infectious disease mortality. After World War II, most countries progressed faster and followed a general trend of converging and increasing life expectancy. Most LMICs entered Stage 2 later than high-income countries (in the second half of the 20th century), but they subsequently moved on to the next stages of ET more rapidly than that of the high-income countries. This faster transition was observed in Brazil (70, 74), Chile (75), China (79, 80, 155), Japan (91), Korea (93), Mexico (60), Peru (76, 77, 171), and Vietnam (104), perhaps because of greater economic growth and improvements in health. By the end of the 20th century, the proportion of ailments related to circulatory disease had increased to 75% in Chile (75). Diabetes mortality in Mexico increased slightly during 1980–2000 (62). By 2004, Mexico had progressed further in the ET, with NCD and injury as a major burden of disease [75% of all deaths in 2004 were due to NCDs, mainly due to ischaemic heart diseases (13%), diabetes mellitus (10%), cerebrovascular diseases (6%), liver cirrhosis (6%), and road traffic accidents (4%)] (60).

A variant of the non-Western transition model has been formulated as the 'protracted polarised transition model', in order to account for the process of epidemiological polarisation. Frenk et al. conceptualised this extended variant of the ET theory based on Latin American data. Latin America had experienced a resurgence of malaria and dengue fever, along with a burden of infectious diseases and NCDs across regions and social classes, with widening health gaps across regions and social classes due to an unequal distribution of health interventions. The epidemiological polarisation element is characterised by the overlap of eras, a persistence of infectious diseases combined with the emergence of NCDs and new epidemic diseases, capturing also increased geographical and social health inequalities. They also showed several main features of the protracted–polarised model; that the period of mortality decline is short in contrast with the classical model; the onset of mortality decline does not begin before the 20th century; infectious diseases are not yet completely controlled, with consequent overlaps of stages and unequal distributions of wealth and health; and re-emergence of old diseases that did not feature in the original ET theory (18).

Evidence from some countries confirmed the protracted epidemiological polarisation (32, 38). Agyei-Mensah and de-Graft Aikins described how Ghana had to deal simultaneously with persistent infectious diseases predominantly among poor people, increasing NCDs predominantly among wealthy people, and the emergence of the HIV/AIDS epidemic (32). Kahn et al. (38) examined trends in age-specific mortality in a rural South African population and reported a 'counter transition' of increasing mortality among children and young adults due to acute diarrhoea and malnutrition, 'epidemiological polarisation' from a greater mortality burden among disadvantaged groups, and a 'protracted transition' with a coexistence of HIV/AIDS and chronic NCDs in older adults (38, 39, 41). Peru also faces a double burden of persistent communicable diseases and NCDs, while dealing with injuries, re-emerging infections (TB, malaria), and HIV/AIDS. The country also suffers from inequalities in health and wealth within the country. This situation created an overlap of stages, broadening the gaps in health status among social classes and geographical regions (76, 77, 171).

Determinants of ET

Omran postulated that 'The shifts in health and disease patterns that characterize the epidemiologic transition are closely associated with the demographic and socio-economic transition that constitute the modernization complex'. Demographic and socio-economic determinants, including sex, income level, education level, marital status, ethnicity, regional differences, as well as wider structural and environmental factors, can influence the transition process and lead to unequal health outcomes. In this review, we observed sex differentials in mortality over time (46, 49, 51, 73, 80, 81, 86, 94, 106, 118, 123, 132, 157); regional disparities in mortality (57, 60, 62, 72, 74, 76, 77, 85, 102, 146, 148, 172); social, economic, and mortality disparities across different ethnic groups (48, 49, 90, 98, 121, 125, 156); and climatic factors as determinants of mortality transition (115, 124, 153).

Although demographic and socio-economic transitions are significantly related to health improvement, some health problems still exist in some countries because the effects of socio-economic development and globalisation can be detrimental. In China, the economic reform that improved people's living environment (i.e. income and health care) had net effects on health. Conversely, economic reform in Russia did not produce any tangible economic and social benefits, and therefore health status was stagnant or deteriorating. In Russia, morbidity and mortality related to infectious disease increased after the reform, which could be explained by a relaxed or deteriorating public health network (152). Despite the improvement of socio-economic status in Chile during 1970–2000, Chile still experienced a persistent burden of tuberculosis, typhoid fever, and hepatitis, and at the

same time, an increasing burden of HIV/AIDS (75). The association between unintentional injuries and economic development differed by age. In population groups aged 45–74 years, unintentional injuries were negatively associated with GNP per capita, reflecting lower rates of injuries in countries with high GDP. Among the group aged 75 years and above, the association was positive with a higher rate of unintentional injuries in countries with high GDP (137). The same pattern was observed for intentional injuries such as suicide (150).

What deviations from the classic ET model have been observed, and do these reveal emerging patterns?

The process of mortality transition can be complex and dynamic, often influenced by demographic, socio-economic, technological, cultural, environmental, and biological drivers of change. We found numerous instances of deviations from the classic ET theory when examining specific geopolitical areas. A general assumption that current LMICs will experience the same transitional patterns as those countries already industrialised, following increased economic growth, is not proven. Some countries have encountered serious obstacles preventing them from completing certain stages of the transition, often tied to their history, economic development, or culture. Vallin and Mesle, referring to Eastern Europe and Africa, challenged the theory and proposed the idea that life expectancies are rapidly converging towards a maximum level (173). Having achieved relatively high life expectancies during the 1970s, a number of central European and former Soviet republics experienced periods of stagnation due to increasing cardiovascular mortality, violence, and alcoholism, with declines in life expectancy following the collapse of the USSR (97, 147, 174, 175). The combination of a period of highly centralised communist health planning, followed by a relatively disorderly transition towards free-market economies, was not a pattern that fitted into ET theory (175). In addition, the global economic crisis has also played important roles in this stagnation. Vallin and Mesle characterised the 'cardiovascular revolution' as a specific stage of ET when life expectancy increases and chronic diseases decline, which some countries in Eastern Europe have not yet achieved. Declining life expectancy in Eastern European countries is also closely linked to unusual distortions in the structure of age-specific mortality (173).

Many LMICs have started to close the gaps between themselves and more advanced countries in terms of economy, education, and health. While doing so, in the late 20th century, new infectious diseases such as HIV/AIDS and avian influenza showed remarkable abilities to spread as global epidemics. Recent outbreaks of avian influenza alerted the world to new susceptibilities to epidemics due to population growth, unplanned urbanisation, anti-microbial resistance, poverty, and lifestyle changes in communities. AIDS still persists as a major global health concern although medical technology has progressed in preventing new HIV infections and AIDS-related deaths. A global summary of the AIDS epidemic in 2011 showed a 15% increase in the number of people living with HIV but a 13% decrease in the number of people newly infected with HIV compared with 2001 (176). Although AIDS reporting mostly focuses on national trends, there are often large variations in HIV prevalence and epidemiological patterns within countries. The epidemic of HIV/AIDS appears to have stabilised in most countries, with sub-Saharan Africa remaining the most heavily affected region, accounting for 71% of newly infected HIV cases in 2010. The AIDS epidemic triggered a decrease, and in some cases a sharp drop, in the life expectancy levels in many African countries, which is bound to influence any interpretation of their progress in ET terms. The major demographic consequences of AIDS are something that Omran could not have possibly anticipated in 1971 but represent a salutary warning about the necessary complexity of any comprehensive and generalisable understanding of ET. The impact of wars and other forms of political violence, frequent in Africa, and perhaps most extreme in Somalia (177), introduces further uncertainty into understanding ET. In other cases, impacts may not be sufficiently population-wide to hugely influence ET. For example, HIV/AIDS in Thailand in the 1990s caused a 35% increase in adult male mortality but did not have a huge impact on life expectancy for the overall population in Thailand (101).

Declines in mortality in East Asian countries varied over space and time due to different political systems, different paces of economic growth, and geographical differences. East Asia underwent a significant mortality reduction in the second half of the 20th century. Japan began this rapid progress (91) and was followed by Taiwan, Hong Kong, Korea, and China (91, 158). Some African countries have advanced at a much slower pace than other countries with similar life expectancy levels. Many African countries have not yet reached the third stage of ET due to slow health care advances, and economic and political crises. Downward trends in mortality have been reversed in some areas as a consequence of the HIV/AIDS pandemic. In sub-Saharan Africa as a whole, as ET progresses problems tend to multiply, giving rise to the concept of the quadruple burden of disease (existing infectious diseases coupled with NCDs, injuries, and HIV/AIDS), which combine to disadvantage poor and marginalised groups (31, 39, 40, 42, 145, 161). We mapped some countries studied in this review to different sub-models which account for differences in timing and pace of ET in LMICs. The rapid sub-model related to countries such as Hong Kong and China, which experienced rapid

Citation: Glob Health Action 2014, **7**: 23574 - http://dx.doi.org/10.3402/gha.v7.23574

industrialisation but remain in the third stage of ET (152, 155, 178–180). The intermediate sub-model refers to middle- or low-income countries, such as India, Vietnam, and Mexico, which still face overlapping infectious diseases and NCDs (17, 58, 60, 62, 63, 104, 181). This suggests that long-term shifts in patterns of mortality never entirely displace infectious diseases by degenerative and man-made disease and ageing, as Omran described in his theory. The slow sub-model refers to countries that are least prepared to handle the triple burden of disease, typically in Africa (38–41, 45).

We identified some new evidence of deviations from ET theory. According to the original theory, infectious diseases will gradually be eliminated as a major population health threat. However, there are examples to the contrary. In Nauru, previously high mortality from infectious diseases transitioned very rapidly into extremely high mortality from diabetes, circulatory disorders, and accidents over a short period, thus denying appreciable increases in life expectancy (132). Meanwhile, Mexico during 1922–1955 faced overlapping burdens of disease and an increasing trend of NCDs among the younger age groups, due to poverty and inability to afford healthcare (17). High mortality burdens from external causes, such as violence, in El Salvador did not follow cause of death patterns elsewhere which were predominated by infectious diseases at early stages of transition (182). Wolleswinkel-van den Bosch et al. showed in a cluster analysis in the Netherlands that the mortality decline in infectious diseases and the rise in non-infectious diseases had different time trends, in contrast with the classical interpretation of the ET that assumes uniform changes (113). Coste et al. evaluated the consistency of trends in causes of deaths in France in the short term. They found that there was a shift in mortality trends among adults during 1988–1999 that was marked by greater HIV infection, injury, and poisoning compared to mortality trends during 1968–1979. Their findings showed patterns of causes of deaths differing substantially from the classical patterns of ET (108). Mortality trends in the Tsimane communities in Bolivia showed decreased instances of deaths due to infectious diseases but increased instances of accidents and violence, particularly in the middle and late adulthood, which demonstrates a different pattern from the classical ET theory (105).

Another study in Hungary during 1971–2007 showed that the mortality patterns were different to what might be expected in the fourth stage of ET, in that female death rates increased due to ischaemic heart disease, while no evidence of emergent or re-emergent diseases were found and deaths related to external injuries declined (183). In Canada, the shifting pattern of causes of death did not entirely fit the theory. Lussier and Choiniere showed that decline in mortality related to cardiovascular disease was mostly concentrated in advanced ages, and the decline did not occur at the same pace for men and women (47). Fragmentation of the ET theory into distinctive stages delimited in time has proven to be inappropriate for Canada since the mid-20th century. The native Indian population in Canada has been shown to follow a different ET trajectory from the general population (52).

All these examples support existing criticisms that ET theory fails to describe the transition in some countries due to complexities of socio-economic, historical, political, and cultural factors that caused deviations from certain stages of the transition. Therefore, Omran's ET theory and its later developments cannot predict changes in mortality and cause-of-death patterns across all countries and time periods, as noted previously in some earlier reviews by Caselli et al. (10) and Gaylin and Kates (21). Despite these findings, we can conclude that relatively little research has concentrated on linking progress in ET to the original theory. The interest in the socio-economic determinants of ET has been greater than in the underlying disease shifts. Studies on mortality transitions in various societies still use theoretical perspectives to understand the process of population change by relating mortality patterns to demographic and socio-economic trends through the development of models (mechanisms of interaction that describe the patterns, determinants, and consequences of health and disease changes in diverse populations).

Limitations of this study

One limitation of this paper is the potential of publication bias. We are aware that many of the published papers are papers that do not confirm to the ET theory, and hence perhaps had higher chances of being accepted for publication as the utility of the theory was debated. Papers with more ambiguous results may never be published. It is a challenge to use mortality to compare ETs across countries or regions, as many factors can influence the mortality transition. The use of age-standardised death rates might make the comparison across countries more compelling, though the question on the comparability of standard populations used remains valid.

Conclusions

Despite many criticisms of Omran's theory, many researchers are still using the ET theory as the framework for their studies. In this review, we observe that the theory fits the transition patterns in some countries but needs further adjustments in other settings. With respect to broad categories, the original ET theory can, to some extent, continue to provide useful estimates of cause-specific mortality changes in countries where the transition started later than in industrialised countries and may contribute usefully to predictions of cause-specific mortality.

There are numerous examples where deviations from the classical ET theory have been observed in specific

geopolitical settings since the theory was introduced four decades ago. Although the current state of knowledge and evidence on historical and contemporary cause-specific mortality changes may still be too sparse to formulate a new evidence-based theory that provides an integrated formulation of the underlying processes of cause-specific mortality change, the need to update the ET theory is clear.

A comprehensive theory must start from the complexity of reality, even though it may not always be possible to measure all necessary parameters; a phenomenological basis will not be adequate. A new evidence-based formulation in terms of patterns of changes in causes of death, and disruptions in health due to emerging risks, is needed, which focuses on the underlying mechanisms and cause-specific mortality changes that result, rather than the current crude description of a decline in infectious diseases and a rise in non-infectious diseases with little reference to underlying determinants. Since comprehensive data are not available on a 100% basis around the world, judicious use will need to be made of the best available sources of evidence, such as those coming, for example, from the long-term surveillance of defined populations by the INDEPTH Network in otherwise uncounted regions (184). At the same time, optimal use needs to be made of data archives which document ET processes already experienced. The underlying processes of mortality change need to be described more specifically with reference to causes of death, speed of mortality changes, factors that cause disruption to health, and the ways in which populations adapt to these disruptions. Comparative analyses using standardised methods in various population groups must provide information for elaborating and refining models of transition, as is needed to handle at least some of the many problems associated with changes in population inequalities.

Main findings

- Experiences of epidemiological transition have varied between countries in terms of timing, pace and underlying mechanisms.
- Elements of epidemiological transition have been described in many countries, but observed transitions have not always followed the pathways described in Omran's original theory.
- The emphasis of epidemiological transition research has changed over the past four decades, with an increasing tendency to study wide-ranging aspects of the determinants of mortality, including risk factors, lifestyle changes, socio-economics and macro factors such as climate change.

Key messages for action

- Though it does not completely describe transition patterns observed in all settings, the original ET theory can continue to provide a useful framework for describing cause-specific mortality changes and may contribute usefully to predictions of cause-specific mortality, particularly in low- and middle-income settings.
- There is need for common consensus on a vital minimum dataset, such as a complete civil registration and documentation of major risk factors and health care coverage in low and middle-income countries to better document changing health patterns and support national health information systems. In the meantime, the lack of comprehensive population health and demographic data necessitates judicious use of the available data sources.
- Revisions to epidemiological transition theory are needed and should be based on the growing empirical evidence base in a wide range of settings.

Authors' contributions

AS has been involved in the whole process of drafting the manuscript, SW, EF, UH, and PB have been involved in revising critically the important intellectual content. All authors read and approved the final manuscript for publication.

Acknowledgements

This work was undertaken within the Umeå Centre for Global Health Research, with support from Forte (formerly FAS), the Swedish Council for Health, Working Life and Welfare (grant no. 2006-1512).

Conflict of interest and funding

The authors have not received any funding or benefits from industry or elsewhere to conduct this study. The authors declared that they have no conflict of interests relating to this study.

References

1. Kirk D. Demographic transition theory. Popul Stud 1996; 50: 361–87.
2. Thompson WS. Population. Am J Sociol 1929; 34: 959–75.
3. Landry A. The demographic revolution. Pol Popul Rev 1987; 13: 731–40.
4. Notestein F. Population – the long view. In: Schultz TW, ed. Food for the world. Chicago: Chicago University Press; 1945, pp. 37–57.
5. Omran AR. The epidemiologic transition. A theory of the epidemiology of population change. Milbank Mem Fund Q 1971; 49: 509–38.

Citation: Glob Health Action 2014, **7**: 23574 - http://dx.doi.org/10.3402/gha.v7.23574

6. Omran AR. The epidemiologic transition theory. A preliminary update. J Trop Pediatr 1983; 29: 305–16.
7. Omran AR. The epidemiologic transition theory revisited thirty years later. World Health Stat Q 1998; 51: 99–119.
8. Olshansky SJ, Ault AB. The fourth stage of the epidemiologic transition: the age of delayed degenerative diseases. Milbank Q 1986; 64: 355–91.
9. Rogers RG, Hackenberg R. Extending epidemiologic transition theory: a new stage. Soc Biol 1987; 34: 234–43.
10. Caselli G, Meslé F, Vallin J. Epidemiologic transition theory exceptions. Genus 2002; 9: 9–51.
11. Armelagos GJ, Brown PJ, Turner B. Evolutionary, historical and political economic perspectives on health and disease. Soc Sci Med 2005; 61: 755–65.
12. Gage TB. Are modern environments really bad for us?: revisiting the demographic and epidemiologic transitions. Am J Phys Anthropol 2005; Suppl 41: 96–117.
13. Weisz G, Olszynko-Gryn J. The theory of epidemiologic transition: the origins of a citation classic. J Hist Med Allied Sci 2010; 65: 287–326.
14. Kunitz SJ. The value of particularism in the study of the cultural, social and behavioral determinants of mortality. In: What we know about health transition. The cultural, social, and behavioral determinants of health. Vol. 1. Proceedings of an International Workshop, Canberra, 1990, pp. 92–109.
15. Ruzicka L, Kane P. Health transition: the course of morbidity and mortality. Determinants Health J 1990; 1: 1–26.
16. Fetter B, Coello-Ramirez P, Rogers J, Nelson M. Forum: the epidemiolocal transition. Forum Health Trans Rev 1997; 7: 235–40.
17. Carolina Martinez S, Gustavo Leal F. Epidemiological transition: model or illusion? A look at the problem of health in Mexico. Soc Sci Med 2003; 57: 539–50.
18. Frenk J, Bobadilla JL, Stern C, Frejka T, Lozano R. Elements for a theory of the health transition. Health Trans Rev 1991; 1: 21–38.
19. Smallman-Raynor M, Phillips D. Late stages of epidemiological transition: health status in the developed world. Health Place 1999; 5: 209–22.
20. Mackenbach JP. The epidemiologic transition theory. J Epidemiol Commun Health 1994; 48: 329–31.
21. Gaylin DS, Kates J. Refocusing the lens: epidemiologic transition theory, mortality differentials, and the AIDS pandemic. Soc Sci Med 1997; 44: 609–21.
22. Caldwell JC, Caldwell P. What have we learnt about the cultural, social and behavioural determinants of health? From selected readings to the first Health Transition Workshop. Health Trans Rev 1991; 1: 3–19.
23. Caldwell JC. Population health in transition. Bull World Health Organ 2001; 79: 159–60.
24. McKeown T. The road to health. World Health Forum 1989; 10: 408–16.
25. Preston S. The changing relation between mortality and socioeconomic development. Pop Stud 1975; 29: 231–48.
26. Caldwell JC. Health transition: the cultural, social and behavioural determinants of health in the Third World. Soc Sci Med 1993; 36: 125–35.
27. Caldwell JC, Caldwell P. Epidemiologic transitions. Kaohsiung J Med Sci 1999; 15: S4–14.
28. Mathers CD, Fat DM, Inoue M, Rao C, Lopes AD. Counting the dead and what they died from: an assessment of the global status of cause of death data. Bull World Health Organ 2005; 83: 171–7.
29. Alter GC, Carmichael AG. Reflections on the classification of causes of death. Contin Chang 1997; 12: 169–73.
30. Berhane Y, Wall S, Fantahun M, Emmelin A, Mekonnen W, Hogberg U, et al. A rural Ethiopian population undergoing epidemiological transition over a generation: Butajira from 1987 to 2004. Scand J Public Health 2008; 36: 436–41.
31. Misganaw A, Mariam DH, Araya T. The double mortality burden among adults in Addis Ababa, Ethiopia, 2006–2009. Prevent Chronic Dis 2012; 9: E84.
32. Agyei-Mensah S, de-Graft Aikins A. Epidemiological transition and the double burden of disease in Accra, Ghana. J Urban Health 2010; 87: 879–97.
33. Chihana M, Floyd S, Molesworth A, Crampin AC, Kayuni N, Price A, et al. Adult mortality and probable cause of death in rural northern Malawi in the era of HIV treatment. Trop Med Int Health 2012; 17: e74–83.
34. Bah SM. Assessing the contribution of age-sex differentials in causes of death due to infectious and parasitic diseases to the trends in age-sex differentials in life expectancy in Mauritius. Soc Biol 1998; 45: 260–72.
35. Abdesslam B. Evolution of rural-urban health gaps in Morocco: 1992–2011. BMC Res Notes 2012; 5: 381.
36. Bovet P. The epidemiologic transition to chronic diseases in developing countries: cardiovascular mortality, morbidity, and risk factors in Seychelles (Indian Ocean). Investigators of the Seychelles Heart Study. Soz Praventivmed 1995; 40: 35–43.
37. Stringhini S, Sinon F, Didon J, Gedeon J, Paccaud F, Bovet P. Declining stroke and myocardial infarction mortality between 1989 and 2010 in a country of the african region. Stroke 2012; 43: 2283–8.
38. Kahn K, Garenne ML, Collinson MA, Tollman SM. Mortality trends in a new South Africa: hard to make a fresh start. Scand J Public Health Suppl 2007; 69: 26–34.
39. Tollman SM, Kahn K, Sartorius B, Collinson MA, Clark SJ, Garenne ML. Implications of mortality transition for primary health care in rural South Africa: a population-based surveillance study. Lancet 2008; 372: 893–901.
40. Sartorius B, Kahn K, Collinson MA, Vounatsou P, Tollman SM. Survived infancy but still vulnerable: spatial-temporal trends and risk factors for child mortality in the Agincourt rural sub-district, South Africa, 1992–2007. Geospat Health 2011; 5: 285–95.
41. Sartorius BK, Kahn K, Vounatsou P, Collinson MA, Tollman SM. Young and vulnerable: spatial-temporal trends and risk factors for infant mortality in rural South Africa (Agincourt), 1992–2007. BMC Public Health 2010; 10: 645.
42. Bradshaw D, Groenewald P, Laubscher R, Nannan N, Nojilana B, Norman R, et al. Initial burden of disease estimates for South Africa, 2000. S Afr Med J 2003; 93: 682–8.
43. Garenne M, Kahn K, Collinson M, Gomez-Olive X, Tollman S. Protective effect of pregnancy in rural South Africa: questioning the concept of "indirect cause" of maternal death. PLoS One 2013; 8: e64414.
44. Kahn K, Tollman SM, Garenne M, Gear JS. Who dies from what? Determining cause of death in South Africa's rural north-east. Trop Med Int Health 1999; 4: 433–41.
45. Garenne M, Gakusi E. Health transitions in sub-Saharan Africa: overview of mortality trends in children under 5 years old (1950–2000). Bull World Health Organ 2006; 84: 470–8.
46. Narh-Bana SA, Chirwa TF, Mwanyangala MA, Nathan R. Adult deaths and the future: a cause-specific analysis of adult deaths from a longitudinal study in rural Tanzania 2003–2007. Trop Med Int Health 2012; 17: 1396–404.
47. Lussier MH, Bourbeau R, Choiniere R. Does the recent evolution of Canadian mortality agree with the epidemiologic transition theory? Demogr Res 2008; 18: 531–68.

48. Trovato F. Mortality differentials in Canada, 1951–1971: French, British, and Indians. Cult Med Psychiatry 1988; 12: 459–77.
49. Trovato F. Canadian Indian mortality during the 1980s. Soc Biol 2000; 47: 135–45.
50. Young TK. Are subarctic Indians undergoing the epidemiologic transition? Soc Sci Med 1988; 26: 659–71.
51. Chen J, Millar WJ. Are recent cohorts healthier than their predecessors? Health Rep 2000; 11: 9–23 (Eng); 9–6 (Fre).
52. Marrett LD, Chaudhry M. Cancer incidence and mortality in Ontario First Nations, 1968–1991 (Canada). Cancer Causes Control 2003; 14: 259–68.
53. Diaz-Briquets S. Determinants of mortality transition in developing countries before and after the Second World War: some evidence from Cuba. Popul Stud 1981; 35: 399–411.
54. Rodriguez-Ojea A, Jimenez S, Berdasco A, Esquivel M. The nutrition transition in Cuba in the nineties: an overview. Public Health Nutr 2002; 5: 129–33.
55. Rosero-Bixby L. Socioeconomic development, health interventions and mortality decline in Costa Rica. Scand J Soc Med Suppl 1991; 46: 33–42.
56. Perry H, Berggren W, Berggren G, Dowell D, Menager H, Bottex E, et al. Long-term reductions in mortality among children under age 5 in rural Haiti: effects of a comprehensive health system in an impoverished setting. Am J Public Health 2007; 97: 240–6.
57. McCaw-Binns A, Alexander SF, Lindo JL, Escoffery C, Spence K, Lewis-Bell K, et al. Epidemiologic transition in maternal mortality and morbidity: new challenges for Jamaica. Int J Gynaecol Obstet 2007; 96: 226–32.
58. Malina RM, Pena Reyes ME, Little BB. Epidemiologic transition in an isolated indigenous community in the Valley of Oaxaca, Mexico. Am J Phys Anthropol 2008; 137: 69–81.
59. Pick JB, Butler EW. Demographic, social, and economic effects on Mexican causes of death in 1990. Soc Biol 1998; 45: 151–71.
60. Stevens G, Dias RH, Thomas KJ, Rivera JA, Carvalho N, Barquera S, et al. Characterizing the epidemiological transition in Mexico: national and subnational burden of diseases, injuries, and risk factors. PLoS Med 2008; 5: e125.
61. Little BB, Malina RM, Reyes ME. Natural selection and demographic transition in a Zapotec-speaking genetic isolate in the Valley of Oaxaca, southern Mexico. Ann Hum Biol 2008; 35: 34–49.
62. Barquera S, Tovar-Guzman V, Campos-Nonato I, Gonzalez-Villalpando C, Rivera-Dommarco J. Geography of diabetes mellitus mortality in Mexico: an epidemiologic transition analysis. Arch Med Res 2003; 34: 407–14.
63. Rivera JA, Barquera S, Campirano F, Campos I, Safdie M, Tovar V. Epidemiological and nutritional transition in Mexico: rapid increase of non-communicable chronic diseases and obesity. Public Health Nutr 2002; 5: 113–22.
64. Gulliford MC. Epidemiological transition in Trinidad and Tobago, West Indies 1953–1992. Int J Epidemiol 1996; 25: 357–65.
65. Mungrue K. The changing face of death in Trinidad and Tobago, before and after independence. West Indian Med J 2012; 61: 452–9.
66. Condran GA, Cheney RA. Mortality trends in Philadelphia: age- and cause-specific death rates, 1870–1930. Demography 1982; 19: 97–124.
67. Fliess KH. Mortality transition among the Wends of Serbin, Texas, 1854–1884: changes in pattern of death from parochial records. Soc Biol 1991; 38: 266–76.
68. Potter LB. Socioeconomic determinants of white and black males' life expectancy differentials, 1980. Demography 1991; 28: 303–21.
69. Mansur Ade P, Lopes AI, Favarato D, Avakian SD, Cesar LA, Ramires JA. Epidemiologic transition in mortality rate from circulatory diseases in Brazil. Arq Bras Cardiol 2009; 93: 506–10.
70. Barros FC, Victora CG, Vaughan JP, Tomasi E, Horta BL, Cesar JA, et al. The epidemiological transition in maternal and child health in a Brazilian city, 1982–93: a comparison of two population-based cohorts. Paediatr Perinat Epidemiol 2001; 15: 4–11.
71. Chaimowicz F. Age transition of tuberculosis incidence and mortality in Brazil. Rev Saude Publica 2001; 35: 81–7.
72. Sastry N. Trends in socioeconomic inequalities in mortality in developing countries: the case of child survival in Sao Paulo, Brazil. Demography 2004; 41: 443–64.
73. Guimaraes RM, Muzi CD. Trend of mortality rates for gastric cancer in Brazil and regions in the period of 30 years (1980–2009). Arq Gastroenterol 2012; 49: 184–8.
74. Moura EC, Pacheco-Santos LM, Peters LR, Serruya SJ, Guimaraes R. Research on chronic noncommunicable diseases in Brazil: meeting the challenges of epidemiologic transition. Rev Panam Salud Publica 2012; 31: 240–5.
75. Albala C, Vio F. Epidemiological transition in Latin America: the case of Chile. Public Health 1995; 109: 431–42.
76. Huicho L, Trelles M, Gonzales F, Mendoza W, Miranda J. Mortality profiles in a country facing epidemiological transition: an analysis of registered data. BMC Public Health 2009; 9: 47.
77. Huynen MM, Vollebregt L, Martens P, Benavides BM. The epidemiologic transition in Peru. Rev Panam Salud Publica 2005; 17: 51–9.
78. Pineros M, Hernandez G, Bray F. Increasing mortality rates of common malignancies in Colombia: an emerging problem. Cancer 2004; 101: 2285–92.
79. Yang G, Kong L, Zhao W, Wan X, Zhai Y, Chen LC, et al. Emergence of chronic non-communicable diseases in China. Lancet 2008; 372: 1697–705.
80. 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.
81. Jiang G, Wang D, Li W, Pan Y, Zheng W, Zhang H, et al. Coronary heart disease mortality in China: age, gender, and urban-rural gaps during epidemiological transition. Rev Panam Salud Publica 2012; 31: 317–24.
82. Chan MF, Van IK, Ng WI. Factors contributing to neonatal mortality rates in Macao: evidence from 1957–2006 data. Nurs Health Sci 2010; 12: 410–14.
83. Gao J, Qian J, Tang S, Eriksson BO, Blas E. Health equity in transition from planned to market economy in China. Health Policy Plan 2002; Suppl 17: 20–9.
84. Schooling CM, Lau EW, Tin KY, Leung GM. Social disparities and cause-specific mortality during economic development. Soc Sci Med 2010; 70: 1550–7.
85. Joshi R, Cardona M, Iyengar S, Sukumar A, Raju CR, Raju KR, et al. Chronic diseases now a leading cause of death in rural India – mortality data from the Andhra Pradesh Rural Health Initiative. Int J Epidemiol 2006; 35: 1522–9.
86. Kumar R, Kumar D, Jagnoor J, Aggarwal AK, Lakshmi PV. Epidemiological transition in a rural community of northern India: 18-year mortality surveillance using verbal autopsy. J Epidemiol Community Health 2012; 66: 890–3.
87. Palanivel C, Yadav K, Gupta V, Rai SK, Misra P, Krishnan A. Causes of death in rural adult population of North India

(2002–2007), using verbal autopsy tool. Indian J Public Health 2013; 57: 78–83.
88. Arokiasamy P, Gautam A. Neonatal mortality in the empowered action group states of India: trends and determinants. J Biosoc Sci 2008; 40: 183–201.
89. Pradhan J, Arokiasamy P. Socio-economic inequalities in child survival in India: a decomposition analysis. Health Policy 2010; 98: 114–20.
90. Keinan-Boker L, Baron-Epel O, Fishler Y, Liphshitz I, Barchana M, Dichtiar R, et al. Breast cancer trends in Israeli Jewish and Arab women, 1996–2007. Eur J Cancer Prev 2013; 22: 112–20.
91. Kuroda T. The demographic transition in Japan. Soc Sci Med 1978; 12: 451–7.
92. Yong V, Saito Y. Are there education differentials in disability and mortality transitions and active life expectancy among Japanese older adults? Findings from a 10-year prospective cohort study. J Gerontol B Psychol Sci Soc Sci 2012; 67: 343–53.
93. Suh I. Cardiovascular mortality in Korea: a country experiencing epidemiologic transition. Acta Cardiol 2001; 56: 75–81.
94. Lee ES. Epidemiologic transition in Korea: a new perspective in population and development studies. Health Policy 1986; 6: 57–72.
95. Lee ES. Mortality transition in Korea: its implications for health policy and education. Health Policy 1986; 6: 57–72.
96. Abu-Rmeileh NM, Husseini A, Abu-Arqoub O, Hamad M, Giacaman R. Mortality patterns in the West Bank, Palestinian Territories, 1999–2003. Prev Chronic Dis 2008; 5: A112.
97. Vishnevsky A, Shkolnikov V, Vassin S. Epidemiological transition in the USSR as mirrored by regional differences. Genus 1991; 47: 7–100.
98. Niti M, Ng TP. Temporal trends and ethnic variations in amenable mortality in Singapore 1965–1994: the impact of health care in transition. Int J Epidemiol 2001; 30: 966–73.
99. Meegama SA. The mortality transition in Sri Lanka. In Determinants of Mortality Change and Differentials in Developing Countries: The Five-Country Case Study Project. Population Studies, No. 94; ST/ESA/SER.A/94. New York: United Nations, Department of International Economic and Social Affairs; 1986, pp. 5–32.
100. Bundhamcharoen K, Odton P, Phulkerd S, Tangcharoensathien V. Burden of disease in Thailand: changes in health gap between 1999 and 2004. BMC Public Health 2011; 11: 53.
101. Hill K, Vapattanawong P, Prasartkul P, Porapakkham Y, Lim SS, Lopez AD. Epidemiologic transition interrupted: a reassessment of mortality trends in Thailand, 1980–2000. Int J Epidemiol 2007; 36: 374–84.
102. Akgun S, Rao C, Yardim N, Basara BB, Aydin O, Mollahaliloglu S, et al. Estimating mortality and causes of death in Turkey: methods, results and policy implications. Eur J Public Health 2007; 17: 593–9.
103. Huong DL, Minh HV, Byass P. Applying verbal autopsy to determine cause of death in rural Vietnam. Scand J Public Health Suppl 2003; 62: 19–25.
104. Hoa NP, Rao C, Hoy DG, Hinh ND, Chuc NT, Ngo DA. Mortality measures from sample-based surveillance: evidence of the epidemiological transition in Viet Nam. Bull World Health Organ 2012; 90: 764–72.
105. Gurven M, Kaplan H, Supa AZ. Mortality experience of Tsimane Amerindians of Bolivia: regional variation and temporal trends. Am J Hum Biol 2007; 19: 376–98.
106. Georgieva L, Powles J, Genchev G, Salchev P, Poptodorov G. Bulgarian population in transitional period. Croat Med J 2002; 43: 240–4.
107. Koupilova I, McKee M, Holcik J. Neonatal mortality in the Czech Republic during the transition. Health Policy 1998; 46: 43–52.
108. Coste J, Bernardin E, Jougla E. Patterns of mortality and their changes in France (1968–99): insights into the structure of diseases leading to death and epidemiological transition in an industrialised country. J Epidemiol Community Health 2006; 60: 945–55.
109. Balogh S, Papp R, Jozan P, Csaszar A. Continued improvement of cardiovascular mortality in Hungary – impact of increased cardio-metabolic prescriptions. BMC Public Health 2010; 10: 422.
110. Kalediene R, Petrauskiene J. Inequalities in mortality by education and socio-economic transition in Lithuania: equal opportunities? Public Health 2005; 119: 808–15.
111. Kalediene R, Petrauskiene J, Starkuviene S. Inequalities in mortality by marital status during socio-economic transition in Lithuania. Public Health 2007; 121: 385–92.
112. Van Poppel F, Schellekens J, Liefbroer AC. Religious differentials in infant and child mortality in Holland, 1855–1912. Popul Stud 2002; 56: 277–89.
113. Wolleswinkel-van den Bosch JH, Looman CW, Van Poppel FW, Mackenbach JP. Cause-specific mortality trends in The Netherlands, 1875–1992: a formal analysis of the epidemiologic transition. Int J Epidemiol 1997; 26: 772–81.
114. Alfonso Sanchez MA, Mendietat VP, Pena JA, Calderon R. Demographic and health patterns in a rural community from the Basque area in Spain (1800–1990). J Biosoc Sci 2002; 34: 541–58.
115. Munoz-Tuduri M, Garcia-Moro C, Walker PL. Time series analysis of the epidemiological transition in Minorca, 1634–1997. Hum Biol 2006; 78: 619–34.
116. Alfonso-Sanchez MA, Calderon R, Pena JA. Opportunity for natural selection in a Basque population and its secular trend: evolutionary implications of epidemic mortality. Hum Biol 2004; 76: 361–81.
117. Alfonso-Sanchez MA, Pena JA, Calderon R. Time trends and determinants of completed family size in a rural community from the Basque area of Spain (1800–1969). J Biosoc Sci 2003; 35: 481–97.
118. Moraes SA, Suzuki CS, Freitas IC, Costa Junior ML. Mortality rates due to diseases of the circulatory system (DCS) in Ribeirao Preto – SP, from 1980 to 2004. Arq Bras Cardiol 2009; 93: 589–96, 637–44.
119. Moreno LA, Sarria A, Popkin BM. The nutrition transition in Spain: a European Mediterranean country. Eur J Clin Nutr 2002; 56: 992–1003.
120. Skold P, Axelsson P. The northern population development; colonization and mortality in Swedish Sapmi, 1776–1895. Int J Circumpolar Health 2008; 67: 27–42.
121. Skold P, Axelsson P, Karlsson L, Smith L. Infant mortality of Sami and settlers in Northern Sweden: the era of colonization 1750–1900. Glob Health Action 2011; 4. DOI: 10.3402/gha.v4i0.8441.
122. Demetrius L. The demographic evolution of human populations: the role of selection and environmental factors. Demography 1989; 26: 353–72.
123. Hemstrom O. Explaining differential rates of mortality decline for Swedish men and women: a time-series analysis, 1945–1992. Soc Sci Med 1999; 48: 1759–77.
124. Carson C, Hajat S, Armstrong B, Wilkinson P. Declining vulnerability to temperature-related mortality in London over the 20th century. Am J Epidemiol 2006; 164: 77–84.
125. Wild S, Fischbacher C, Brock A, Griffiths C, Bhopal R. Mortality from all causes and circulatory disease by country of

birth in England and Wales 2001–2003. J Public Health 2007; 29(2): 191–8.
126. Burnley IH. Inequalities in the transition of ischaemic heart disease mortality in New South Wales, Australia, 1969–1994. Soc Sci Med 1998; 47: 1209–22.
127. Burnley IH, Rintoul D. Inequalities in the transition of cerebrovascular disease mortality in New South Wales, Australia 1969–1996. Soc Sci Med 2002; 54: 545–59.
128. Carter K, Cornelius M, Taylor R, Ali SS, Rao C, Lopez AD, et al. Mortality trends in Fiji. Aust N Z J Public Health 2011; 35: 412–20.
129. Vigneron E. The epidemiological transition in an overseas territory: disease mapping in French Polynesia. Soc Sci Med 1989; 29: 913–22.
130. Vigneron E. Epidemiological transition and geographical discontinuities: the case of cardiovascular mortality in French Polynesia. Soc Sci Med 1993; 37: 779–90.
131. Carter K, Soakai TS, Taylor R, Gadabu I, Rao C, Thoma K, et al. Mortality trends and the epidemiological transition in Nauru. Asia Pac J Public Health 2011; 23: 10–23.
132. Schooneveldt M, Songer T, Zimmet P, Thoma K. Changing mortality patterns in Nauruans: an example of epidemiological transition. J Epidemiol Community Health 1988; 42: 89–95.
133. Davis P. Health patterns in New Zealand: class, ethnicity and the impact of economic development. Soc Sci Med 1984; 18: 919–25.
134. Sandiford P. Getting back the missing men of Aotearoa: declining gender inequality in NZ life expectancy. J Prim Health Care 2009; 1: 270–7.
135. Ahmed N, Andersson R. Unintentional injury mortality and socio-economic development among 15–44-year-olds: in a health transition perspective. Public Health 2000; 114: 416–22.
136. Moniruzzaman S, Andersson R. Relationship between economic development and suicide mortality: a global cross-sectional analysis in an epidemiological transition perspective. Public Health 2004; 118: 346–8.
137. Moniruzzaman S, Andersson R. Relationship between economic development and risk of injuries in older adults and the elderly. A global analysis of unintentional injury mortality in an epidemiologic transition perspective. Eur J Public Health 2005; 15: 454–8.
138. Plitponkarnpim A, Andersson R, Jansson B, Svanstrom L. Unintentional injury mortality in children: a priority for middle income countries in the advanced stage of epidemiological transition. Inj Prev 1999; 5: 98–103.
139. Heuveline P, Guillot M, Gwatkin DR. The uneven tides of the health transition. Soc Sci Med 2002; 55: 313–22.
140. Serow WJ, Cowart ME, Camezon J. Epidemiologic transition theory and aging: Hispanic populations of North America and the Caribbean. J Health Hum Serv Adm 1998; 20: 333–47.
141. Shah A. The possible evidence for an epidemiological transition hypothesis for elderly suicides. Int Psychogeriatr 2010; 22: 219–26.
142. Viner RM, Coffey C, Mathers C, Bloem P, Costello A, Santelli J, et al. 50-year mortality trends in children and young people: a study of 50 low-income, middle-income, and high-income countries. Lancet 2011; 377: 1162–74.
143. Murray CJ, Lopez AD. Mortality by cause for eight regions of the world: Global Burden of Disease Study. Lancet 1997; 349: 1269–76.
144. Murray CJ, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. Lancet 1997; 349: 1436–42.
145. Lopez AD, Mathers CD. Measuring the global burden of disease and epidemiological transitions: 2002–2030. Ann Trop Med Parasitol 2006; 100: 481–99.
146. Dhillon PK, Jeemon P, Arora NK, Mathur P, Maskey M, Sukirna RD, et al. Status of epidemiology in the WHO South-East Asia region: burden of disease, determinants of health and epidemiological research, workforce and training capacity. Int J Epidemiol 2012; 41: 847–60.
147. Hofmarcher MM. Is public health between East and West? Analysis of wealth, health and mortality in Austria, Central and Eastern European Countries and Croatia relative to the European Union. Croat Med J 1998; 39: 241–8.
148. Pérez-Farinós N, López-Abente G, Pastor-Barriuso R. Time trend and age-period-cohort effect on kidney cancer mortality in Europe, 1981–2000. BMC Public Health 2006; 3: 119.
149. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012; 380: 2095–128.
150. Moniruzzaman S, Andersson R. Cross-national injury mortality differentials by income level: the possible role of age and ageing. Public Health 2008; 122: 1167–76.
151. Kaasik T, Andersson R, Horte LG. The effects of political and economic transitions on health and safety in Estonia: an Estonian-Swedish comparative study. Soc Sci Med 1998; 47: 1589–99.
152. Liu Y, Rao K, Fei J. Economic transition and health transition: comparing China and Russia. Health Policy 1998; 44: 103–22.
153. Fellman J, Eriksson AW. Regional, temporal, and seasonal variations in births and deaths: the effects of famines. Soc Biol 2001; 48: 86–104.
154. Phillips DR. Problems and potential of researching epidemiological transition: examples from Southeast Asia. Soc Sci Med 1991; 33: 395–404.
155. 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.
156. Trovato F. Aboriginal mortality in Canada, the United States and New Zealand. J Biosoc Sci 2001; 33: 67–86.
157. Weden MM, Brown RA. Historical and life course timing of the male mortality disadvantage in Europe: epidemiologic transitions, evolution, and behavior. Soc Biol 2006; 53: 61–80.
158. Zhao Z, Kinfu Y. Mortality transition in East Asia. Asian Popul Stud 2005; 1: 3–30.
159. Mackenbach JP. Convergence and divergence of life expectancy in Europe: a centennial view. Eur J Epidemiol 2013; 28: 229–40.
160. Adjuik M, Smith T, Clark S, Todd J, Garrib A, Kinfu Y, et al. Cause-specific mortality rates in sub-Saharan Africa and Bangladesh. Bull World Health Organ 2006; 84: 181–8.
161. Fottrell E, Kahn K, Ng N, Sartorius B, Huong DL, Van Minh H, et al. Mortality measurement in transition: proof of principle for standardised multi-country comparisons. Trop Med Int Health 2010; 15: 1256–65.
162. Pendleton BF, Yang SO. Socioeconomic and health effects on mortality declines in developing countries. Soc Sci Med 1985; 20: 453–60.
163. Taylor R, Bampton D, Lopez AD. Contemporary patterns of Pacific Island mortality. Int J Epidemiol 2005; 34: 207–14.
164. Winegarden CR, Murray JE. Effects of early health-insurance programs on European mortality and fertility trends. Soc Sci Med 2004; 58: 1825–36.

165. Lewis ME. Impact of industrialization: comparative study of child health in four sites from medieval and postmedieval England (A.D. 850–1859). Am J Phys Anthropol 2002; 119: 211–23.
166. Murray CJ, Bobadilla JL. Epidemiological transitions in the former socialist economies: divergent patterns of mortality and causes of death. In: Bobadilla JL, Costello CA, Mitchell F, eds. Premature death in the new independent states. Washington, DC: The National Academic Press; 1997, pp. 184–219.
167. Kintner HJ. Recording the epidemiologic transition in Germany, 1816–1934. J Hist Med Allied Sci 1999; 54: 167–89.
168. Spijker JJ, Camara AD, Blanes A. The health transition and biological living standards: adult height and mortality in 20th-century Spain. Econ Hum Biol 2012; 10: 276–88.
169. Bah SM. Quantitative approaches to detect the fourth stage of the epidemiologic transition. Soc Biol 1995; 42: 143–8.
170. Verastegui E, Mohar A. Colorectal cancer in Mexico: should a middle income country invest in screening or in treatment? Eur J Health Econ 2010; 10: S107–14.
171. Fraser B. Peru's epidemiological transition. Lancet 2006; 367: 2049–50.
172. Gupte MD, Ramachandran V, Mutatkar RK. Epidemiological profile of India: historical and contemporary perspectives. J Biosci 2001; 26: 437–64.
173. Vallin J, Mesle F. Convergences and divergences in mortality. A new approach to health transition. Demogr Res 2004; 2: 11–44.
174. Shkolnikov V, McKee M, Leon D, Chenet L. Why is the death rate from lung cancer falling in the Russian Federation? Eur J Epidemiol 1999; 15: 203–6.
175. Veenema TG. Health systems and maternal and child survival in Central Asian Republics. J Nurs Scholarsh 2000; 32: 301–6.
176. World Health Organization (2011). Global HIV/AIDS response: epidemic update and health sector progress towards universal access. Geneva: World Health Organization.
177. Guha-Sapir D, Ratnayake R. Consequences of ongoing civil conflict in Somalia: evidence for public health responses. PLoS Med 2009; 6: e1000108.
178. Gu D, Dupre ME, Warner DF, Zeng Y. Changing health status and health expectancies among older adults in China: gender differences from 1992 to 2002. Soc Sci Med 2009; 68: 2170–9.
179. Cook IG, Dummer TJ. Changing health in China: re-evaluating the epidemiological transition model. Health Policy 2004; 67: 329–43.
180. Wu P, Cowling BJ, Schooling CM, Wong IO, Johnston JM, Leung CC, et al. Age-period-cohort analysis of tuberculosis notifications in Hong Kong from 1961 to 2005. Thorax 2008; 63: 312–16.
181. Guerra-Godinez JC, Larrosa-Haro A, Coello-Ramirez P, Tostado HR, Rivera-Chavez E, Castillo de Leon YA, et al. Changing trends in prevalence, morbidity, and lethality in persistent diarrhea of infancy during the last decade in Mexico. Arch Med Res 2003; 34: 209–13.
182. Avilles L. Epidemiology as discourse: the politics of development institutions in the Epidemiological Profile of El Salvador. J Epidemiol Community Health 2001; 55: 164–71.
183. Kovacs K, Hablicsek L. Cause specific mortality trends in Hungary in the light of the theory of the epidemiological transition. Budapest, Hungary: Demographic Research Institute, HSCO; 2010, 63–74.
184. Sankoh O, Byass P. The INDEPTH Network: filling vital gaps in global epidemiology. Int J Epidemiol 2012; 41: 579–88.

Global Health Action

PART II

The evolution of disease: anthropological perspectives on epidemiologic transitions

Molly Kathleen Zuckerman[1]*, Kristin Nicole Harper[2], Ronald Barrett[3] and George John Armelagos[4]

[1]Department of Anthropology and Middle Eastern Cultures, Cobb Institute of Archaeology, Mississippi State University, Mississippi State, MS, USA; [2]Department of Environmental Health Sciences, Columbia University Medical Center, New York, NY, USA; [3]Department of Anthropology, Macalester College, Saint Paul, MN, USA; [4]Department of Anthropology, Emory University, Atlanta, GA, USA

Background: The model of epidemiologic transitions has served as a guiding framework for understanding relationships between patterns of human health and disease and economic development for the past several decades. However, epidemiologic transition theory is infrequently employed in epidemiology.
Objective: Moving beyond Omran's original formulation, we discuss critiques and modifications of the theory of epidemiologic transitions and highlight some of the ways in which incorporating epidemiologic transition theory can benefit theory and practice in epidemiology.
Design: We focus on two broad contemporary trends in human health that epidemiologic transition theory is useful for conceptualizing: the increased incidence of chronic inflammatory diseases (CIDs), such as allergic and autoimmune diseases, and the emergence and reemergence of infectious disease.
Results: Situating these trends within epidemiologic transition theory, we explain the rise in CIDs with the hygiene hypothesis and the rise in emerging and reemerging infections with the concept of a third epidemiologic transition.
Conclusions: Contextualizing these trends within epidemiologic transition theory reveals implications for clinical practice, global health policies, and future research within epidemiology.

Keywords: *epidemiologic transitions*; *Omran*; *epidemiology*; *hygiene hypothesis*; *infectious disease*

Responsible Editors: Nawi Ng, Umeå University, Sweden; Barthélémy Kuate Defo, University of Montreal, Canada.

*Correspondence to: Molly Kathleen Zuckerman, Department of Anthropology and Middle Eastern Cultures, Cobb Institute of Archaeology, Mississippi State University, P.O. Box AR, Mississippi State, MS 39762, USA, Email: Mkz12@msstate.edu

This paper is part of the Special Issue: *Epidemiological Transitions – Beyond Omran's Theory*. More papers from this issue can be found at http://www.globalhealthaction.net

Received: 8 November 2013; Revised: 12 January 2014; Accepted: 21 January 2014; Published: 15 May 2014

For the past several decades, the model of epidemiologic transitions has served as a guiding framework for understanding relationships between patterns of human health and disease and economic development (1). As originally proposed by Omran (2), an epidemiologic transition is a trend wherein a high burden of mortality from infectious disease—primarily epidemic 'childhood' diseases such as pertussis and measles—is replaced by one of chronic and non-communicable diseases (NCDs), such as cardiovascular disease, cancer, and diabetes. Omran's 'classic' model was originally formulated to capture the changes in cause-specific mortality that followed the Industrial Revolution in the United States and in Western Europe. However, in a modified form, the transition is ongoing in many developing low- and middle-income countries (LMICs), which carry a high burden of mortality from infectious diseases and NCDs (3).

Recently, scholars have placed the classic transition within an expanded evolutionary framework. This recognizes a 'first' transition coincident with the Neolithic period and the agricultural revolution and a 'third' transition of emerging and reemerging infectious diseases occurring in the modern era (1). In this expanded framework, Omran's classic transition becomes the 'second epidemiologic transition', as it is referred to here.

Epidemiologic transition theory provides a model for the dynamics among economic, social, demographic, and ecological factors and the evolution and spread of disease (4), explaining major trends in the human diseasescape and granting insight into ultimate causes of a given trend and therefore potential solutions (5). Consequently, it has become paradigmatic in public health policy (3, 6–9), demography (10, 11), biological and medical anthropology (12–14), and economics (15). However, it

Citation: Glob Health Action 2014, **7**: 23303 - http://dx.doi.org/10.3402/gha.v7.23303

has had much less impact in epidemiology (5, 16, 17). This is because epidemiologists are generally concerned with the study of one or a few specific diseases within restricted time frames. Identifying a causal pathogen or characterizing a novel condition necessitates attention to its specific properties, such as statistical risk factors or disease ecology. This epidemiological approach translates into conceptualizing diseases, including emerging diseases, as singular entities attributable to more immediate and proximate causes, not as components of broader health trends attributable to longer-term and ultimate causes (5). However, as we discuss here, Omran's classic theory and its modifications can help to broaden the epidemiologist's understanding of the complex, multiple dimensions of health and disease over time, revealing potential proximate as well as ultimate causes, prevention strategies, and predictions of future epidemiologic trends and, in doing so, contribute to improving population health (5, 16).

Moving beyond Omran's original formulation, this paper critiques and modifies the theory of epidemiologic transitions, highlighting some of the ways that an understanding of epidemiologic transitions can benefit theory and practice in epidemiology. We also focus on two broad contemporary trends in human health that are relevant to epidemiologic transition theory: the increased incidence of chronic inflammatory diseases (CIDs), such as allergic and autoimmune diseases, and the emergence and reemergence of infectious diseases. We situate these trends within an expanded epidemiologic transition theory, employing the hygiene hypothesis to explain the rise in CIDs and the concept of a third epidemiologic transition to explain the rise in emerging and reemerging infections. We further discuss the implications that this approach generates for preventive medicine, global health policies, and future research in epidemiology.

Present investigation

The theory of epidemiologic transitions

Epidemiologic transition theory models the changes in cause-specific mortality that accompanied the industrialization-associated demographic transition, the declines in mortality and fertility, and the resulting population growth (18) (see Fig. 1). Demographic transition theory is a simplified, descriptive, multistage model of this transition (14). Stage 1 is largely preindustrial and features high mortality and fertility, variable but generally low life expectancy, and slow but stable population growth interrupted by large periodic fluctuations in mortality (crisis mortality). Stage 2 involves declining 'normal' mortality but persistently high fertility. Stage 3 features continued declines in normal mortality as well as fertility, increased life expectancy, and more sustained population growth. Declining normal mortality began in the mid-nineteenth century in western and northern Europe and the United States and around 1920 in LMICs, such as Chile. Crisis mortality continued into the twentieth century, perhaps ending only with the 1918–1919 Spanish influenza outbreak (14). Stage 4 features low, still declining mortality, as observed in developed countries after World War II (19). Mortality declines continue, especially among the elderly, but slowly, likely due to already low mortality among infants, children, and young adults (14).

As originally formulated (2), the epidemiologic transition theory builds on demographic transition theory with a consideration of changes in cause specific mortality across four stages. Stage 1, the 'Age of Pestilence and Famine', is preindustrial and is characterized by high frequencies of epidemic infectious disease and crisis mortality. Stages 2 and 3, the 'Age of Receding Pandemics', involve a change from epidemic to endemic infectious disease and decreasing crisis mortality. Stage 4, the 'Age of Degenerative and Man-Made [*sic*] Diseases', ushers in a high burden of NCDs, the result of age-related degenerative processes and anthropogenic factors including environmental hazards and nutritional and behavioral patterns associated with industrialization and urban living. The neat dichotomy between infections and NCDs is blurred by the chronic course of some infections, such as tuberculosis, and increasing recognition of the role of infection and inflammatory processes in many chronic conditions, such as cervical cancer (20) and coronary heart disease (21). However, rather than precluding the use of transition theory, this complication highlights the need to consider the historical and evolutionary relationships between humans and pathogens to understand current patterns of human health.

Modifications and critiques of epidemiologic transition theory

Multiple modifications have been made to Omran's original theory. Omran (2), attempting to accommodate variations from the theory, initially characterized three models for the progress of the transition. The 'classical model', featuring slow progression from high to low mortality and fertility, occurred in Western Europe, and was driven by economic development and advances in sanitation, public health, and medical knowledge. The 'accelerated model', seen in Japan, features more rapid progress, purportedly driven by the same factors. The 'contemporary model' applies to LMICs, wherein mortality has declined but fertility remains high and NCDs do not yet constitute the primary epidemiologic burden. Beyond the expansions suggested by Barrett and colleagues (1), Olshansky and colleagues have added a fourth stage, the 'Age of Delayed Degenerative Diseases', which recognizes a shift in NCDs to the elderly even as life

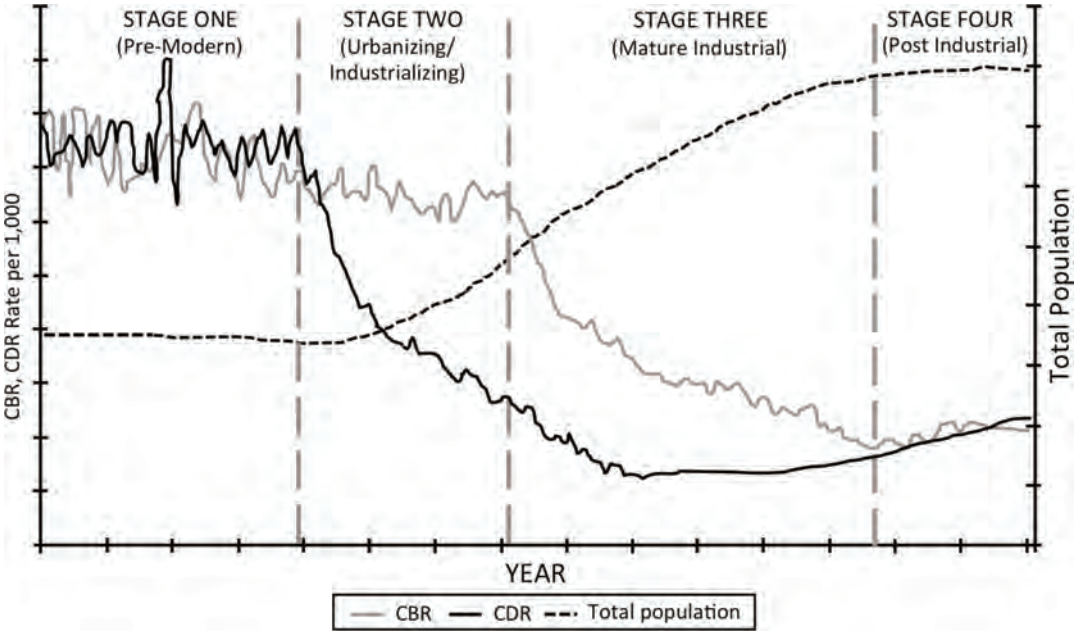

Fig. 1. The demographic transition model (Source: K. Montgomery).

expectancy increases, and a fifth, encompassing morbidity and mortality from HIV/AIDS and other emerging infectious diseases (22, 23). Graziano (24) has proposed an alternative fifth stage, the 'Age of Obesity and Inactivity', in which increasing levels of weight gain and obesity alter the pattern of NCDs among the elderly (represented by Olshansky and Ault's fourth stage).

From anthropological and epidemiological perspectives, the original epidemiologic transition theory has several shortcomings. The proposed explanations are largely speculative, lacking reliable data or relying on small sample sizes that are not likely to be generalizable (16). Many have also argued that the theory insufficiently addresses social factors, such as poverty (12); the differential progress of the transition within and across various demographic subgroups—such as in relation to sex, gender, race, and location (12, 25); and the profound contemporary impact of emerging and reemerging infectious diseases, such as HIV/AIDS and multidrug resistant tuberculosis (MDR TB)(12, 25, 26). It also fails to differentiate adequately between the *risk* of dying from any given cause or set of causes and the *proportion* of overall mortality due to various causes (14, 25). Additionally, by focusing on mortality, the theory largely neglects issues of morbidity, disability, and quality of life. This approach is sharply divergent from the increasing use of more holistic definitions of health and the use of broader measures of life expectancy, such as disability adjusted life years (DALYs), by health researchers and organizations (16). Saliently, Caldwell (6: 160) and others have also critiqued Omran's models for failing to recognize 'the global nature and historical sequence of the mortality transition as it spread', positing instead that each society exhibits its own particular model. Despite these issues, however, the theory remains profoundly useful for investigating variation in patterns over time and among locations, for conceptualizing historical patterns, and for predicting future trends (17).

Epidemiologic transition theory and epidemiology

Although epidemiology largely maintains its focus on single disease conditions and proximate causes, the rise of the socioecological model in epidemiology represents an opportunity for integrating epidemiologic transition theory into epidemiologic theory and practice. This model replaces the earlier 'epidemiologic triad' and 'causal web' models of disease risk, which paid inadequate attention to the multidimensional quality of factors that affect health and disease (27), particularly the cultural, political, and economic factors that act as ultimate causes of disease (28, 29). Instead, the socioecological model recognizes that a broad array of systems and interrelated determinants of health exist, acting either synergistically or antagonistically (16). The model vertically expands the domain of epidemiological studies 'upward' to incorporate biological, behavioral, mental, social, and environmental systems and factors (such as policy and economic environments) as well as 'downward' to the molecular and genetic levels. Additionally, it extends the domain horizontally to consider trends over time, ranging from the developmental issues addressed by life course epidemiology to evolving associations among the various levels (16: 6).

Citation: Glob Health Action 2014, **7**: 23303 - http://dx.doi.org/10.3402/gha.v7.23303

Epidemiologic transition theory provides the broad, longitudinal, and historical perspectives that the early models lacked (16). In epidemiology, the increasing adoption of and reliance upon socioecological models has also opened up a spectrum of possibilities for investigations of disease risk that include social, cultural, policy, and economic factors, as well as how these factors have changed over time (5, 16). The unique capacity of epidemiologic transition theory to identify ultimate causes in disease risk means that transition theory models can be usefully incorporated into socioecological models in epidemiology. As Fleischer and McKeown (16: 10) discuss, such incorporation would be particularly useful for social epidemiologists because it would shed additional light on the ways in which 'upstream', global determinants differentially impact the health-states of vulnerable, disadvantaged populations. Although the focus on singular disease conditions, their risk factors, and proximate causes is useful in some investigations, epidemiologists are gradually recognizing that addressing patterns of health-states might yield greater insights into more upstream, shared determinants and therefore have a profound impact upon population health (16).

Contemporary trends in health addressed through epidemiologic transition theory

CIDs and emerging and reemerging infectious disease

Epidemiological transition theory can be usefully applied to a variety of broad trends in patterns of health and disease in contemporary human populations. Here we focus on two specific trends: the increased incidence of CIDs and the emergence and reemergence of infectious diseases. We use these examples to demonstrate how the interpretive lens and attention to ultimate causes generated by epidemiologic transition theory can be used to inform clinical practice, global health policies, and future research within epidemiology.

The hygiene hypothesis and the rise of CIDs

Since the 1950s, numerous researchers have noted increased rates of CIDs, namely allergic and autoimmune diseases, in many developed nations (30). These include allergic diseases, such as asthma and atopic dermatitis, and autoimmune diseases, such as multiple sclerosis (MS) and inflammatory bowel disease. For instance, the prevalence of atopic dermatitis has doubled or tripled in developed countries over the past three decades, affecting 15 to 30% of children and 2 to 10% of adults (31). Part of these increases may be attributable to detection bias via improved diagnoses and access to medical resources in many locations, but this does not entirely explain the rapid rise in the incidence of these conditions, particularly those such as MS that are easily diagnosed (32).

Many epidemiologists have noticed that the increase in CIDs is concomitant with a decrease in epidemic infections that occurred during the second epidemiologic transition. However, rather than highly virulent pathogens, such as smallpox or measles (33), the phenomenon has been attributed to the diminished exposure to environmental microorganisms (specifically helminthic parasites, chronic viruses and nonlethal bacterial infections, environmental saprophytes) and to declines in the mass and diversity of gut microbiota, which also occurred as part of the second transition (34). Situated within epidemiologic transition theory, these observations have given rise to the 'hygiene hypothesis', which claims that in developed nations, diminished childhood exposure during childhood or even diminished prenatal (maternal) exposure (35) to these microorganisms has resulted in immunoregulatory failure, manifesting as an increased incidence of CIDs. According to this theory, prior to industrialization, humans lived in a state of 'evolved dependence' with these microbes—our 'old friends'—that was critical for successful immunological functioning, specifically the development of an immunomodulatory response that maintains tolerance of self-antigens and abrogates autoimmune diseases (34). In effect, the lifestyle changes—sanitary improvements, pasteurization, use of antibiotics, and improved hygiene—that contributed to the second transition may have produced a substantial trade-off in health and quality of life, with developed nations exchanging a high burden of infectious disease for a higher burden of CIDs (36).

Using the hygiene hypothesis to identify potential ultimate causes of CIDs has several direct implications for epidemiology. Although the identified causes are broadly environmental and evolutionary, they directly translate into proximate causes and therefore specific risk factors and targets for preventive medicine. Informed by the hygiene hypothesis, researchers have identified causal relationships among lifestyle changes, infectious burden, and the incidence of allergic and autoimmune diseases. They continue to investigate these dynamics using animal models of autoimmune and allergic diseases and, to a lesser extent, clinical intervention studies (32). For example, the incidence of spontaneous type 1 diabetes (T1D) is directly correlated with the sanitary conditions of animal facilities for non-obese diabetic (NOD) mice, with a low infectious burden translating into a high T1D incidence (37). Another study found that intranasal exposure of pregnant mice to cowshed derived, non-pathogenic bacterium, *Acinetobacter lwoffii* F78, protected against the development of experimental asthma in their progeny (35). Other researchers have found that in humans, intentional infection with the swine-derived helminth, *Trichuris suis*, ameliorated symptoms in patients with active Crohn's disease as well as ulcerative colitis (38, 39). Although research findings are not yet determinative—for instance, helminth eradication has been found to increase atopic skin sensitization (40) while

improving asthma symptoms in the same population (41)—studies such as these suggest that approaching the broad trend of increased incidence of CIDs through an evolutionary, historical lens generates findings that can be directly applied to improving population health. Research focused on determining which types of microbial exposures exert a protective effect on developing allergic sensitization (and when during the life course they occur) can translate into improved identification of risk factors, targeted strategies for disease prevention, and foci for preventive medicine.

A practical outcome of this approach can be found in the recent proposed ruling by the Food and Drug Administration (FDA) that manufacturers of antibacterial soap that contains the chemicals triclosan and triclocarban demonstrate that they are safer and more effective than soap and water. This ruling is the product of years of mounting fears that the chemicals in the soaps, the use of which has greatly proliferated in the past few decades, may disrupt normal development of the reproductive system and metabolism and promote drug-resistant infections, among other issues (42). This action reflects increasing public health awareness that interfering with long-standing balances between humans and nonpathogenic microbes in their environments can have pervasive and substantial health effects.

Addressing the ancient determinants of emerging infections

One of the chief advantages of an expanded framework of epidemiologic transitions, such as that described by Barrett and colleagues (1), is that it allows us to backtrace the ultimate determinants of current infections over long stretches of time back into ancient history and prehistory. Looking back to the Neolithic period, changes in subsistence, settlement, and social organization associated with the agricultural revolution created conditions that selected for the emergence of acute, epidemic, 'crowd' infectious diseases, such as smallpox, measles, and pertussis, which were among the chief causes of human morbidity and mortality up until the second epidemiologic transition (1). Agriculture allowed for the production of high calorie foods at the cost of dietary diversity (43), and agricultural communities that rely heavily on a few crops often experience compromised nutrition that can predispose them to infection (44, 45). In addition, domesticated animals served as the source for many novel human pathogens, such as measles and smallpox, as did the expansion of farmers into new terrain suited to the transmission of diseases such as malaria and yellow fever (46). The transition to large and densely settled societies also allowed for the ongoing transmission of more virulent and shorter-lived infections, such as crowd diseases, within and between human populations. Furthermore, the creation of social hierarchies led to the unequal distribution of basic resources for healthy living, thereby creating reservoirs for new and recurring infections in impoverished communities (12, 47).

Today, we see the same themes of subsistence, settlement, and social organization in the emergence and reemergence of infectious diseases. The hyperurbanization of domesticated animals, combined with the use of antibiotic growth factors, is contributing to increases in selective conditions for the entry of zoonotic pathogens from animal to human populations (48). Human settlements have consolidated to the point that the majority of the global population now lives in urban environments with ample opportunities for ongoing disease transmission (49). Additionally, globalization has linked the health problems of impoverished communities with other populations throughout the globe, in developed nations and LMICs, such that humans can now be said to live within a single infectious disease ecology. The scale and speed of human activities have greatly increased since the Neolithic, but the activities themselves are qualitatively the same, despite the novelty of some pathogens.

This longer term, evolutionary perspective has practical applications for the prevention and control of new and drug-resistant infections. Recognizing that the agricultural revolution brought new infections into human populations through increased exposure to animals, researchers have examined the entry of novel contemporary pathogens through human settlement in uninhabited wilderness areas and subsistence practices such as bushmeat hunting. Research has shown that bushmeat handling and consumption of nonhuman primates in particular has provided a highly effective pathway for the transmission of zoonotic pathogens into human populations, both in the past and today (50); several infections in humans, such as HIV/AIDS (via simian immunodeficiency viruses [SIVs]) (51) and a wide variety of human T-lymphotropic viruses (HTLVs), which are associated with leukemia, lymphoma, and HTLV-associated myelopathy (52), have been linked to pathogens present in nonhuman primates. Understanding these relationships and the pathological consequences of interspecies interactions has substantial implications for public health. As in Neolithic times, the wide array of subsistence practices that increase exposure between humans and animals increases the probability that nonhuman pathogens will, through a series of evolutionary steps, evolve the ability to infect humans, and later, evolve the capacity for sustained transmission within and between human populations. Recognizing these dynamics helps to inform international economic and environmental policies and technological innovations aimed at reducing these exposures (53).

Similarly, researchers can extend lessons from the intensification of agriculture to preventing, or at least reducing, the evolution of drug resistant infections.

Citation: Glob Health Action 2014, **7**: 23303 - http://dx.doi.org/10.3402/gha.v7.23303

The FDA announced in December 2013 that it was phasing out the use of antibiotics as growth promoters in cows, pigs, and chickens (54). Although this policy is voluntary for drug manufacturers and subject to loopholes, it represents one of the first attempts by the U.S. government to address one of the ultimate causes driving the third epidemiologic transition, specifically the reemergence of previously controlled infectious diseases. The European Union has long since recognized this issue, having passed regulations against the use of human antibiotics, and their analogs, as growth factors in livestock (55). This policy change and the research underlying it reflect an increased understanding in the public health community that drug resistance not only poses a large and growing threat to human health but that this trend fits into a larger evolutionary pattern of selective relationships between humans, animals, and pathogens. Ten thousand years ago, domestication of animals gave human populations zoonotic infections; in the present, our interactions with domesticated animals are conferring drug resistant infections upon us.

Conclusion: global health and the third epidemiologic transition

The current trends of novel, virulent, and drug resistant infectious diseases represent a third epidemiologic transition in human disease. To some extent, this can be seen as a convergence of disease patterns associated with the two previous epidemiologic transitions. As with the first transition, physical and social environments are once again selecting for the emergence and transmission of acute epidemic infections in human populations. As with the second transition, NCDs continue to rise in populations rich and poor, in developed nations and LMICs alike. In a combination of these two trends, acute and chronic diseases are leading to the emergence of global syndemics, wherein the interaction between two or more health conditions—such as among diabetes, cardiopulmonary disease, and severe acute respiratory syndrome (SARS)—has a multiplicative rather than additive impact on human well-being (56).

An expanded theory of epidemiologic transitions not only deepens understanding of these global health trends, it also informs policies and programs for prevention, detection of novel conditions and pathogens, identification of risk factors, and control of diseases. Shifting from primarily vertically organized programs aimed at single diseases and proximate causes, epidemiologists and other health researchers can develop complementary horizontally based programs aimed at combinations of infectious and chronic diseases and the upstream determinants that they have in common. The U.S. Centers for Disease Control and Prevention (CDC) has begun this process with the merging of previously independent programs and divisions within its organization (57). Similarly, observers of polio eradication programs, for example, have increasingly begun to recognize that treatment and vaccination cannot be successfully implemented without addressing broader health determinants, such as those responsible for the first and second epidemiologic transitions (58). Applying these lessons to current global health challenges may not eradicate all human diseases, but it could generate and marshal support for novel, broadscale interventions that target the ultimate causes for many important health conditions. At that point, we hope to write optimistically about the positive lessons evident a fourth epidemiologic transition.

Main findings

- Integrating epidemiological transition theory into theory and practice in epidemiology can broaden epidemiologists' understanding of the complex, multiple dimensions of health and disease over time. This can reveal potential proximate and ultimate causes, prevention strategies, and predictions of future epidemiologic trends, therefore contributing to improvements in population health.
- Contextualizing disease trends, such as the increased incidence of chronic inflammatory diseases (CIDs), including allergy and autoimmune diseases, and the emergence and re-emergence of infectious disease, within epidemiological transition theory, and attending to historical and evolutionary relationships between humans and pathogens, has direct implications for clinical practice, global health policies, and future research within epidemiology.

Key messages for action

- Situating the increased incidence of CIDs within the hygiene hypothesis, a theory closely associated with epidemiologic transition theory that ties the trend to immunoregulatory failure brought on by decreased early life exposure to non-pathogenic environmental microbes, can be used to identify specific risk factors, targeted strategies for disease prevention, and foci for preventive medicine in global efforts to reduce CIDs.
- Addressing the emergence and re-emergence of infectious disease, such as HIV/AIDs and MDR-TB, within long term evolutionary perspectives can reveal the ancient and prehistoric ultimate determinants of current infections, generating practical applications for the prevention and control of new and drug-resistant infections, such as attendance to the role of

bush-meat and environmental degradation in the emergence and spread of disease.

- The critical insight from epidemiologic transition theory that interference with long-standing balances between humans and non-pathogenic microbes in their environments can have pervasive and substantial health effects has directly informed recent action in public health, such as the FDA's proposed rulings discouraging the use of antibiotics in livestock feed and antibacterial compounds in soaps.

Conflict of interest and funding

The authors have not received any funding or benefits from industry or elsewhere to conduct this study.

References

1. Barrett R, Kuzawa C, McDade T, Armelagos G. Emerging and re-emerging infectious diseases: the third epidemiologic transition. Annu Rev Anthropol 1998; 27: 247–71.
2. Omran A. The epidemiologic transition. Milbank Mem Fund Q 1971; 49: 509–38.
3. Marinho F, Soliz P, Gawryszewski V, Gerger A. Epidemiological transition in the Americas: changes and inequalities. Lancet 2013; 381: S89.
4. Armelagos G, Barnes K. The evolution of human disease and the rise of allergy: epidemiological transitions. Med Anthropol 1999; 18: 187–213.
5. Harper K, Armelagos G. The changing disease-scape in the third epidemiological transition. Int J Environ Res Public Health 2010; 7: 675–97.
6. Caldwell J. Population health in transition. Bull World Health Organ 2001; 79: 159–60.
7. Girard DZ. The cost of epidemiological transition: a study of a decrease in pertussis vaccination coverage. Health Policy 2005; 74: 287–303.
8. Huicho L, Trelles M, Gonzales F, Mendoza W, Miranda J. Mortality profiles in a country facing epidemiological transition: an analysis of registered data. BMC Publ Health 2009; 9: 47.
9. Szreter S. Rethinking McKeown: the relationship between public health and social change. Am J Publ Health 2002; 92: 722–5.
10. McKeown T, Record R. Reasons for the decline of mortality in England and Wales during the nineteenth century. Popul Stud 1962; 16: 94–122.
11. Kirk D. Demographic transition theory. Popul Stud 1996; 50: 361–87.
12. Armelagos G, Brown P, Turner B. Evolutionary, historical and political economic perspectives on health and disease. Soc Sci Med 2005; 61: 755–765.
13. Munoz-Tuduri M, Garcia-Moro C, Walker PL. Time series analysis of the epidemiological transition in Minorca, 1634–1997. Hum Biol 2006; 78: 619–34.
14. Gage TB. Are modern environments really bad for us?: revisiting the demographic and epidemiologic transitions. Yearbk Phys Anthropol 2005; 48: 96–117.
15. Morand OF. Economic growth, longevity and the epidemiological transition. Eur J Health Econ 2004; 5: 166–74.
16. Fleischer N, McKeown R. The second epidemiologic transition from an epidemiologist's perspective. In: Zuckerman M, ed. Modern environments and human heath: revisiting the second epidemiologic transition. Hoboken, NJ: Wiley-Blackwell; 2014, pp. 353–69.
17. Mackenbach JP. The epidemiological transition theory. J Epidemiol Community Health 1994; 48: 329–31.
18. Thompson W. Population. Am J Sociol 1929; 34: 959–75.
19. Vallin J. Mortality in Europe from 1720 to 1914: long-term trends and changes in patterns by age and sex. In: R Schofield, D Reher, A Bideau eds. The decline of mortality in Europe. Oxford: Clarendon Press; 1991, pp. 38–67.
20. Walboomers J, Jacobs M, Manos M, Xavier Bosch F, Kummer J, Shah K, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol 1999; 189: 12–9.
21. Humphrey L, Buckley D, Freeman M, Helfand M. Periodontal disease and coronary heart disease incidence: a systematic review and meta-analysis. J Gen Intern Med 2008; 23: 2079–86.
22. Olshansky S, Ault A. The fourth stage of the epidemiologic transition: the age of delayed degenerative diseases. Milbank Q 1986; 64: 355–91.
23. Olshansky S, Carnes B, Rogers R, Smith L. Emerging infectious diseases: the fifth stage of the epidemiologic transition? World Health Stat Q 1998; 51: 207–17.
24. Graziano J. Fifth phase of the epidemiologic transition: the age of obesity and inactivity. JAMA 2010; 303: 275–6.
25. McKeown R. The epidemiologic transition: changing patterns of mortality and population dynamics. Am J Lifestyle Med 2009; 3: 19S–25S.
26. Sanders J, Fuhrer G, Johnson M, Riddle M. The epidemiological transition: the current status of infectious diseases in the developed world versus the developing world. Sci Prog 2008; 91: 1–37.
27. Krieger N. Embodiment: a conceptual glossary for epidemiology. J Epidemiol Community Health 2005; 59: 350–5.
28. Turshen M. The political ecology of disease. Rev Radic Polit Econ 1977; 9: 45–60.
29. Farmer P. Social inequalities and emerging infectious diseases. Emerg Infect Dis 1996; 2: 259–69.
30. Munoz-Lopez F. Validity of the hygiene hypothesis [comment]. Allergol Immunopathol 2006; 34: 129–30.
31. Williams H, Flohr C. How epidemiology has challenged 3 prevailing concepts about atopic dermatitis. J Allergy Clin Immunol 2006; 118: 209–13.
32. Okada H, Kuhn C, Feillet H, Bach J. The 'hygiene hypothesis' for autoimmune and allergic diseases: an update. Clin Exp Immunol 2010; 160: 1–9.
33. Strachan DP. Hay fever, hygiene, and household size. Br Med J 1989; 299: 1259–60.
34. Rook GA. 99th Dahlem conference on infection, inflammation and chronic inflammatory disorders: Darwinian medicine and the 'hygiene' or 'old friends' hypothesis. Clin Exp Immunol 2010; 160: 70–9.
35. Conrad ML, Ferstl R, Teich R, Brand S, Blümer N, Yildirim AO, et al. Maternal TLR signaling is required for prenatal asthma protection by the nonpathogenic microbe Acinetobacter lwoffii F78. J Exp Med 2009; 206: 2869–77.
36. Zuckerman M, Armelagos G. The hygiene hypothesis and the second epidemiologic transition. In: Zuckerman M, ed. Modern environments and human health: revisiting the second epidemiologic transition. Hoboken, NJ: Wiley-Blackwell; 2014, pp. 303–21.
37. Bach JF. The effect of infections on susceptibility to autoimmune and allergic diseases. N Engl J Med 2002; 347: 911–20.
38. Summers R, Elliott D, Urban J, Jr, Thompson R, Weinstock J. Trichuris suis therapy in Crohn's disease. Gut 2005; 54: 87–90.

39. Summers R, Elliott D, Urban J, Jr, Thompson R, Weinstock J. Trichuris suis therapy for active ulcerative colitis: a randomized controlled trial. Gastroenterology 2005; 128: 825–32.
40. Lynch N, Hagel I, Perez M, Di Prisco M, Lopez R, Alvarez N. Effect of anthelmintic treatment on the allergic reactivity of children in a tropical slum. J Allergy Clin Immunol 1993; 92: 404–11.
41. Lynch N, Palenque M, Hagel I, DiPrisco M. Clinical improvement of asthma after anthelminthic treatment in a tropical situation. Am J Respir Crit Care Med 1997; 156: 50–4.
42. Tavernise S. F.D.A. questions safety of antibacterial soaps. New York: The New York Times; 2013, pp. A21.
43. Reid W, Miller K. Food crops and biodiversity. Keeping options alive: the scientific basis for the conservation of biodiversity. Washington, DC: World Resources Institute; 1989.
44. Scrimshaw N. Historical concepts of interactions, synergism and antagonism between nutrition and infection. J Nutr 2003; 133: 316–21.
45. Scrimshaw N. INCAP studies of nutrition and infection. Food Nutr Bull 2010; 31: 54–67.
46. Wolfe N, Dunavan C, Diamond J. Origins of major human infectious diseases. Nature 2007; 447: 279–83.
47. Flannery K. The creation of inequality: how our prehistoric ancestors set the stage for Monarchy, Slavery, and Empire. Cambridge, MA: Harvard University Press; 2012.
48. Davis M. The monster at our door: the global threat of Avian Flu. New York: Owl Books; 2006.
49. Moore M, Gould P, Keary B. Global urbanization and impact on health. Int J Hyg Environ Health 2003; 206: 269–78.
50. Harper K, Zuckerman MK, Turner BL, Armelagos GJ. Primates, pathogens, and evolution: a context for understanding emerging disease. In: J Brinkworth, K Pechenkina eds. Primates, pathogens, and evolution. New York: Springer Publishing; 2013, pp. 389–409.
51. Hahn B, Shaw G, Cock K, Sharp P. AIDS as a zoonosis: scientific and public health implications. Science 2000; 287: 607–14.
52. Wolfe N, Heneine W, Carr J, Garcia A, Shanmugam V, Tamoufe U, et al. Emergence of unique primate T-lymphotropic viruses among central African bushmeat hunters. Proc Natl Acad Sci Unit States Am 2005; 102: 7994–9.
53. Wolfe N. Viral storm: the dawn of a new pandemic. New York: Times Books; 2011.
54. Tavernise S. F.D.A. restricts antibiotics use for livestock. New York: The New York Times; 2013, pp. A1.
55. Amyes S. Magic bullets, lost horizons: the rise and fall of antibiotics. New York: Taylor & Francis; 2001.
56. Singer M. Ecosyndemics: global warming and the coming plagues of the 21st century. In: A Swedlund, A Herring eds. Plagues and epidemics: infected spaces past and present. London: Berg; 2010, pp. 21–37.
57. CDC. Available from http://www.cdc.gov/nchhstp/programintegration [cited 11 July 2013].
58. Closser S. Chasing polio in Pakistan: why the world's largest health initiative may fail. Nashville, TN: Vanderbilt University Press; 2010.

Global Health Action

PART III

Reorienting women's health in low- and middle-income countries: the case of depression and Type 2 diabetes

Emily Mendenhall[1]* and Lesley Jo Weaver[2]

[1]Science, Technology, and International Affairs Program, School of Foreign Service, Georgetown University, Washington, DC, USA; [2]Department of Anthropology, Emory University, Atlanta, Georgia, USA

Women's health in low- and middle-income countries (LMICs) has historically focused on sexual and reproductive health. However, understanding how women acquire, experience, and treat non-reproductive health conditions, such as non-communicable diseases, has become a fundamental public health concern. Special attention to the social determinants of LMIC women's health can provide socially and culturally relevant knowledge for implementation of policies and programs for women increasingly confronting these 'New Challenge Diseases'. This article uses the example of depression and Type 2 diabetes comorbidity to illustrate how attending to the social determinants of mental and physical health beyond the reproductive years contributes to a more holistic agenda for women's health. For instance, we must address the plurality of experiences that shape women's health from social determinants of depression, such as gendered subjugation within the home and public sphere, to the structural determinants of obesity and diabetes, such as poor access to healthy foods and health care. Attending to the complexities of health and social well-being beyond the reproductive years helps the women's global health agenda capture the full spectrum of health concerns, particularly the chronic and non-communicable conditions that emerge as life expectancy increases.

Keywords: *women's health*; *depression*; *Type 2 diabetes*; *life course*; *social determinants*; *epidemiological transition*

Responsible Editors: Nawi Ng, Umeå University, Sweden; Barthélémy Kuate Defo, University of Montreal, Canada.

*Correspondence to: Emily Mendenhall, Science, Technology, and International Affairs Program, School of Foreign Service, Georgetown University, Washington, DC, USA, Email: em1061@georgetown.edu

This paper is part of the Special Issue: *Epidemiological Transitions – Beyond Omran's Theory*. More papers from this issue can be found at http://www.globalhealthaction.net

Received: 10 September 2013; Revised: 12 November 2013; Accepted: 13 December 2013; Published: 9 January 2014

The contemporary landscape of women's health in low- and middle-income countries (LMICs) is more complex than public health approaches in previous decades reflected, when the focus was primarily on sexual and reproductive health. As populations age, no longer are sexual and reproductive health the dominant themes that shape how women can live longer, healthier lives. Instead, a combined perspective of the social determinants of mental and physical health across the life course comes to the forefront. Understanding how women acquire, experience, and treat non-reproductive health conditions, such as non-communicable diseases, over the course of their lives is particularly important for women living in resource-constrained settings, who are socially and economically marginalized and often experience limited access to healthcare. This article uses the example of depression and Type 2 diabetes comorbidity to illustrate how attending to the social determinants of mental and physical health beyond the reproductive years contributes to a more holistic agenda for women's health.

This shift in priorities requires that we break down the traditional distinctions between 'chronic' and 'acute', 'communicable' and 'non-communicable' diseases because in fact they often occur together. For instance, diabetes and tuberculosis not only coexist within a given population but also can coexist within a single individual. Likewise, over- and under-nutrition can exist simultaneously in communities, households, or even individuals during different phases of their lives. In light of these complex scenarios, Knaul and Frenk have suggested that we rethink public health paradigms for the challenges of aging populations as 'New Challenge Diseases' rather than 'non-communicable diseases' (1). This approach requires that we move beyond diseased-focused silos in public health. Instead, we must address the plurality of experiences that shape women's health from social determinants of depression, such as gendered subjugation within the home and public sphere, to the structural determinants of obesity and diabetes, such as poor access to healthy foods and health care. Women living in LMICs require special attention not only because their

Citation: Glob Health Action 2014, **7**: 22803 - http://dx.doi.org/10.3402/gha.v7.22803

experiences are unique to women living in affluent nations but also because such limited research is available on their social and health problems. Bias of research from high-income nations may construe LMIC women's experiences and contribute not only to knowledge displaced from women's social experiences but also policies and programs that do not reflect the social, economic, and cultural factors surrounding women's mental and physical health problems in LMICs.

As opposed to traditional disease-based approaches in medicine and public health, a life course approach encompasses the powerful role of social and economic determinants of health in women's lives from infancy to old age (2, 3). This approach is particularly important for women who may experience disproportionate social disadvantage, gendered discrimination, and chronic, untreated depression when compared to men (4). Indeed, new global data demonstrates that women's health is overall poorer than their male counterparts around the world (5), and this is largely due to socially driven inequalities. Recognizing this is crucial for understanding and managing chronic diseases, which typically have complex etiologies rooted in long-term lifestyle choices as well as intergenerationally heritable characteristics, both genetic and behavioral. A life course perspective acknowledges, for instance, that social and economic problems related to poverty both fuel poor health and result from it, creating cycles that are difficult to break.

We present complexities of the comorbidity between Type 2 diabetes and depression to illustrate the need for a life course perspective in women's health. Type 2 diabetes, an adult-onset chronic disease, is widely known as a disease of 'modernization' that is emerging in LMICs and shifting from affluent to lower income groups all over the world (6, 7). Biologists and epidemiologists identify depression as both a cause and consequence of diabetes (8, 9), while medical social scientists have elucidated some of the complex socioeconomic and psychophysiological pathways linking the two chronic conditions (10). Despite increasing diabetes prevalence in LMICs, the research on social experiences of those living with diabetes, depression, and their comorbidity is limited. The few existing qualitative studies suggest that experiences differ between men and women (11) as well as between income groups (12).

Social and economic determinants of women's health are fundamental in the relationship of depression and diabetes, particularly among people of lower socioeconomic status (6, 13). As underscored in the 2010 Global Burden of Disease studies, experiences of social problems such as various forms of interpersonal abuse, and psychological problems such as depression and anxiety, have escalated either by detection or actual incidence among women on a global scale (5). Stress throughout the life course rooted in childhood trauma, abuse, or the chronicity of poverty may be key risk factors for depression and/or poor eating and activity patterns that lead to obesity and its complications, such as Type 2 diabetes (10). Complicating matters is the dual burden associated with living in poverty in rapidly modernizing cities that make unhealthy foods accessible and affordable, fueling obesity epidemics in LMIC settings (7). These inequalities create a negative feedback loop, whereby social and economic problems increase the likelihood of developing depression, diabetes, and their overlap, and these illnesses together promote the development of diabetes-related complications such as loss of limbs or eyesight and subsequent physical disability, further compounding socioeconomic inequalities (10). Finally, because of stigma and limited mental healthcare services in LMICs (14), women experiencing this comorbidity are more likely to seek care for diabetes than for depression, leaving half of the comorbidity unaddressed (11).

In India, home to the second largest population of people with Type 2 diabetes in the world (13), recent epidemiological and qualitative data suggest that the illness is becoming more prevalent among the middle classes and working poor (15). In tandem, mental healthcare is limited (16). Despite active research and policies aimed at addressing chronic and mental health diseases in India (17), there remains a large gap in knowledge about how these conditions afflict various Indian communities in their everyday lives, especially poor women. Qualitative research on depression and diabetes in India indicates that lower income people experience higher rates of social stress and depression, and poorer access to health care (12). Such research also underscores the powerful role that gendered social roles play in shaping women's mental health and diabetes outcomes (18). For example, gendered behavioral norms orient Indian women strongly toward the care of others, and therefore away from the self-care activities that are usually integral to diabetes management (11). Maintaining these other-care-oriented roles appears to be good for diabetic women's mental, but not physical, health.

The recognition of social forces as part of diabetes and depression etiology in India and other LMICs presents new challenges for public health because it underscores that medicating these complex illnesses does not fully address them. Finding a better public health solution to comorbidities like Type 2 diabetes and depression will likely only occur when we understand the limitations, and harness the power of, cultural beliefs and social conditions to shape behaviors that affect chronic diseases: how people eat, move, and medicate; how economic conditions may function as a barrier to treatment; and how depression may complicate a chronic disease, both socially and physically (7).

The comorbidity between depression and diabetes among women in LMICs is but one example of the ways in which women's non-reproductive health concerns deserve more prominence in global health. It also presents a strong

case for increased attention to social and psychological determinants of women's health over the life course. The present lack of such perspectives in women's global health may result from limited funding for non-reproductive issues, lack of interest, or may simply be another manifestation of the great information gap between high-income countries and LMICs. Regardless, it should be a priority of future research, programming, and policy.

Focusing on health, not disease

Why should diabetes and depression comorbidity be on the women's health agenda? Depression has only become a major global health concern in the past decade, and has proven very difficult to address, not least because of stigma and limited human resources for mental healthcare. This is especially true for women in LMICs, whose access to mental healthcare may be virtually non-existent, and whose care-seeking behaviors and budgets typically include little, if any, room for mental healthcare. Moreover, most LMICs' health systems are poorly equipped to meet the complex prevention and management challenges associated with chronic conditions like diabetes and mental illnesses because, until very recently, infectious diseases were the dominant population health concerns.

The Movement for Global Mental Health's often-cited slogan, that there is 'no health without mental health' (14), emphasizes the need for integrated mental and physical healthcare systems to combat the next generation of public health problems. This would require an ideological and organizational shift in biomedicine, which has until recently viewed physical and mental health as separate categories of pathology requiring separate treatments, but would likely open up new avenues for cost-effective treatment. The WHO mental health Gap Action Program (mhGAP), for instance, suggests steps by which mental illness diagnosis and treatment can be integrated into primary care settings, and many initiatives are working to actualize this goal in LMICs (e.g. PRIME: http://www.prime.uct.ac.za/). With relatively little additional investment, basic mental healthcare could also be integrated into existing diabetes care guidelines. Such an approach is particularly important for women who face a higher burden of social problems and mental illness, which influence diabetes self-care and health outcomes. Yet, until a more integrative approach is adopted within clinics and public health agendas, healthcare silos will dominate global health dialogues, funding structures, and disease-focused (as opposed to *health*-focused) campaigns.

As the co-occurrence of mental and physical health problems gains recognition in the public health agenda, a more nuanced understanding of sociocultural influences on women's lifetime health is crucial. This is particularly important in LMIC settings where women face not only great social disadvantage but also an increasing burden of mental and physical health problems. A life course perspective requires acknowledging that women's mental and physical health are closely linked with cultural beliefs, social experiences (both past and present), and economic conditions over time. It also recognizes that women's health status shapes their social and economic conditions, for better or worse. Strategic points of intervention can improve women's social and emotional well-being across decades, which could then empower them to identify and care for their own health problems more effectively. In this way, integrating a social and psychological approach into health agendas, from the clinical to the policy level, can make a big impact.

Main findings

- Moving beyond disease-focused silos in public health requires that we attend to the plurality of experiences that shape women's health from social determinants of depression, such as gendered subjugation within the home and public sphere, to the structural determinants of obesity and diabetes, such as poor access to health foods and health care.
- Complexities demonstrated by the comorbidity of depression and type 2 diabetes illustrate the need for a life course perspective in women's health; social and economic factors serve as both causes and consequences of these co-conditions.
- The recognition of social forces as part of diabetes and depression aetiology in low- and middle-income countries presents new challenges for public health because it underscores that medicating these complex illnesses does not fully address them; this requires that we understand the limitations, and harness the power of, cultural beliefs and social conditions to shape behaviors that affect chronic diseases.

Key messages for action

- Integrating a social and psychological approach into health agendas, from the clinical to the policy level, can make a big impact.
- Strategic points of intervention can improve women's social and emotional well being across the life course, which could then empower them to identify and care for their own health problems more effectively.
- With relatively little additional investment, basic mental healthcare (as illustrated in the WHO mental health Gap Action Program (mhGAP)) can be integrated into existing diabetes care guidelines; such an approach is particularly important for women who face a higher burden of social problems and mental illness, which influence diabetes self-care and health outcomes.

Citation: Glob Health Action 2014, **7**: 22803 - http://dx.doi.org/10.3402/gha.v7.22803

Conflict of interest and funding

No conflict of interests declared.

References

1. Knaul FM, Frenk J. Strengthening health systems to address New Challenge Diseases (NCDs). HSPH News. Fall 2011.
2. Lynch J, Smith G. A life course approach to chronic disease epidemiology. Annu Rev Public Health 2005; 26: 1–35.
3. Worthman C, Kohrt B. Receding horizons of health: biocultural approaches to public health paradoxes. Soc Sci Med 2005; 61: 861–78.
4. Patel V, Kleinman A. Poverty and common mental disorders in developing countries. Bull World Health Organ 2003; 81: 609e–15.
5. Murray CJL, 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.
6. Leone T, Coast E, Narayanan S, Aikins A. Diabetes and depression comorbidity and socioeconomic status in low and middle income countries (LMICs): a mapping of the evidence. Global Health 2012; 8: 39–49.
7. Popkin BM, Adair LS, Ng SW. Global nutrition transition and the pandemic of obesity in developing countries. Nutr Rev 2012; 70: 3–21.
8. Golden S, Lazo M, Carnethon M, Bertoni A, Schreiner P, Roux A, et al. Examining a bidirectional association between depressive symptoms and diabetes. JAMA 2008; 299: 2751–9.
9. Musselman D, Betan E, Larsen H, Phillips L. Relationship of depression to diabetes types 1 and 2: epidemiology, biology, and treatment. Biol Psychiatry 2003; 54: 317–29.
10. Mendenhall E. Syndemic suffering: social distress, depression, and diabetes among Mexican immigrant women. Walnut Creek, CA: Left Coast Press; 2012.
11. Weaver LJ. When family comes first: diabetes, social roles, and coping among women in North India. Atlanta, GA: Emory University; 2013.
12. Mendenhall E, Shivashankar R, Tandon N, Ali MK, Narayan KMV, Prabhakaran D. Stress and diabetes in socioeconomic context: a qualitative study of urban Indians. Soc Sci Med 2012; 75: 2522–9.
13. International Diabetes Federation. IDF diabetes atlas. Brussels, Belgium: IDF; 2011.
14. Prince M, Patel V, Saxena S, Maj M, Maselko J, Phillips MR, et al. No health without mental health. Lancet 2007; 370: 859–77.
15. Deepa M, Anjana R, Manjula D, Narayan K, Mohan V. Convergence of prevalence rates of diabetes and cardiometabolic risk factors in middle and low income groups in urban India: 10-year follow-up of the Chennai Urban Population Study. J Diabetes Sci Technol 2011; 5: 918–27.
16. Patel V. The future of psychiatry in low- and middle-income countries. Psychol Med 2009; 39: 1759–62.
17. Patel V, Chatterji S, Chisholm D, Ebrahim S, Golapakrishna G, Mathers C, et al. Chronic diseases and injuries in India. Lancet 2011; 377: 413–28.
18. Weaver LJ, Hadley C. Social pathways in the comorbidity between Type 2 diabetes and mental health: concerns in a pilot study of urban middle- and upper-class Indian women. Ethos 2011; 39: 211–25.

Global Health Action

PART III

Understanding epidemiological transition in India

Suryakant Yadav* and Perianayagam Arokiasamy

International Institute for Population Sciences, Mumbai, India

Background: Omran's theory explains changing disease patterns over time predominantly from infectious to chronic noncommunicable diseases (NCDs). India's epidemiological transition is characterized by dual burden of diseases. Kumar addressed low mortality and high morbidity in Kerala, which seems also to be true for India as a country in the current demographic scenario.

Methods: NSS data (1986–1987, 1995–1996, 2004) and aggregated data on causes of death provided by Registrar General India (RGI) were used to examine the structural changes in morbidity and causes of death. A zero-inflated poisson (ZIP) regression model and a beta-binomial model were used to corroborate the mounting age pattern of morbidity. Measures, namely the 25th and 75th percentiles of age-at-death and modal age-at-death, were used to examine the advances in mortality transition.

Objective: This study addressed the advances in epidemiological transition via exploring the structural changes in pattern of diseases and progress in mortality transition.

Results: The burden of NCDs has been increasing in old age without replacing the burden of communicable diseases. The manifold rise of chronic diseases in recent decades justifies the death toll and is responsible for transformation in the age pattern of morbidity. Over time, deaths have been concentrated near the modal age-at-death. Modal age-at-death increased linearly by 5 years for females ($r^2 = 0.9515$) and males ($r^2 = 0.9020$). Significant increase in modal age-at-death ascertained the dominance of old age mortality over the childhood/adult age mortality.

Conclusions: India experiences a dual burden of diseases associated with a remarkable transformation in the age pattern of morbidity and mortality, contemporaneous with structural changes in disease patterns. Continued progress in the pattern of diseases and mortality transition, accompanied by a linear rise in e_x, unravels a compelling variation in advances found so far in epidemiological transition witnessed by the developed nations, with similar matrices for India.

Keywords: *noncommunicable diseases*; *communicable diseases*; *disease patterns*; *mortality transition*; *epidemiological transition*

Responsible Editors: Nawi Ng, Umeå University, Sweden; Barthélémy Kuate Defo, University of Montreal, Canada.

*Correspondence to: Suryakant Yadav, International Institute for Population Sciences, Mumbai, India, Email: suryakant11@gmail.com

This paper is part of the Special Issue: *Epidemiological Transitions – Beyond Omran's Theory*. More papers from this issue can be found at http://www.globalhealthaction.net

Received: 2 November 2013; Revised: 26 January 2014; Accepted: 3 February 2014; Published: 15 May 2014

The parallel processes of demographic and epidemiological transition are currently occurring at remarkable speed in India. From a comparative perspective, in the absence of sufficient longitudinal data from the vital registration system, the study of epidemiological transition has received much less attention to adequately understand the major shifts in mortality and morbidity patterns (1–3). Since the 1970s, researchers have primarily focused on India's demographic transition due to the availability of good long-term demographic trend data on mortality, fertility, and population growth from a sample registration system (SRS) (4–9). Aside from this, the limited number of studies that have addressed the broader domain of India's health transition (comprising mortality and epidemiological transition) have been narrowly focused because of data constraints. In general, the historical trends in mortality conditions of low- and middle-income countries have varied from those of high-income countries. For example, in underdeveloped nations such as India, mortality began to fall (10, 11) around the 1950s but accelerated through the 1970s and 1980s. For example, although the infant mortality rate (IMR) fell by two-thirds, from 129 to 44/1,000 live births, the crude death rate was reduced by half, from 14.9 to 7.1/1,000 persons, during the period 1970–2011 (12, 13). In an era of such emerging trends, India and its states have transitioned from high mortality conditions to medium/low mortality conditions in the last three decades.

Accumulated data from multiple sources suggest that, consistent with the current phase of demographic and epidemiological transition, the pace of 'India's Health Transition' has been swift (14, 15). Furthermore, the

Citation: Glob Health Action 2014, **7**: 23248 - http://dx.doi.org/10.3402/gha.v7.23248

progress in health transition has spread over all geographical regions of India where the rise in morbidity is accompanied by a marked fall in mortality. Nevertheless, data also suggest notable heterogeneity among states with the highest morbidity rates (such as 255/1,000 persons in Kerala) and the lowest (in poorer states such as Jharkhand where the rate is 33/1,000 persons). Further data from India's National Sample Survey Organization (NSSO) revealed an enormous increase in India's morbidity level during the last two decades (16–18). The rise of noncommunicable diseases (NCDs) has been associated with risk factors such as low physical activity, use of tobacco, low intake of fruits and vegetables, high body mass index (BMI), and so on. However, the government of India (GOI) has recognized that, irrespective of these risk factors and of socioeconomic status, NCDs are extremely common among older people. Between 1995–1996 and 2004, the crude morbidity rate increased by more than 60 and 90%, respectively, among rural and urban populations. In both rural and urban areas, the rise in morbidity level has been common across demographic and social spectrum:—among females, males, social groups, monthly per capita expenditure (MPCE) classes, and so on (16, 19).

In order to understand India's progress in epidemiological transition, in this paper we assess that country's structural changes in patterns of morbidity and mortality. Omran's (20, 21) theory of epidemiological transition provides a useful basis for examining such shifts in mortality and morbidity. The theory describes the process of changing disease patterns over time—predominantly from infectious to chronic NCDs (20, 21). However, the theory has been benchmarked primarily based on the experiences of developed nations where the burden of chronic NCDs has replaced the burden of communicable diseases. Contrastingly, in developing countries such as India, communicable diseases have been an additional burden due to the mounting number of NCDs (15, 16, 22). As a result of such confounding patterns (with e_0 equaling 66.1 years in 2006–2010), epidemiological transition in India may be moving through the *Age of Receding Pandemics* and the *Age of Degenerative and Man-Made Diseases* (23). Progression to a particular stage of epidemiological transition is important to investigate while India is in the midst of this swift transition.

Overall, India has been experiencing rapid structural changes in disease patterns within the short span of the last three decades. Advances in mortality and morbidity transition to later stages point toward an upheaval in epidemiological transition. Visaria (15) has brought to light the phenomenon of the dual burden of disease and its linkages with epidemiological transition. In rural India, the burden of NCDs increased from 35.9 to 54.9%, and the burden of communicable diseases declined from 47.7 to 22.1%, during the transitional period from the 1970s to the mid-1990s (15). Although the overall burden of communicable diseases has declined, this burden is still very significant. More importantly, the share of communicable diseases has not been entirely replaced by chronic NCDs. In addition to Visaria's study (15), John et al. (24) have recently demonstrated the heavier burden of communicable diseases, which is as high as 30% of the total. Therefore, India's current stage of epidemiological transition can be characterized by low mortality, high morbidity, and by the double burden of communicable diseases and NCDs. These broad mortality and morbidity trends suggest that India represents a major contrast in the process of epidemiological transition particularly in light of conclusions based on studies of developed nations.

Viewed from an analytical context, the structural changes in disease patterns are concomitant with the transformation in the age pattern of morbidity and mortality. The age pattern of mortality (25) has been flattening over decades; however, the age pattern of morbidity is not commensurate with the transformation in age pattern of mortality. Rather, in recent decades, the age pattern of morbidity has been mounting as a result of the increasing prevalence of NCDs and communicable diseases (19, 26, 27). Altogether, these intervening components stimulate the progression of mortality transition in the country (28) thereby increasing the progress of epidemiological transition. This fundamental structural shift possibly signals advances in epidemiological transition (26, 29, 30). Researchers have acknowledged this phenomenon and have studied this variation in epidemiological transition in India (an underdeveloped nation) compared to the progress of epidemiological transition so far observed in developed nations. Regardless, rarely any of the studies have addressed the India's variation in its progress of epidemiological transition from the globally established course of epidemiological transition. Therefore, it is critical to fill this theoretical research gap for underdeveloped nations. This paper attempts to accomplish this by investigating fundamental processes such as the age pattern of morbidity, structural changes in disease patterns, and mortality transition to assess the course of India's epidemiological transition. The specific objectives explored are: 1) the prevalence of chronic diseases among the population aged 60 and above, 2) structural changes in causes of death, and 3) transformation in distribution of age-at-death and modal age-at-death. The first two objectives explain the correlation between morbidity and causes of death, and the third objective connects it with mortality transition. Broadly, the study addresses whether the structural changes in disease patterns conform to the progression in mortality transition and hence to the advances in epidemiological transition.

Background

The burden of communicable diseases had been dominant around the world until the mid-twentieth century.

In Africa and Asia (including India), the trilogy of smallpox, plague, and malaria remained the 'top killer' diseases until the 1960s. However, by the late 1970s, smallpox was eradicated in India through sustained vaccination programs; better sanitation and housing reduced the overall burden of communicable diseases (31). Nevertheless, by the mid-1990s, other communicable diseases, such as tuberculosis of lungs (5.3%), gastroenteritis/dysentery (3%), and pneumonia (4.7%) were responsible for a considerable number of deaths, signifying the burden of communicable diseases. Overall, the manifold rise in morbidity rates in recent decades concomitant with a decline in mortality rates is a manifestation of the progression of 'health transition' (32); this is a phenomenon where the survival rate of the population led to an increase in the level of morbidity and a decline in mortality. Kumar (14) explored the rapidly advancing phenomenon of 'health transition' in Kerala because this state had been experiencing consistently high morbidity prevalence levels with the steep fall in mortality (14). The phenomenon expedites the process of expansion of morbidity, which is found to be true at the national level by residence and by sex (26, 33).

The advances in mortality transition were explored by Ranjan Chaurasia (28) for the period 1970–2005 based on a measure of 'entropy' in life expectancy tables. He found the level of mortality continues to be high based on international standards. Among females, the pace of mortality transition slowed down. Comparatively, among urban males, mortality transition was found to be comparatively faster.

Today, modal age-at-death is widely used for testing mortality transition. Canudas-Romo (34), while testing the shifting mortality hypothesis, studied the modal age-at-death using various models, such as the Gompertz, Logistic, Siler, and Log-Siler models. The Siler model revealed a rise in modal age-at-death alongside a shift in distribution of age-at-death, although the Gompertz model only showed a shift in the distribution of age-at-death. The earlier models assume a reduction in IMR over time, while the later model assumes IMR is at the lowest possible value. The results were affirmative; modal age-at-death addressed the change in mortality. Nevertheless, a declining IMR and adult mortality rate have an impact on the modal age-at-death and on subsequent transformation in distribution of age-at-death. Modal age-at-death, similar to entropy and life expectancy, differentiates the pace of mortality transition among different categories of population. Additionally, modal age-at-death is a measure of length of life and is relevant for measuring dispersion of distribution of age-at-death (35). Modal age-at-death is useful for understanding changes in life expectancy and corresponding changes in the age pattern of mortality. The reduction in dispersion above the modal-age-at-death accompanied by the concentration of death in a narrower age-interval indicates a high to low mortality transition. Trends in C50—a measure of mortality compression—and modal age-at-death demonstrated virtual convergence among the developed countries in the last stage of epidemiological transition (35–37). Modal age-at-death is sensitive, simple, and convenient to use and has multiple applications, and therefore is preferred over other measures. Accordingly, it has been used globally to assess phenomena such as mortality transition and demographic and epidemiological transition. It is worthwhile to mention that the transformation in distribution of age-at-death is a simpler way to address phenomena such as mortality compression, inequality in age-at-death, and mortality transition (38).

Data and methods

In this study, we have used multiple data sources and integrated them together for national level analysis. First unit level cross-sectional NSSO data for 1986–1987, 1995–1996, and 2004 on morbidity of aged (60+) persons were used to understand the changes in the prevalence rate of chronic diseases over time. The NSSO has retained the specific section on morbidity and ailing persons for the aged population for the three time periods (1986–1987, 1995–1996, and 2004); however, data on morbidity/NCDs for all ages is available only for the time periods 1995–1996 and 2004. The coverage of chronic diseases specific to the aged population remained same for the three time periods. Also, the prevalence of chronic diseases has escalated remarkably among the aged population signaling rapid changes in the morbidity profile of India (17, 18, 26). Therefore, morbidity analyses were done for the population aged 60 and above for the three data points. The prevalence rate of chronic diseases—hypertension, joint and bones, asthma, heart disease, cancer and other tumors, urinary problems, and diabetes—was adjusted for age, sex, residence, living alone, dependency, hospitalization, education, MPCE, and region (north, east, northeast, west, south, and central) using a ZIP regression model. We used a beta-binomial model to examine the changes in summary event rate of chronic diseases and also in total NCDs over time for the older population. The occurrence of chronic diseases/NCDs in a household of an older person age 60 or above with a reference period of the last 15 days will be a success, otherwise failure of the event. Thus, the probability of the occurrence of chronic diseases/NCDs for an older person in a household follows binomial distribution. However, the probability of occurrence of chronic diseases/NCDs across households varies and follows a beta distribution. Therefore, the application of a compound distribution —beta-binomial distribution—allows estimation of occurrence probability for the summary event rate of chronic diseases/NCDs for the older population across households (39). NSS 1986–1987,

Citation: Glob Health Action 2014, **7**: 23248 - http://dx.doi.org/10.3402/gha.v7.23248

1995–1996, and 2004 provide information for 8,478; 7,016; and 18,261 households, respectively, for ailing older persons aged 60 and above.

Second, we used aggregated data by broad age groups from a survey of causes of death (SCD) for rural areas (40) and mortality statistics of causes of death (MSCD) for urban areas (41) to understand the structural changes in causes of death. Medical certification of causes of death (MCCD) provided mortality statistics mostly for urban areas (42). For the period 2001–2003, RGI provided mortality statistics for both rural and urban populations (22). These data have been used together to construct the distribution of deaths attributable to communicable diseases and to NCDs for rural and urban India, respectively, and to examine the transformation in distribution of deaths.

Third, an age specific death rate (ASDR) provided by RGI (25) was used to construct new life tables based on the methodology proposed by the United Nations (43–45). The ${}_5a_x$ values of the life tables were very close to 2.5, which confirms the uniform distribution of age-at-death in the age groups. Therefore, the d_x column of the life table, for the age group 10 and above (46), was disaggregated into single years using the King-Karup method (47). Furthermore, the single-year distribution of age-at-death was smoothed using the cubic spline method to remove the erratic fluctuations (48, 49). Henceforth, the truncated and smoothed distribution is referred to as distribution of age-at-death. The transformation in distribution of age-at-death, trends in the 25th and 75th percentiles of age-at-death and at modal age-at-death are mainly examined to understand the progress of mortality transition in India (34, 35). The 25th and 75th percentiles of age-at-death refer to the age corresponding to 25 and 75% of deaths in distribution of age-at-death; modal age-at-death refers to the age corresponding to the modal value of the distribution of age-at-death.

Results

In this section, we present the results of our analysis on 1) structural changes in morbidity and causes of death and 2) transformation in distribution of age-at-death and modal age-at-death. The results overall describe the linkage between morbidity pattern and causes of death and the consequent progress in epidemiological transition.

Structural changes in morbidity

Figure 1 displays the age pattern of chronic diseases—by residence and by sex—for 1986–1987, 1995–1996, and 2004. The age pattern of morbidity reveals a mounting concentration of morbidity prevalence in the 60–64 and older age groups. The rising gradient of morbidity prevalence in the older ages peaked at the old–old (70–79) ages of 75–79. As a result of such a sudden and steep rise in the prevalence of chronic diseases, the overall morbidity began to take a distinguishing shape in 1995–1996 compared to the pattern observed among the developed nations. The prevalence rate of chronic diseases

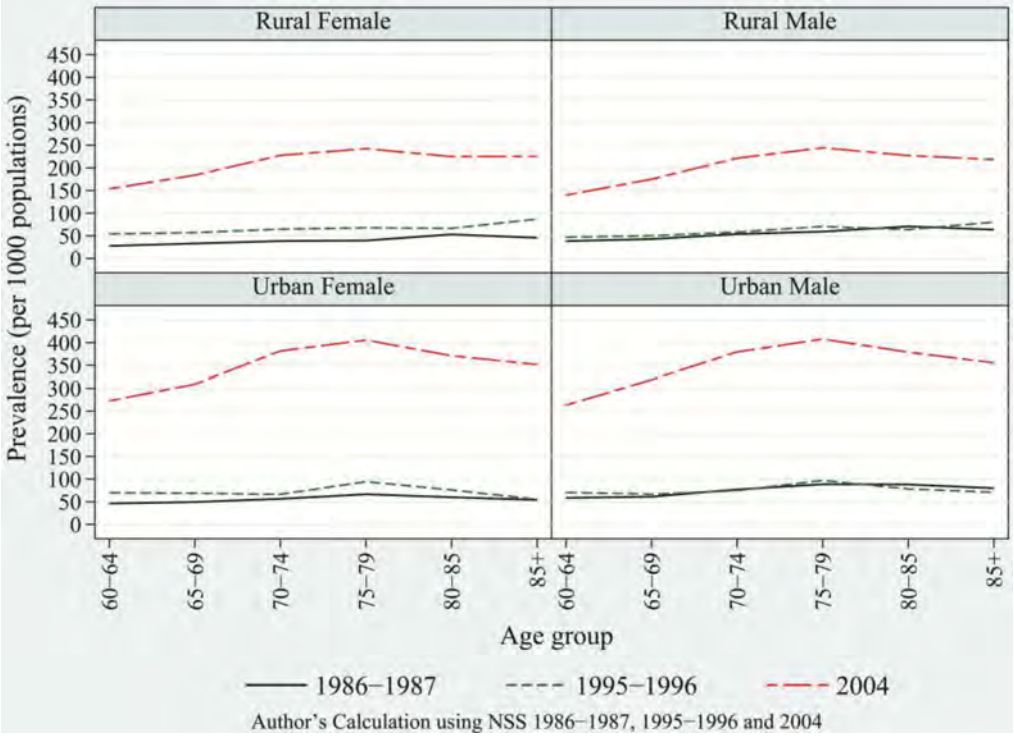

Fig. 1. Age pattern of prevalence rate of chronic diseases by sex and residence, India for 1986–1987, 1995–1996, and 2004.

for ages 60+ increased from 48.5 in 1986–1987 to 69.5 in 1995–1996 and rose to 260.8 in 2004; overall, this indicated a more than five-fold rise in the prevalence of chronic diseases. Table 1 provides trends in the prevalence rate of each chronic disease among the aged population by sex and residence. By sex and residence, the prevalence rate of chronic diseases was highest among urban males and was characterized by a high prevalence among old–old and oldest of old (ages 80 and above). Among young–old (ages 60–69), the prevalence rate was relatively low compared to old–old and oldest of old. The prevalence rate for cardiovascular diseases (CVDs), which includes heart disease and hypertension, increased from 14.9 in 1986–1987 to 19.5 in 1995–1996 and escalated to as high as 63 in 2004, representing four-fold rise during the period of 1986–1987 to 2004.

In general, the age pattern of communicable diseases does not correspond with the changing age pattern of chronic disease. This is because the change in the prevalence rate of communicable diseases has not been sufficient to mark any structural changes in the overall age pattern of morbidity (26). In contrast, the enormous rise of degenerative diseases compared to any other categories of diseases, and the increased fatality rates due to them, have been largely responsible for deaths in older ages and, consequently, have contributed to the structural changes in disease patterns. The overall rise in degenerative diseases in recent decades indicates progress in the epidemiological transition to a later stage.

It is important to validate the enormous rise of chronic diseases because a remarkable increase in chronic NCDs has been revealed. The beta-binomial model was tested to estimate the summary event rate of chronic diseases, which takes into account the variation in the chance of occurrence of chronic NCDs across households and the binomial distribution followed by chronic diseases in each household (Table 2) (39). The summary event rate of chronic diseases increased from 4.01% (0.08%) in 1986–87 to 4.6% (0.22%), and to 12.7% (0.23%) in 2004. The trends in summary event rate of chronic diseases reveal a steep rise between 1995–1996 and 2004 compared with the marginal increase between 1986–1987 and 1995–1996. The parameters of the beta-binomial model give alpha and beta estimates of 0.05 and 1.18, respectively, for 1995–1996 and 0.16 and 1.1, respectively, for 2004. The likelihood ratio test for overdispersion was significant at the 1% level of significance for both years and thus confirmed the applicability of the beta-binomial model over the binomial model. The test was done for NCDs for 1995–1996 and for 2004 (3). The summary event rate of chronic NCDs rose by more than three times between 1995–1996 and 2004, similar to the manifold rise of chronic diseases.

The parameters of the beta-binomial model and the test of overdispersion are given in Table 3. Between 1995–1996 and 2004, chronic diseases and NCDs show a similar manifold rise that corroborates the rises in the prevalence rate obtained from the regression model. In sum, the several-fold rises of chronic diseases certainly transformed the age pattern of morbidity. This transformation in the age pattern of morbidity has been rapid since the mid-1990s. In recent decades, degenerative diseases have been responsible for poor health in old ages with an increase in life expectancy in old ages (26). Accordingly, they have been accountable for the heavy burden of deaths in old ages. This phenomenon is indicative of the increasing share of deaths due to NCDs and of the corresponding advances in epidemiological transition.

Table 1. Prevalence rate of chronic diseases among older population by residence and sex, India, 1986–1987, 1995–1996, and 2004

	1986–1987				1995–1996				2004			
	Rural		Urban		Rural		Urban		Rural		Urban	
Chronic diseases	F	M	F	M	F	M	F	M	F	M	F	M
Asthma/cough and acute bronchitis*	10.3	17.9	11.2	19.4	14.5	18.4	8.4	12.2	19.2	33.7	18.8	30.6
Problems of joints and bones	10.7	11.2	9.0	8.3	27.6	18.0	22.3	14.0	52.9	33.3	61.1	41.1
Hypertension	7.1	6.7	19.8	18.8	11.1	6.5	20.4	16.1	28.6	21.4	80.0	60.2
Heart disease	2.1	3.3	3.8	6.3	3.4	5.7	7.0	15.4	10.0	14.0	31.7	45.3
Urinary problems	0.5	1.3	0.8	2.1	1.7	2.6	2.2	3.7	3.3	6.8	3.1	7.3
Diabetes	1.3	2.2	4.3	7.3	3.7	4.7	12.1	15.5	14.3	15.4	49.6	57.3
Cancer and other tumors**	0.4	0.5	0.6	0.7	1.5	1.8	1.7	1.8	3.6	2.4	4.6	3.2

Source: Author's calculation from NSS 1986–1987, 1995–1996, and 2004; F: female; M: male.
*In 1995–1996, information was provided for cough and acute bronchitis. In 1986–1987 and 2004, information was provided for asthma.
**In 1986–1987, information was provided for cancer only. In 1995–1996 and 2004, information was provided for cancer and other tumors.

Citation: Glob Health Action 2014, **7**: 23248 - http://dx.doi.org/10.3402/gha.v7.23248

Table 2. Summary event rate for chronic diseases, India, 1986–1987, 1995–1996, and 2004

Year	Event rate (SE)	%	Dispersion (SE)	%	α	β	2 (LBB-LB)
1986–1987	4.01	0.08	60.70	4.50	0.06	1.57	870.85***
1995–1996	4.60	0.22	80.30	9.20	0.05	1.18	440.65***
2004	12.70	0.23	79.10	5.30	0.16	1.10	930.42***

Source: Author's calculation from NSS 1986–1987, 1995–1996, and 2004. LBB: Log Likelihood of Beta Binomial model; LB: Log Likelihood of Binomial model.
***Significant @ 1% level of significance.

Structural changes in causes of death

Structural changes in overall disease patterns over a period of 30 years and more, with the concomitant transformation in the age pattern of morbidity and mortality, resulted in significant structural changes in causes of death (50–53). During the 1970s and 1980s, infectious and parasitic diseases were the dominant cause of death in India. Among rural populations, diarrhea (diseases of the digestive system), cough (disorders of the respiratory system), and fever were responsible for 8, 23 and 15.5%, respectively, of the total deaths in 1972. The share of these diseases declined to 36 and 31% (40), respectively, in 1982 and 1997 with a considerable share of deaths attributable to tuberculosis of the lungs, gastroenteritis/dysentery, and pneumonia. Among the urban population, infectious and parasitic diseases, diseases of the respiratory system, and diseases of the digestive system, were responsible for 26, 10 and 6%, respectively, of total deaths in 1975. In 1985, these diseases accounted for 36.4% of total deaths, but this figure declined to 28.1% in 1995 (41). Among the urban population, the burden of respiratory tuberculosis and pneumonia accounted for an average of 5 and 3%, respectively, of total deaths from 2001 to 2004 (42).

Over time, with the modest decrease in the burden of communicable diseases, NCDs emerged as a major cause of death. The mortality burden attributed to NCDs has been increasing over time without replacing the burden attributed to communicable diseases. Among rural populations, the burden of NCDs increased from 12% in 1977 to 13% in 1982 and 28% in 1997 (40). The decade of the 1990s witnessed a major rise in the burden of NCDs. Among urban populations, the share of NCDs increased from 29% in 1975 to 35% in 1985 and 36% in 1995 (41). In general, the burden of NCDs among urban populations as compared to rural populations has been consistently higher. Among urban populations, the major changes in disease patterns were caused by infectious and parasitic diseases, which declined steeply by 10% during 1975–1995, and diseases of the circulatory system, which increased asymptotically to the highest share and engrossed the largest share of 21% of total burden of diseases in 1995 (Figs. 2 and 3).

Table 3. Summary event rate for noncommunicable disease, India, 1995–96 and 2004

Year	Event rate (SE)	%	Dispersion (SE)	%	α	β	2 (LBB-LB)
1995–1996	5.10	0.22	77.20	8.60	0.06	1.23	439.08***
2004	16.00	0.25	79.20	4.90	0.20	1.06	930.42***

Source: Author's calculation from NSS 1995–1996 and 2004.
***Significant @ 1% level of significance.

Diseases of the circulatory system have been the prominent killers in recent decades as compared to infectious and parasitic diseases, which were the top killer diseases until the 1980s. In contrast to urban populations, throughout the period 1972–1997, rural populations' top killer diseases had been those of the respiratory system, though the rate has been declining over time. Apart from these prominent causes of deaths, injuries, poisoning, and death caused by other external factors have been significantly responsible for changes in disease patterns. Among rural and urban populations, the burden of this category of diseases increased from an average of 4 and 8.5%, respectively, in early1970s to an average of 11 and 12%, respectively, in late 1990s. In later half of 1990s, among both rural and urban populations, males in general were at greater risk of dying than females. In particular, there were more deaths among males due to infectious and parasitic diseases and to diseases of the circulatory system. These have been the prominent killer diseases, and they have largely shaped the disease patterns.

The recent report on causes of death from 2001–2003 provides percentage share of causes of death by major categories of diseases. The results assert that CVDs are the top killer diseases in rural as well as urban areas and more profoundly among urban males. By age groups, for the age groups of 25–69 and 70+, deaths due to CVDs were highest among urban males. For example, deaths due to chronic NCDs were higher by 4 and 2%, among urban males aged 25–69 and 70+, respectively, than among urban females (22). In 2006, studies by Joshi et al. conducted in the eastern and western Godavari regions of Andhra Pradesh demonstrated similar findings. Death data were recorded using a verbal autopsy method from a well-established cause of death surveillance system for a rural area. Joshi et al. affirm chronic NCDs as the leading cause of death in this area, which indicates the rapid progress of epidemiological transition in rural India (30).

Figures 4 and 5, for rural and urban population, respectively, show the distribution of deaths attributable to communicable diseases and NCDs by broader age groups separately by residence for selected years (40, 41). For each category of disease by residence, the distribution

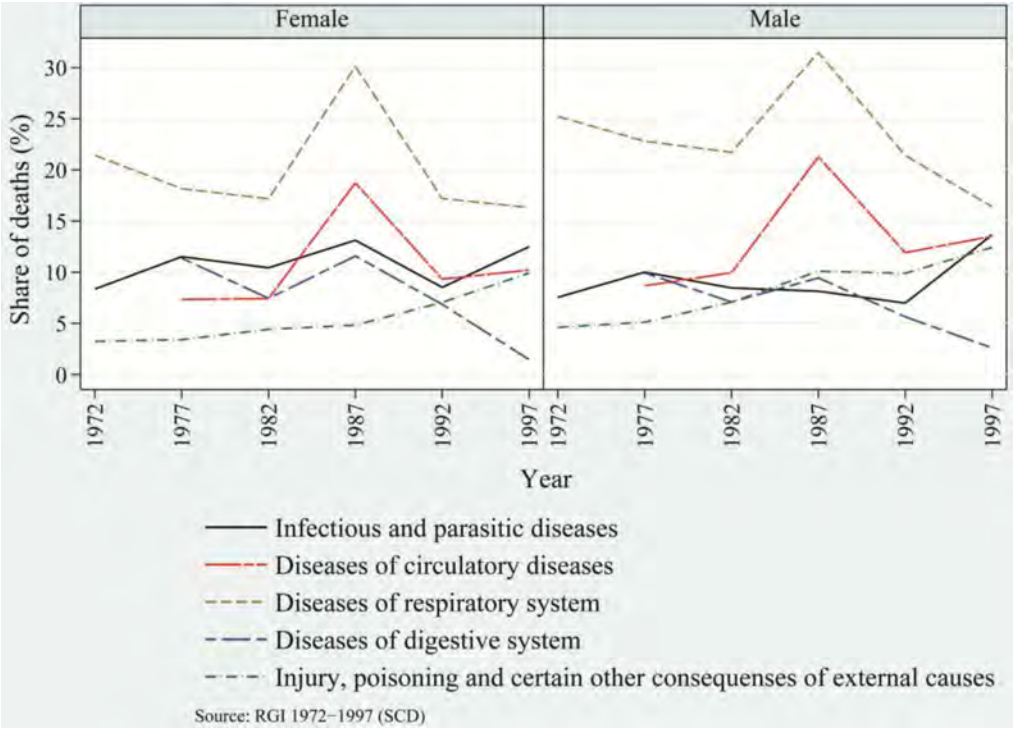

Fig. 2. Share of major causes of death among females and males, rural India, 1975–1995: Infectious and parasitic diseases, circulatory diseases, respiratory diseases, diseases of the digestive system, injury, poisoning, and certain other external causes, such as senility.

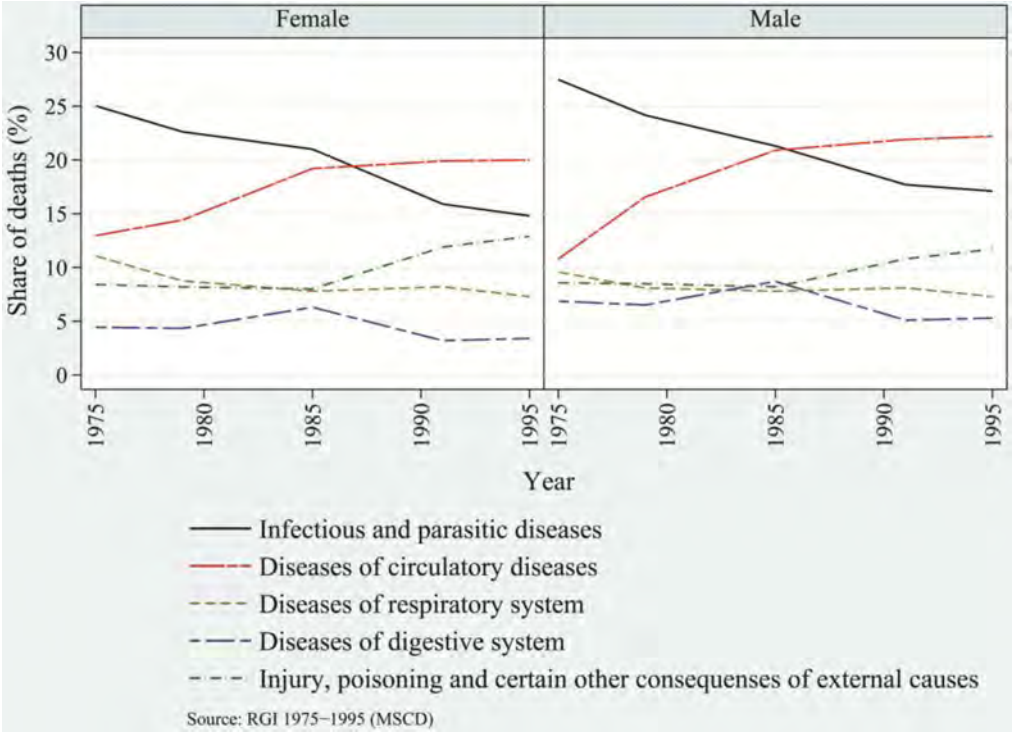

Fig. 3. Share of major causes of death among females and males, urban India, 1972–1997: Infectious and parasitic diseases, circulatory diseases, respiratory diseases, diseases of the digestive system, injury, poisoning, and certain other external causes.

Citation: Glob Health Action 2014, **7**: 23248 - http://dx.doi.org/10.3402/gha.v7.23248

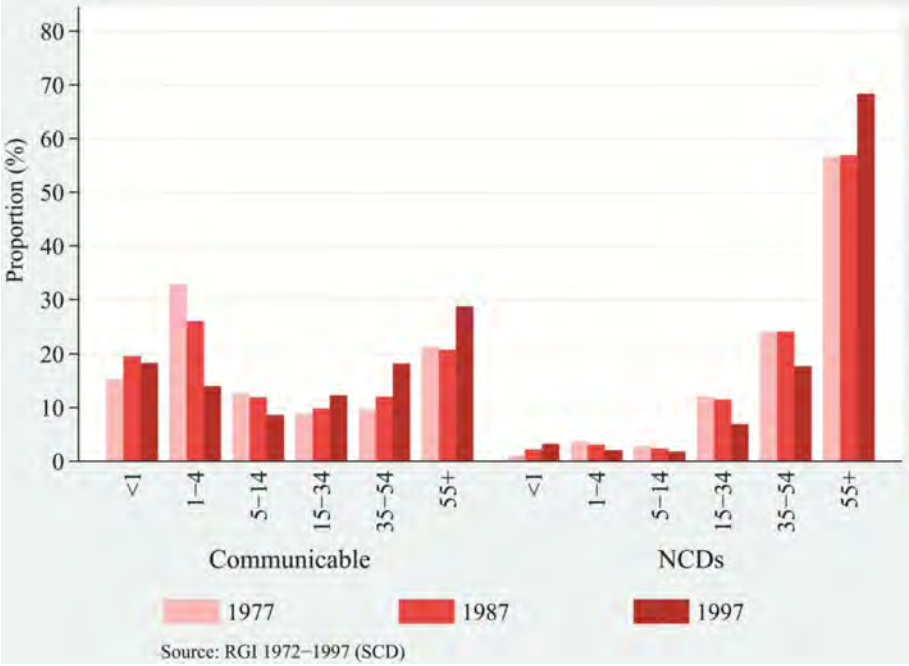

Fig. 4. Distribution of deaths attributed to communicable diseases and NCDs, by broad age groups, rural India, 1977, 1987, and 1997.

of death is comparable. Over time, the transformation in the distribution of deaths is apparent. Among rural and urban populations, the burden of deaths due to communicable diseases had been dominant for children aged 0–4. In general, among the adults and the aged population (15 and above age groups), the proportion of deaths due to communicable diseases had been increasing over the time period (Figs. 4 and 5, left panel). The proportion of deaths in adult and old ages was prominent between the mid-1980s and the mid-1990s, indicating faster progress in the shifting of deaths to higher ages from lower ages.

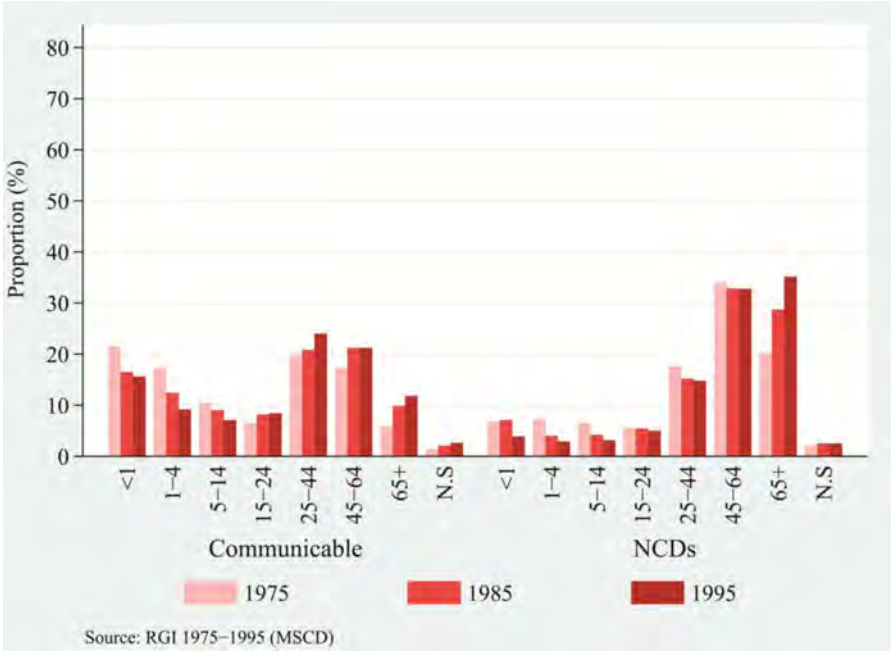

Fig. 5. Distribution of deaths attributed to communicable diseases and NCDs, by broad age groups, urban India, 1975, 1985, and 1995.

Nevertheless, the transformation in distribution of deaths attributed to communicable diseases seemed modest. Relatively, the transformation in distribution of deaths attributed to NCDs seemed prominent. The shape of distribution of deaths attributed to communicable diseases had been relatively uniform across the ages in comparison to the highly skewed distribution of death attributed to NCDs, depicted in Figs. 4 and 5. Compared with the lower burden of NCDs for children's age groups, older rural and urban populations bore a comparatively higher burden of NCDs. Among rural and urban populations, deaths had been increasingly concentrated in the ages of 55 and above and of 45 and above, respectively (Figs. 4 and 5, right panel). The transformation in distribution of deaths attributed to NCDs unravels a much larger proportion of deaths drifting toward old ages. Comparatively, urban populations experienced a higher burden of NCDs than rural populations over a wide range of ages. There was a rapid increase in the concentration of deaths in older ages from the mid-1980s and the mid-1990s, signaling a rapid transformation in distribution of deaths attributable to NCDs. The fall of mortality rates in adult ages and the resulting increase to 8% in 2011 (54) in the proportion of deaths among the aged population demonstrate that the marginal lives in old ages are prone to an elevated risk of dying. Therefore, the age structural transition supplements the transformation in age-at-death and structural changes in disease patterns.

These results confirm that India did not see any fall in the burden of diseases with the progress of demographic and epidemiological transition, but that it moved to a more challenging stage of double burden of disease dominated by burden of NCDs in old ages. However, as may be seen in the progressive stage of epidemiological transition, the rise in the burden of NCDs tends to correspond with the enormous rise in the prevalence rate of chronic diseases. This empirical evidence supports the structural changes in causes of death vis-à-vis transformation in the age pattern of morbidity and mortality. The growing burden of chronic diseases in older ages is a manifestation of structural changes in the disease patterns. These changes in the age patterns of morbidity and in the causes of death structures signal progress in epidemiological transition of India, albeit as a notable variant from the pattern observed for developed countries.

Mortality transition in India

India experienced a significant rise in life expectancy at birth (e_0) during the reference period of the study. The e_0 of India among females and males–increased by 14 and 19 years, respectively, from 1970–75 to 2006–10 (13, 23), signifying marked improvement in the health status of the population. The states of India display more insightful patterns. Among rural females of Kerala, life expectancy reached 77.2 years in 2006–2010, which is no less remarkable progress compared to developed nations. Over time, the IMR and CDR have fallen considerably. In addition, the adult mortality rate fell from 274 to 213 (per 1,000 populations) during 1990–2008 (55). Hence, the population is surviving to higher ages, and as a result, deaths are occurring at higher ages. Aside from this, the growing burden of chronic diseases in older ages and the concomitant structural changes in the disease patterns impel the transformation in the age pattern of mortality.

Changes in the age pattern of mortality are better understood in terms of distribution of age-at-death. In principle, transformation in distribution of age-at-death complies with the transformation in age pattern of mortality (38). Over time, the fact that deaths are occurring at higher ages is apparent (Fig. 6). Of all the ages, deaths tend to be concentrated near the mode of the distribution of age-at-death, which generally falls in the old ages. As this process progressed, the distribution of age-at-death tended to be unimodal in the 1990s, instead of the bimodal pattern seen in the 1970s and 1980s; this resulted in a definite shape to the distribution of age-at-death. The drifting of deaths toward higher ages and the concomitant concentration of deaths in old ages turned the distribution of age-at-death to a bell-shaped curve in recent decades. This transformation in distribution of age-at-death displays the impressive progression of mortality transition and its ramifications on epidemiological transition.

The progression in mortality transition can be examined by using relevant measures such as the 25th percentile of age-of-death, the 75th percentile of age-at-death, and the modal age-at-death. Table 4 provides the trends of the 25th and 75th percentiles of age-at-death by population categories: rural-urban and female-male at intervals of five years during the period of 1970–2007. Among urban females, the 25th and 75th percentiles of age-at-death increased linearly from age 60 and age 80, respectively, in 1970–1974; and from age 68 and age 85, respectively, in 2003–2007. Among urban males, the 25th and 75th percentiles of age-at-death increased linearly from age 57 and age 78, respectively, in 1970–1974; and from age 62 and age 82, respectively, in 2003–2007. Similar linear increases in percentiles of age-at-death were observed for rural females and males. Though the linearity was similar for all population categories, the 25th and 75th percentiles of age-at-death for urban females remained consistently higher throughout the time period when compared to other categories of population.

As supported by the findings of Wilhelm Lexis in 1877 (49), life table deaths move up in the left-hand slope of the distribution of age-at-death, implying a reduction in premature mortality. Among rural and urban females, the 25th percentile of age-at-death shifted to the right by 8 and 10 years, respectively. This shift had been passive

Citation: Glob Health Action 2014, **7**: 23248 - http://dx.doi.org/10.3402/gha.v7.23248

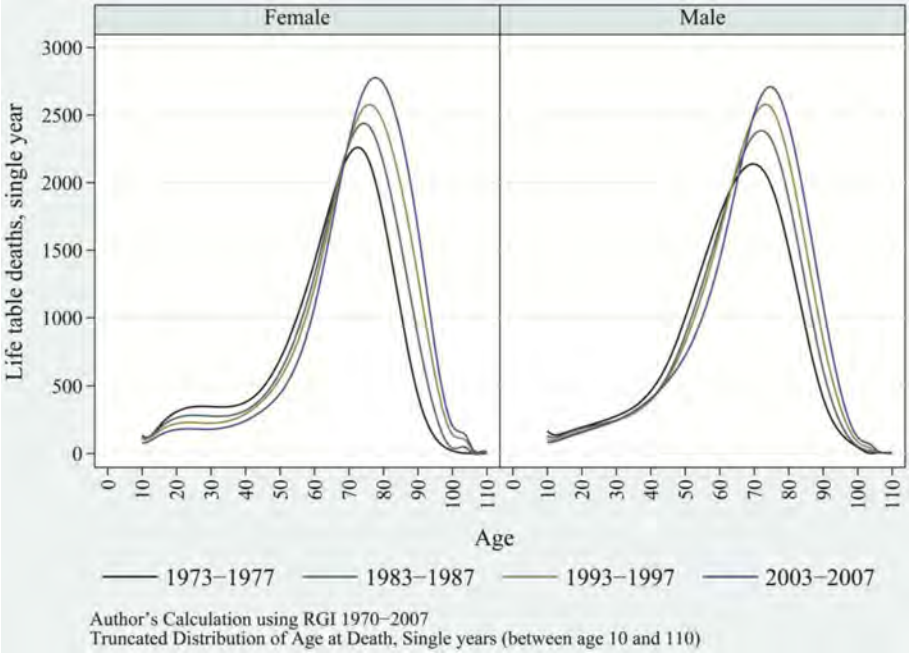

Fig. 6. Distribution of age-at-death and modal age-at-death for selected years, female and male, India, 1973–1977, 1983–1987, 1993–1997, and 2003–2007.

among males as compared to females. As life table deaths drifted to higher ages, the left-hand slope of the distribution of age-at-death tends to be more vertical, moving most deaths near the modal age-at-death. Therefore, greater reduction in premature mortality, or alternatively the rapid fall in adult mortality, results in the concentration of deaths in old ages and the overall transformation in age-at-death. Correspondingly, the right-hand slope of the distribution of age-at-death also tends to be vertical; however, this shift is weaker compared to developed nations (Fig. 6) (35).

As a result of this process, most deaths are concentrated near the modal age-at-death, leading to rises in the modal values of distribution of age-at-death. With the rise in modal values, the modal age-at-death increased linearly, $r^2 = 0.9615$ for females and $r^2 = 0.9120$ for males. Among urban females and males, modal age-at-death in 1970–1974 increased linearly from age 75 and age 71.5, respectively, to age 80.5 and age 76, respectively, in 2003–2007. Among rural females and males, the modal age-at-death increased from age 73 and age 70, respectively, in 1970–1974, to age 77.5 and age 75, respectively, in

Table 4. Trends in 25th and 75th percentiles of DOD, by residence and sex, India

	25th percentile of DOD				75th percentile of DOD			
Year	Rural female	Rural male	Urban female	Urban male	Rural female	Rural male	Urban female	Urban male
1970–1974	54	54	60	57	76	75	80	78
1974–1978	55	54	60	57	76	75	80	77
1979–1983	58	57	63	59	79	76	81	79
1984–1988	59	57	62	59	79	77	82	78
1989–1993	61	58	65	60	79	78	82	79
1994–1998	62	58	65	61	81	78	84	80
1999–2003	63	58	66	61	82	79	84	82
2003–2007	64	60	68	62	83	80	85	82
Adj. R^{2*}	0.9678	0.8694	0.9362	0.9414	0.9502	0.9451	0.9099	0.8485

Source: Author's calculation from RGI (25); DOD: distribution of age-at-death.
*Adj. R^2 is based on 34 observation of percentile age-at-death between 1970 and 2007.

2003–2007 (Fig. 7). Throughout the reference periods, females—irrespective of their residence—showed a consistently higher modal age-at-death than did their male counterparts. This is because of higher e_x among females than males, especially in old ages. The simultaneous process of accumulation of deaths near the modal age-at-death and the shift in modal age-at-death toward higher ages casts a thin shape to the distribution of age-at-death. Therefore, this suggests a decline in variance in age-at-death or the process of compression of mortality over the time period (36, 46, 49).

Comparisons of modal age-at-death between India and some developed nations (such as Sweden, Switzerland, the UK, France, Italy, and Japan) reveal that India's modal age-at-death in recent decades corresponds to modal age-at-death of developed countries in the 1930s and 1940s (Fig. 8a and 8b) (37). Therefore, India significantly lags behind developed nations in the progress of mortality transition—that is, in demographic and epidemiological transition (28). Nevertheless, in recent decades, a sustained and rapid decrease in mortality rates and the linear increase in e_x, the 25th and 75th percentiles, and in the modal age-at-death provide an irreversible and definite shape to the distribution of age-at-death manifested as advances in mortality transition.

The progression in mortality transition is evident and phenomenal. Most older age deaths have been characterized by and attributed to the increasing prevalence of chronic diseases and fatality from them. In addition, the increasing burden of communicable diseases in old ages augments the number of total deaths. Altogether, the advances in mortality transition are concomitant with the structural changes in disease patterns vis-à-vis mounting age patterns of morbidity and mortality. In sum, these phenomena provide evidence of rapid progress in demographic and epidemiological transition.

Discussion and conclusion

In this study, we examined the structural changes in patterns of diseases vis-à-vis advances in mortality transition, which unraveled the considerable progress of epidemiological transition. From 1970 to 2007, India moved swiftly from the dominance of child and adult mortality to a progressive phase dominated by old age mortality. Although the initial periods of the 1970s and 1980s were characterized by a larger burden of communicable diseases, the burden of chronic NCDs emerged as a major cause of old age deaths in the later period even though communicable diseases responsible for deaths have been interchanged. Currently, the burden of communicable diseases remains substantially higher and is accountable for more than a 30% share of all deaths (15, 22, 32).

Although the burden of communicable diseases impacted a variety of ages during the 1970s and 1980s, by the mid-1990s, the burden of communicable diseases had increased considerably in adult and old ages. At the same time, the burden of chronic NCDs has been increasing in old ages; therefore, those who survived through their

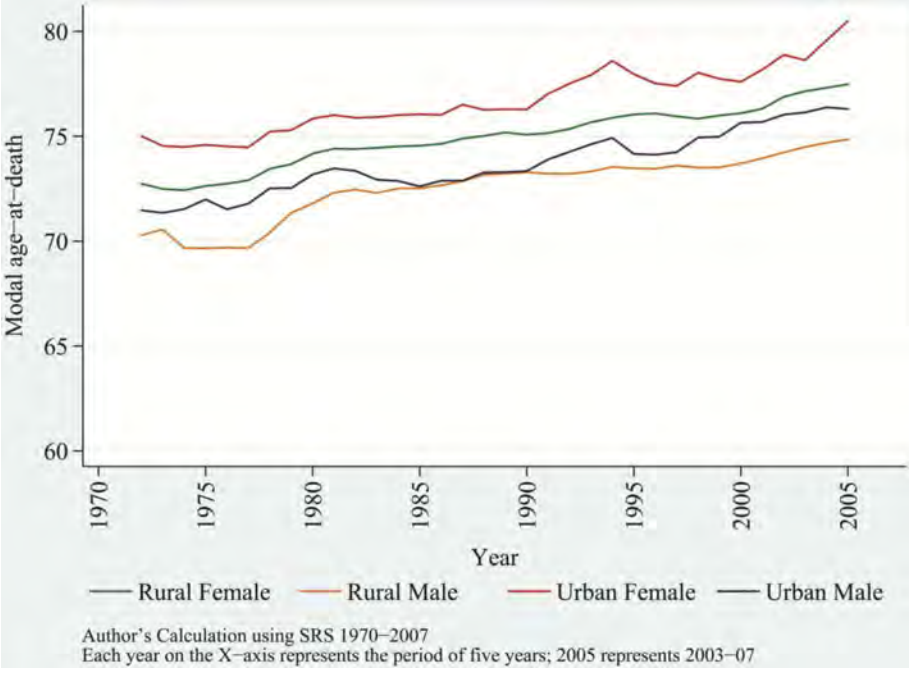

Fig. 7. Trends in modal age-at-death by sex and residence, India, from 1970–1974 to 2003–2007; rural female, rural male, urban female, and urban male.

Citation: Glob Health Action 2014, **7**: 23248 - http://dx.doi.org/10.3402/gha.v7.23248

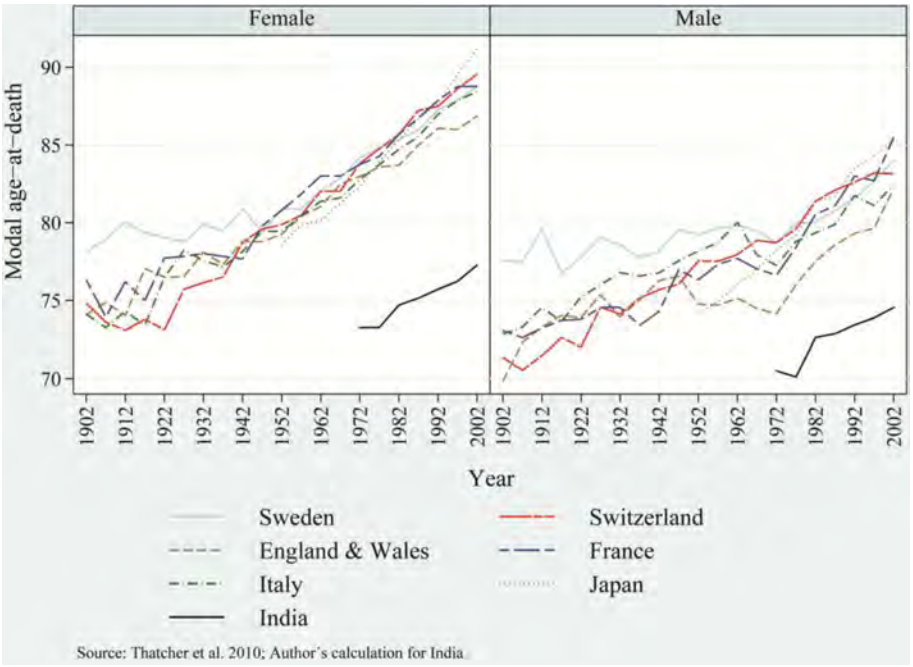

Fig. 8. Trends in modal age-at-death among females and males of selected developed countries and India, including Sweden, Switzerland, the UK, France, Italy, and Japan.

early years to middle and old age are susceptible to an elevated risk of dying. The mounting age pattern of chronic diseases in recent decades corroborates the higher death toll attributed to NCDs in old ages, and the emerging age pattern of mortality confirms the process of postponement of death. Among urban aged females, the prevalence rate of hypertension and problem of joints and bones rose from 20.4 and 22.3, respectively, in 1995–1996 to 80 and 32, respectively, in 2004. Among urban males, chronic diseases such as heart disease and diabetes rose from 15.4 and 15.5, respectively, in 1995–1996, to 45 and 57, respectively, in 2004 (Table 1). Hence, in recent decades, the enormous increase in the prevalence rate of chronic diseases and resulting fatalities has been responsible for structural changes in disease patterns.

The results showed remarkable progress in mortality transition consistent with structural changes in disease patterns. The emerging bell shape of the distribution of age-at-death manifests as an increasing bulk of deaths near the modal age-at-death. The phenomenon is being propelled by the linear increase in e_x, in the 25th and 75th percentiles of age-at-death, and in the modal age-at-death. The modal age-at-death shifted by almost 5 years compared to an increase of 2.7 years in e_{70+} during 1970–2007. The shift in modal age-at-death evidently demonstrates the dominance of old age mortality over the childhood/adult age mortality (34). Globally, the older age mortality is characterized by the heavy burden of NCDs. However, India has moved to a more challenging stage of demographic and epidemiological transition and has experienced a double burden of diseases. India's evolving stage of epidemiological transition generally is not seen among developed nations, where mortality and morbidity are compressed in later years of life. In India, death is concentrated in later years of life but is accompanied by a mounting burden of morbidity. Comparatively, India's mortality conditions lag behind those of developed nations by almost 60–70 years. Continued progress in mortality transition and structural changes in the disease patterns accompanied by a linear rise in e_x indicates that there is a compelling variation in advances found so far in epidemiological transition witnessed by the developed nations, with similar matrices for India.

Amidst demographic and epidemiological transition, India is experiencing a remarkable transformation in the age pattern of morbidity and mortality, the structural changes in disease patterns and consequent double burden of diseases and in mortality transition. Progression in such fundamental and integral demographic processes is evidence of advances in epidemiological transition in the last four decades. All the categories of population, that is, rural-urban and female-male are advancing in epidemiological transition. Urban females have experienced the greatest structural changes in disease patterns and mortality transition compared to other population categories and are leading in epidemiological transition.

Main findings

- India is currently experiencing the double burden of communicable and non-communicable diseases. In recent decades, the age pattern of morbidity has been rising, primarily due to increased prevalence of chronic diseases, resulting in significant structural changes in disease patterns.
- Modal age-at-death is rising in India, demonstrating the dominance of old age mortality over childhood/adult mortality; this is mostly attributable to the high prevalence of chronic diseases and ensuing fatality.
- Similar to developed nations, most Indians now live to old age; however, they are experiencing increasing number of years lived with disability as a consequence of increasing morbidity. Continued reductions in mortality and structural changes in disease patterns strongly indicate epidemiological transition in India, as these patterns begin to emulate those seen in developed nations.

Key messages for action

- India is experiencing rapid health transition, including increased life expectancy at old ages ($e_{60 \text{ and above}}$). However, the older population is living in poor health. Comprehensive health interventions are required for prevention and control of chronic diseases.
- Urban females lead in epidemiological transition with higher life expectancy and high morbidity rates, and thus require urgent interventions.
- In terms of mortality transition, India lags behind developed nations. The combination of a double burden of disease with high morbidity rates presents challenges for improving the overall health status of the population and necessitates a comprehensive policy and action to prevent and control this burden and promote healthy ageing.

Conflict of interest and funding

The authors have not received any funding or benefits from industry or elsewhere to conduct this study.

References

1. Gribble JN, Preston SH. The epidemiological transition: policy and planning implications for developing countries [abstract]. National Research Council. Washington, DC: The National Academies Press; 1993.
2. RAND. Policy and health in Asia: demographic and epidemiological transitions. Available from: http://www.rand.org/content/dam/rand/pubs/research_briefs/RB5036/RB5036.pdf [cited 24 January 2014].
3. Pool I, Wong LR, Vilquin E. Age structural transitions: challenges for development. Paris: Committee for International Cooperation in National Research in Demography; 2006.
4. Bhat PNM, Navneetham K. Recent trends in age-specific mortality in India. J Inst Econ Res 1991; 26: 49–69.
5. Karkal M. Differentials in mortality by sex. Econ Polit Wkly 1987; 22: 1343–7.
6. Krishnaji N, James KS. Gender differentials in adult mortality; with notes on rural-urban contrasts. Econ Polit Wkly 2002; 37: 4633–7.
7. Navaneetham K. Mortality decline in India: an analysis of regional and temporal variations. Demogr India 1993; 22: 53–63.
8. Ranjan A. Mortality change in India: 1970–85. Demogr India 1993; 22: 97–112.
9. Roy TK, Lahiri S. Recent levels and trends in mortality in India and its major states: an analysis of SRS data. In: Srinivasan K, Mukerji S, eds. Dynamics of population and family welfare. International Institute for Population Sciences, pp. 279–349. Bombay: Himalaya Publishing House; 1987.
10. Kingsley D. The amazing decline of mortality in under developed areas. Am Econ Rev 1956; 46: 305–18.
11. Padmanabha P. Mortality in India: a note on trends and implications. Econ Polit Wkly 1982; 17: 1285–90.
12. RGI (2012b). SRS bulletin, sample registration system. Registrar General, India. 47(2). Vital Statistics Division, R. K. Puram, New Delhi.
13. RGI (2009b). Compendium of India's fertility and mortality indicators 1971–2007. New Delhi: Ministry of Home Affairs.
14. Kumar BG. Low mortality and high morbidity in Kerala reconsidered. Popul Dev Rev 1993; 19: 103–21.
15. Visaria L. Mortality trends and the health transition. In: Dyson T, Cassen R, Visaria L, eds. Twenty-first century India—population, economy, human development, and the environment, pp. 32–56. New Delhi: Oxford University Press; 2004.
16. Planning Commission (2011). Report of the working group on disease Burden for 12th five year plan: WG-3(2): non communicable diseases. Available from: http://planningcommission.nic.in/aboutus/committee/wrkgrp12/health/WG_3_2non_communicable.pdf [cited 20 December 2011].
17. NSSO (2006). Morbidity, health care and the conditioned of the aged. NSS 60th round, January–June 2004, Report No. 507. New Delhi: Ministry of Statistics and Programme Implementation, GOI.
18. NSSO (1998). Morbidity and treatments of ailments. NSS 52nd round, July 1995–June 1996, Report No. 441. New Delhi: Ministry of Statistics and Programme Implementation, GOI.
19. Ministry of Health and Family Welfare (MOHFW) (2007). Select health parameters: a comparative analysis across the National Sample Survey Organization (NSSO) 42nd, 52nd, and 60th rounds. Ministry of Health and Family Welfare, Government of India, In Collaboration with the WHO Country Office for India. Available from: http://s3.amazonaws.com/zanran_storage/whoindia.org/ContentPages/112354873.pdf [cited 28 April 2012].
20. Omran AR. The epidemiologic transition. A theory of the epidemiology of population change. Milbank Q 1971; 49: 509–38.
21. Omran AR. The epidemiological transition theory: a preliminary update. J Trop Pediatr 1983; 29: 305–16.
22. RGI (2009a). Report on causes of death in India 2001–2003. New Delhi: Office of Registrar General, Ministry of Home Affairs.
23. RGI (2012a). SRS based abridged life tables 2003–07 to 2006–2010. SRS analytical studies, Report No. 1 of 2012. New Delhi: Registrar General, India.

24. John TJ, Dandona L, Sharma VP, Kakkar M. Continuing challenge of infectious diseases in India. Lancet 2011; 377: 252–69.
25. RGI (1970–2007). Sample registration system: statistical reports. New Delhi: Office of the Registrar General of India.
26. Arokiasamy P, Yadav S. Changing age patterns of morbidity vis-à-vis mortality in India. J Biosoc Sci 2013: 1–18. Available on CJO 2013 DOI: 10.1017/S002193201300062X.
27. James KS. India's demographic change: opportunities and challenges. Science 2011; 333: 576–80.
28. Ranjan Chaurasia A. Mortality transition in India 1970–2005. Asian Popul Stud 2010; 6: 47–68.
29. Quigley MA. Commentary: shifting Burden of disease – epidemiological transition in India. Int J Epidemiol 2006; 35: 1530–1.
30. Joshi R, Cardona M, Iyengar S, Sukumar A, Raju CR, Raju KR, et al. Chronic diseases now a leading cause of death in rural India—mortality data from the Andhra Pradesh rural health initiative. Int J Epidemiol 2006; 35: 1522–9.
31. Banthia J, Dyson T. Smallpox in nineteenth-century India. Popul Dev Rev 1999; 25: 649–80.
32. Johansson SR. The health transition: the cultural inflation of morbidity during decline of mortality. Health Trans Rev 1991; 1: 39–65.
33. Gruenberg EM. The failure or success. Milbank Q 1977; 55(1): 3–24.
34. Canudas-Romo V. The modal age at death and shifting mortality hypothesis. Demogr Res 2008; 19: 1179–204.
35. Kannisto V. Mode and dispersion of the length of life. Population 2001; 13: 159–71.
36. Kannisto V. Measuring the compression of mortality. Demogr Res 2000; 3: 24.
37. Thatcher AR, Cheung SLK, Horiuchi S, Robine J-M. The compression of deaths above the mode. Demogr Res 2010; 22: 505–38.
38. Edwards RD, Tuljapurkar S. Inequality in life spans and a new perspective on mortality convergence across industrialized countries. Popul Dev Rev 2005; 31: 645–74.
39. Young-Xu Y, Chan KA. Pooling overdispersed binomial data to estimate event rate. BMC Med Res Methodol 2008; 8: 58.
40. RGI (1972–1997). Survey of causes of death. New Delhi: Office of the Registrar General of India.
41. RGI (1975–1995). Mortality statistics of causes of death. New Delhi: Office of the Registrar, General of India.
42. RGI (2001–2004). Medical certification of causes of death. New Delhi: Office of the Registrar General of India and Census Commissioner, India, Ministry of Home Affairs.
43. United Nations (1982). Model life tables: for developing countries. Population Studies, No. 77. New York: Department of International Economic and Social Affairs.
44. United Nations (1988a). MORTPAK. The United Nations software package for mortality measurement (Batch oriented software for the mainframe computer). New York: United Nations, pp. 114–18.
45. United Nations (1988b). MORTPAK-LITE. The United Nations software package for mortality measurement (Interactive software for the IBM-PC and compatibles). New York: United Nations, pp. 111–14.
46. Wilmoth JR, Horiuchi S. Variability of age at death within human populations. Demography 1999; 36: 475–95.
47. Siegel JS, Swanson DA. The methods and materials of demography. 2nd ed, Boston: Academic Press; 2004.
48. Kostaki A, Panousis V. Expanding an abridged life table. Demogr Res 2001; 5: 1–22.
49. Cheung SLK, Robine J-M, Paccaud F, Marazzi A. Dissecting the compression of mortality in Switzerland, 1876–2005. Demogr Res 2009; 41: 569–98.
50. Murray CJL, Lopez AD. Global and regional cause-of-death patterns in 1990. Bull World Health Organ 1994; 72: 447–80.
51. Murray CJL, Chen LC. Understanding morbidity change. Popul Dev Rev 1992; 18: 481–503.
52. Kurpad AV, Mony P, Vaz M. Chronic disease in India-where next? Proceeding International Symposium Building Leadership Skills in Food and Nutrition Essential for National Development CFTRI, Mysore, India, 23–25 June 2006.
53. Gupte MD, Ramachandran V, Mutatkar RK. Epidemiological profile of India: historical and contemporary perspectives. J Biosci 2001; 26: 437–64.
54. United Nations (2009). World population prospects: the 2008 revision. New York: Department of Economic and Social Affairs.
55. United Nations (2012). Adult mortality rate (probability of dying between 15 and 60 years per 1000 population). UNdata, Statistics. Available from: http://data.un.org/Data.aspx?d=WHOandf=MEASURE_CODE%3AWHOSIS_000004 [cited 20 May 2012].

Global Health Action

PART III

Changes in mortality and human longevity in Kerala: are they leading to the advanced stage?

Muttikkal B. Thomas* and Kuriath S. James

Population Research Centre (PRC), Institute for Social and Economic Change (ISEC), Bangalore, India

Background: During the last century, Kerala witnessed drastic mortality reduction and high improvement in longevity. This achievement is often compared with that of developed countries. However, how far the early advantages in mortality reduction have further enhanced in Kerala remains unknown. In most developed countries, advanced stage of mortality reduction and further increase in longevity was achieved mainly due to the mortality shift from adult and older ages to oldest ages (Olshansky and Ault 1986).
Objectives: Considering the lack of comprehensive study on the change in longevity in Kerala, this study focuses on discovering (i) the historical time-periods that provided the biggest gain to life expectancy and also the beneficiaries (by age group and sex) and (ii) the contributions of major groups of causes of death in mortality reduction and consequent improvement in longevity.
Methodology and data: The study uses the methodology proposed by Olshansky and Ault in 1986. It used methods such as Temporary Life Expectancy (TLE), Annual Relative Change in TLE, Decomposition of changes in longevity among different age groups (gender and spatial) and causes of deaths, for the analysis. It used data from various sources such as Census, Civil Registration System (CRS) and Directorate of Health Services (DHS), as well as survey data from Sample Registration System (SRS) and Medically Certified Causes of Deaths (MCCD) for this study.
Finding and conclusion: The study found that overall mortality dramatically declined in the state in the recent decades. Younger ages have contributed the most for this reduction. Therefore, further mortality reduction is possible in adult and early old ages. However, the contribution of these ages to life expectancy was lower than that of youngsters until 1991–2000 especially among males. This may indicate a slow progress towards the advanced stage of epidemiological transition characterized by high prevalence of non-communicable diseases. The paper concludes that although the health issues of infants, children, and mothers in the reproductive age group, are effectively addressed through various policies in Kerala, the state needs to focus more on the health problems of adults, especially males.

Keywords: *epidemiological transition*; *mortality and longevity*; *Kerala*; *advance stage of mortality*

Responsible Editors: Nawi Ng, Umeå University, Sweden; Barthélémy Kuate Defo, University of Montreal, Canada.

*Correspondence to: Muttikkal B. Thomas, Population Research Centre (PRC), Institute for Social and Economic Change (ISEC), Bangalore 560072, India, Email: benson@isec.ac.in

This paper is part of the Special Issue: *Epidemiological Transitions – Beyond Omran's Theory*. More papers from this issue can be found at http://www.globalhealthaction.net

Received: 29 September 2013; Revised: 12 January 2014; Accepted: 21 January 2014; Published: 15 May 2014

In the past century, Kerala witnessed a remarkable decline in mortality and considerable advancement in life expectancy. This change was often compared to the pattern observed in developed countries (1, 2). Interestingly, Kerala has achieved low mortality despite low per capita income and higher incidence of malnourishment (3, 4). This remains a paradox to the development theorists. The changes in Kerala were highlighted as unique and referred to as the 'Kerala Model' of development (3, 5, 6).

There are conflicting arguments on the rapid decline in mortality and the reasons thereof. On one hand, it is postulated that state intervention since 1956 has played a major role in achieving reduction in mortality. According to a group of scholars, including Panikar and Soman, the increase in life expectancy was the result of superior medical care through primary health institutions and other measures, such as provision of clean water, sanitary facilities, and an efficient public distribution system, introduced since the formation of the state in 1956 (4, 7). The primary health institution includes primary healthcare services such as vaccination, direct medical care for infectious diseases, and perinatal, maternal, and child care, besides raising general awareness about health. In a comprehensive study, Caldwell (8) described that the role

Citation: Glob Health Action 2014, **7**: 22938 - http://dx.doi.org/10.3402/gha.v7.22938

of the state was inevitable for the high rate of mortality decline in Kerala. He pointed out the role of state-supported healthcare and education system, ensuring accessibility to public health to all, universal immunization, and the provision of antenatal and postnatal services, as a noteworthy aspect in achieving a high rate of mortality decline in Kerala. Another study by Krishnan (9) also enunciated the role of the state by citing the evidence of improved health outcomes in the Malabar region after the expansion of state health facilities in that region. Contrary to this, some scholars argued that the beginning of the massive reduction in mortality in Kerala could be traced to the era before the formation of the state. Kerala was formed on November 1, 1956, by merging Travancore and Cochin, two princely states, with Malabar (a part of erstwhile Madras presidency) on a vernacular basis. They argued that social and cultural improvement, especially through education, climatic conditions, the scattered pattern of settlements, and the mysterious disappearance of a major cause of death, namely plague, were the prime factors for reduction in mortality (5, 10, 11). However, there was no detailed historical assessment of mortality reduction and increased life expectancy in Kerala that could lend clarity to this debate. Moreover, the recent changes in the pattern of mortality were not analyzed carefully.

A comprehensive assessment of mortality trends necessitates an investigation into the expected pattern of change based on the experience in developed countries. Mortality reduction is closely linked to the shifts in disease pattern. In the first stage of mortality reduction, the shift occurs in the cause of death pattern from infectious to chronic-degenerative diseases (12). This leads to a distribution of death from younger to older age groups (>50 years). However, the transition later moves from the older age groups to the oldest age group, which is also known as the age of delayed degenerative diseases (13). At this stage, there is further postponement of death in older to the oldest age groups (ages more than 80) as a result of bringing down deaths from degenerative diseases from adult ages to older ages (13). Similar stage of epidemiological transition is also later suggested by Omran where it was described as Age of declining cardiovascular mortality with ageing, life style modification as well as more death from emerging and resurgent diseases. He also predicted a fifth stage, namely Age of aspired quality of life, with paradoxical longevity and (futuristic stage) persistent inequities.

Although it is well known that Kerala has moved from younger age group mortality to older age groups, the extent of transition to the oldest age group (delayed degenerative disease pattern), and its major causes of death remain unknown. Even though various studies have shown clear dominance of lifestyle/chronic diseases as causes of death and morbidity, they were restricted only to particular social and economic groups or to a specific period of time (14–16). Nevertheless, a preliminary study was conducted by Thomas in 2012 (17) aimed at looking into the changes in mortality by using point estimates such as mortality rates and causes of death, among others, and it was found that there was a transition in mortality to the adult age groups. Similarly, causes of death shifted to chronic degenerative lifestyle diseases from infections and primary healthcare–oriented diseases. Hence, it was necessary to conduct an in-depth investigation into this problem to identify the exact contribution of different age groups to life expectancy. Secondarily, there is also necessity to identify the contribution of major causes of death to the improvement in life expectancy in Kerala.

Considering the aforementioned aspects, this paper examines the pattern of mortality changes and human longevity in Kerala since the beginning of the past century. It aims to capture the dynamics of mortality reduction over the decades to explore the levels and trends in the transition process. Further, it investigates how far Kerala has moved from older age group mortality to the oldest age group mortality as experienced in the developed countries. Specifically, the analysis attempts to establish the time-periods that provided the biggest gain to life expectancy and also the beneficiaries (by age group and sex) and the extent of mortality reduction. It was also of interest to know the contributions of major groups of causes of death in mortality reduction.

Data and methodology

A major difficulty in measuring mortality transition in Kerala emerges from the unavailability of a single reliable data set in the past century. The only reliable data prior to the 1960s are the decennial census in India (13, 14, 18–21). However, the inception of the Sample Registration System (SRS) in Kerala provided an alternative and more reliable data on mortality after the mid-1960s. Therefore, the study mainly uses these two data sets for the purpose of estimation, that is, the first main set of data is from Monograph No: 7, Census of India 1961 (1911–20 to 1951–60) (18, 19) and rest of the data (1971–80 to 2001–08) are from the SRS. However, there was no readily available data for the period from 1961 to 70. Therefore, the study uses figures from the Western Model of Life Tables by Coale, Demeny, and Vaughan (22) for the decade 1961–70 as proxy considering the life expectancy of Kerala as estimated by Bhat (23). The study also uses data from the survey of Medically Certified Causes of Deaths (MCCD) for estimating the contribution of major group of causes of deaths in the advancement of longevity in the state. The MCCD data were obtained from the Registrar General of India's Report on Medical Certification of Cause of Death for various years, published by the Office of the Registrar General and Census in New Delhi.

It is understandable that the data used for the study are limited by their quality. The census data used to analyze the patterns of mortality are constrained by the inaccuracy

in the mortality estimations due to two factors. First, to make reliable estimates of mortality from census data, a supplementary data on infant and child mortality are essential. These are not always readily available. Second, the census data are often subject to bias due to rampant misreporting of age by the respondents (13). Similarly, the proxy mortality rates taken from the 'Western Model of Life Table' may have slight variations between different age intervals. At the same time, the quality of mortality data from SRS may also be restricted by their own sample size. Likewise, the data from MCCD are also handicapped by their lack of quality and non-availability. Considering this fact, this study uses information on causes of death from MCCD that is available from 1976 for urban areas in the state.

The paper follows the methodology forwarded by Olshansky and Ault (16) considering the possibility of mortality reduction from adult and early old age groups to the oldest age group. It compares the change in absolute value of mortality level by the increase or decrease in life expectancy between different periods. However, a major lacunae of life expectancy indicator is that it measures the mortality level for an open age interval x and above. Therefore, it is often limited to the problems of data reliability in older age groups and the restriction on limits of human life span (24). To avoid these problems, the paper analyses the relative risk of mortality transition for closed age interval (x, x + n) by using Temporary Life Expectancy (TLE) and index of Annual Relative Change (ARC) of TLE in the second section. Finally, the paper analyses changes in relative importance of death in older age groups in Kerala by estimating the rate of survival to older age groups, median age of death, and the exact contribution of each group (by age and sex) toward life expectancy at different time intervals. The study also decomposed the improvement in life expectancy by major group of causes of death to identify the dominance of major diseases.

For the purpose of analysis, the study classifies the male and female population into three major groups as youngsters (aged 0–15 years), adults (aged 15–60 years), and old (> 60 years). However, the adults and the aged population are again sub-divided into young adults (aged 15–40 years), old adults (aged 40–60 years), older age group (aged 60–80, and the oldest age group (aged 80 +) for cross examination while considering the vulnerability of diseases (25, 26). Similarly, the historical periods are also divided into different decadal intervals.

An overview of changes in death rates and life expectancy in Kerala

Magnitude of absolute change in life expectancy

Through the reduction in mortality rates, an impressive level of life expectancy has been achieved in Kerala since 1911–20. Such a change can be seen over the time periods as well as among the age and sex groups. Table 1 records the levels and changes in the life expectancy in the state over the past century. It shows that the life expectancy at birth of males rose from 25.5 years in 1911–20 to 70.7 years in 2001–05. Similarly, female life expectancy increased from 27.4 years to 77.1 years. This improvement gave an advantage of 45.2 years for males and 49.7 years for females within 88 years of time – an average annual increase of about 0.51 and 0.56 years, respectively. However, the average annual increase within each decade reflects disparities among the decades, that is, the pace of increase in life expectancy was not constant but varied from decade to decade as shown in the Table 1. Except in 1951–80 for both sex groups and 1921–40 for males and 1971–90 for females, the annual contribution of absolute years to life expectancy was below the overall average showing a low pace in those decades. However, the two decades of 1951–60 and 1961–70 recorded a high pace in absolute changes of life expectancy. Perhaps, the actual value of the increment will be more in terms of relative changes considering restriction due to the limits of human life span (24).

Similarly, the magnitude of changes in life expectancy is different among the age groups, as shown in Table 2. The early decades registered a high, absolute change in life expectancy in the younger age groups. Nevertheless, a slowdown in gains in life expectancy in the younger age groups and comparatively high gains in life expectancy in the advanced age groups are visible in the recent. For instance, the absolute change in life expectancy for females at birth during 1911–20 to 1941–50 was 15.6 years (56.9%) while it was 4.1 years (48.8%) at age at 60. But in recent decades (1971–80 to 2001–08), the absolute changes in life expectancy at birth was by

Table 1. Levels and changes of life expectancies at birth in Kerala during 1911–20 to 2001–08

Period		Absolute increase		Annual average years added	
From	To	Male	Female	Male	Female
1911–20	2001–08	45.25	49.71	0.51	0.56
1911–20	1921–30	4.05	5.29	0.41	0.53
1921–30	1931–40	5.49	5.22	0.55	0.52
1931–40	1941–50	4.58	5.06	0.46	0.51
1941–50	1951–60	4.63	5.17	0.46	0.52
1951–60	1961–70	9.90	9.41	0.99	0.94
1961–70	1971–80	7.83	7.10	0.78	0.71
1971–80	1981–90	4.53	7.64	0.45	0.76
1981–90	1991–00	2.68	2.76	0.27	0.28
1991–00	2001–08	1.56	2.07	0.19	0.26

Source: Calculated from Namboodiri, 1968; Coale, Demeny and Vaughan, 1983; Bhat, 1987; SRS various years.

Citation: Glob Health Action 2014, **7**: 22938 - http://dx.doi.org/10.3402/gha.v7.22938

Table 2. Life expectancy at different age levels in Kerala during 1931–40 to 2001–08, by sex

	1911–20	1921–30	1931–40	1941–50	1951–60	1961–70	1971–80	1981–90	1991–00	2001–08
Male										
0	25.5	29.5	35.0	39.6	44.2	54.1	62.0	66.5	69.2	70.7
15	29.6	33.8	35.5	38.5	41.5	49.0	53.2	55.0	55.9	57.0
40	15.1	17.0	18.4	20.6	22.7	28.6	30.4	32.0	32.7	33.7
60	6.9	8.6	9.7	10.6	11.5	14.4	15.0	16.6	16.7	17.4
80	–	–	3.8	4.0	4.3	5.1	5.3	6.6	6.9	6.9
Female										
0	27.4	32.7	37.9	43.0	48.1	57.6	64.7	72.3	75.1	77.1
15	28.9	33.7	36.8	40.3	43.6	51.6	56.3	60.6	61.6	63.4
40	16.6	17.8	22.4	24.3	26.3	31.2	33.3	37.0	37.8	39.4
60	8.4	9.0	11.5	12.5	13.4	16.0	16.3	19.2	19.9	21.2
80	–	–	4.2	4.5	4.8	5.6	5.3	6.9	7.3	7.9

Source: Calculated from Namboodiri (18); Coale, Demeny, and Vaughan (22); Bhat (23); SRS various years.

12.4 years (19.2%), 4.9 years (30.6%) at age 60, and 2.6 years (49.1%) at age 80 in Kerala. This indicates a major shift in the relative gain at different ages toward increase in life expectancy. There is an obvious reversal of gains in life expectancy at the younger ages has given way to a drastic increase in the life expectancy at the advanced ages. Nevertheless, this change in advanced age groups was relatively lower among males than females. Between 1971–80 and 2001–08, the life expectancy of males increased only by 8.7 years at birth (14.0%), by 2.4 years (16.0%) at age 60, and by 1.6 years (30.2%) at age 80 indicating delay in mortality transition among them.

The sex difference in life expectancy, especially at birth, is also a notable feature of life expectancy changes in Kerala as indicated in Table 2. The table shows a narrow level of discrepancy in life expectancy at birth during the early decades, which widened in the more recent decades. For instance, in 1911–20, the sex difference at birth was only 1.9 years and rose to 6.4 years in the 2001–08 period. It is important to note that this sex difference in life expectancy has remained at a higher level since 1981–90, with females surging ahead in gaining greater mortality decline and improvement in life expectancy than males. This change is very different from the US experience where the sex difference in life expectancy was projected to end in the fourth stage of epidemiological transition (13). In short, the analysis of absolute change in life expectancy shows that life expectancy in Kerala has improved over the decades in absolute terms. However, there are differences in the magnitude of such changes across the time periods as well as among age and sex groups.

The absolute life expectancy estimates, which can be used as a proxy to the rate of mortality, suffer from problems of relative magnitudes. In other words, the possible change in life expectancy indicators depends on the level of already achieved life expectancy (24). Moreover, the estimation of life expectancy is constrained by technical problems of unreliability of data, especially pertaining to the older age groups (generally more than 65 years) in the developing countries (25). Besides, there is a possibility of a counteracting effect, which neutralizes the contribution of different age groups – a high mortality increase experienced by one group is offset by a reduction in the other group (27). This point to a certain level of arbitrariness in results (with respect to mortality changes) which was estimated from the absolute life-expectancy indicators. These lacunae in estimation call for an alternative measure to capture the changes in mortality considering its relative risks at different age structures, life expectancy levels, and the quality of data at different age groups. Therefore, we use TLEs and the index of ARC of TLE for further analysis.

Transition in relative risk of mortality by age and sex

Temporary life expectancies

Table 3 shows TLE at select exact age intervals by sex for Kerala between 1931–40 and 2001–08. The TLE ($_i^e x$) from age x to (x+i) is the average number of years that a 'group of persons' alive at exact age x will live from age x to (x+ i) years (24)

$$\text{i.e. } {}_ie_x = \left(\frac{T_x - T_{x+i}}{l_x}\right)$$

where, $_i^e x$ represents the TLE from age x and $x+$ n. T_x and T_{x+i} are the number of person-years lived at age l_x and older, and $x+i$ and older, respectively. l_x is the number of survivors at age x in a life table (as radix) of 100,000.

Table 3 records a substantial increase in TLE over the decades in Kerala. However, the nature and pace of this change are different across age intervals as well

Table 3. Observed temporary life expectancies at selected exact age intervals by sex (1931–40 to 2001–08)

	1931–40	1941–50	1951–60	1961–70	*1971–80	*1981–90	*1991–00	*2001–08
Male								
0–80	34.96	39.49	44.01	53.40	60.95	64.76	67.24	68.56
0–15	10.85	11.48	12.06	12.94	13.79	14.34	14.69	14.77
15–40	22.12	22.61	23.04	23.71	24.46	24.53	24.60	24.62
40–60	14.32	15.29	16.11	17.88	18.45	18.54	18.77	18.89
60–80	9.33	10.11	10.88	13.03	13.48	14.23	14.20	14.70
Female								
0–80	37.75	42.69	47.68	56.41	63.30	69.49	71.82	73.13
0–15	11.43	12.04	12.62	13.20	13.77	14.42	14.73	14.78
15–40	21.58	22.21	22.76	23.81	24.52	24.68	24.74	24.78
40–60	15.88	16.47	17.11	18.34	19.06	19.38	19.44	19.57
60–80	10.87	11.66	12.39	14.13	14.49	15.94	16.22	16.74

Source: Calculated from Namboodiri (18); Coale, Demeny, and Vaughan (22); Bhat (23); SRS various years.
*The findings from 1971 onward have been taken from ref. 17.

as decade intervals. It can be observed that the TLE between birth and age 80 increased to 68.56 years and 73.13 years in the 2001–08 period for both males and females, respectively. It means that male babies born in Kerala during the 2001–08 period can be expected to live 68.56 years, whereas female babies can be expected to live 73.13 years, provided the mortality rates in the period when they were born remains constant throughout their lives. Nevertheless, though the expected lifespan in Kerala has almost doubled during this period, it still leaves room of 11.44 years and 6.87 years, for both male and female groups, respectively, for further improvement.

The change in TLE was also different among various age groups. It shows that younger age groups benefited more from the TLE changes than the older age groups over the decades. It should be noted that the subgroups (0–15, 15–40, 40–60) had already attained almost maximum TLE (difference was below 0.5 years) before 1971–80, except for males in the age group of 40–60. Also, though the older age group (60–80) achieved improvement in TLE, there is still ample room for improvement. It also indicates that the possibility of further mortality reduction is concentrated in the older age groups (40–60, 60–80). Moreover, it should be noted that the changes in TLE showed a severe stagnation for males, especially in these age groups in the past four decades in Kerala.

Though the data show an improvement in TLE in almost all age groups over the decades, the changes are similar for both sexes across the different age groups. Difference between the TLE at birth and age 80 can be taken as an instance where it shows a difference of 4.57 years in 2001–08. It also shows that there is a miniscule sex difference in the younger age groups (0–15, 15–40) and high disparities in the older age groups (40–60 and 60–80). It denotes that the overall difference in TLE between the ages 0 and 80 years would be due to the disparities between males and females in the older age groups. It is also worth mentioning that the reduction in difference in TLE in the younger age groups could be a result of effective control over the causes of death in the younger age groups. Perhaps, this has not happened in the case of the older age groups. A comparatively lower TLE for males than females (females almost near the maximum) in the age group 40–60 years and 60–80 years means that the mortality shift was unequal and that more males than females lagged in those age groups.

Index of ARC of in TLE

The index of ARC in TLE during 1931–40 to 2001–08 in Kerala is recorded in Table 4. It represents the percentage change in two mortality measures in their observed reduction in deaths in relation to the total possible reduction (24). In other words, it shows the change in number of years lived between two periods considering the maximum possibility of reduction in that age group.

The ARC can be calculated as: ${}_i\mathrm{ARC}_x^n = [1 - (1 - {}_i\mathrm{RC}_x^n)^{\frac{1}{n}}] \cdot 100$, where ${}_i\mathrm{ARC}_x^n$ represents Index of ARC in TLE; n is the number of years, and ${}_i\mathrm{RC}_x^n$ is the observed change in TLE in relation to the maximum possible change,

where ${}_i\mathrm{RC}_x^n$ can be calculated as, ${}_i\mathrm{RC}_x^n = \frac{{}_i e_x^{t+n} - {}_i e_x^t}{i - {}_i e_x^t}$, where ${}_i e_x^{t+n} - {}_i e_x^t$ is the absolute change of TLE of years of life between two particular ages; i is the maximum possible TLE between the age intervals.

The figures in Table 4 indicate high variations in the pace of improvements in TLE across sex and time-period. By and large, it can be said that significant changes happened between 1951–60 and 1981–90. It may be recalled that our analyses of Crude Death Rates (CDR) and absolute changes in life expectancy have pointed to a similar effect for the period after state formation. Almost all the ages and sex subgroups were at their best performance in terms

Citation: Glob Health Action 2014, **7**: 22938 - http://dx.doi.org/10.3402/gha.v7.22938

Table 4. Annual Relative Change Index of TLE at selected age intervals by sex in between different decades (1931–40 to 2001–08)

	1931–40| 1941–50	1941–50| 1951–60	1951–60| 1961–70	1961–70| 1971–80	*1971–80| 1981–90	*1981–90| 1991–00	*1991–00| 2001–08
Male							
0–80	1.05	1.18	2.98	3.28	2.21	1.76	1.36
0–15	1.62	1.78	3.52	5.14	5.89	7.30	3.68
15–40	1.86	1.96	4.10	8.38	1.33	1.60	0.73
40–60	1.86	1.90	5.88	3.09	0.57	1.71	1.31
60–80	0.76	0.81	2.65	0.66	1.22	−0.06	1.12
Female							
0–80	1.24	1.43	3.10	3.40	4.52	2.47	2.16
0–15	1.86	2.17	2.73	3.71	7.25	7.52	2.42
15–40	1.99	2.18	6.14	8.77	3.94	2.01	2.24
40–60	1.52	1.99	5.40	5.53	4.12	0.88	3.46
60–80	0.89	0.92	2.56	0.64	3.01	0.70	1.83

Source: Calculated from Namboodiri (18); Coale, Demeny, and Vaughan (22); Bhat (23); SRS various years.
*The findings from 1971 onward have been taken from ref. 17.

of TLE in that period. It could be due to the provision of better healthcare measures that responded to the causes of death at different time-periods in Kerala.

However, the performance of ARCs in TLE indices is negligible in the recent decades. This rapid decline in pace, especially in younger age groups, may be due to the fact that the TLE in younger age groups have rapidly approached the size of age interval and the limits of further decline. However, it is also seen in Table 7 that the ARC in TLE in the advanced age groups for both sexes (40–60 and 60–80) was comparatively lower than that of the early age groups after 1971–80 in most cases. This is interesting because we expected more changes in the older age groups considering the possibility of the advanced mortality reduction indicating a shift of mortality from the older to the oldest age groups.

Sex difference in pace of improvement in TLE is another issue of concern. It was recorded that in all the periods except 1951–60 to 1971–80 at the 0–15 and 60–80 age groups, 1981–90 to 1991–2000 at the 40–60 age group, and 1991–2000 to 2001–08 at the 0–15 age group, the pace of change was higher for females than males. Moreover, though the difference in the pace was low in 1931–40 to 1941–50, it stabilized thereafter at a higher level till 1981–90. Later, it dropped further for almost all age groups, then picked-up again in the 2001–08 period. This phenomenon was more intensive in the younger age groups than the in the older age groups. Nevertheless, the low performance of pace indices of TLE among older age groups, especially with regard to males than females, may be an indication of a lag in moving toward the stage of advanced mortality reduction in Kerala.

To sum up, the analysis of changes in the relative risk in mortality reflects an impressive change in TLE in Kerala during the past century. It was noted that the younger age groups contributed almost their maximum capacity to TLE by the reduction in mortality. However, the contribution was low from the older age groups which differed from our expectation. Moreover, there were high disparities prevalent between males and females in TLE, especially in the older age groups. This could be due to a low rate of increase in male TLE during those periods. Similar finding is also evident in the annual index of TLE. Notably, there was low rate of change in the indices of TLE after 1981–90 in most cases in advanced age groups.

Changes in relative importance of mortality at older ages

The epidemiological transition theory put forward by Olshansky and Ault (13) emphasized the importance of mortality changes in the older age groups. In order to analyze the importance of mortality in the older age groups, this section attempts to discuss the changes in the median age of death and the proportion of survival to older age groups. Further, we analyze the contribution of each age group toward increments in life expectancy to identify the recent dynamics of mortality change and thereby understand the possibility of advanced stages of epidemiological transition.

Median age at death and proportion survival to older age groups

Over the decades, it has been observed that the proportion of survival of population from birth to the oldest age group has increased in Kerala. Table 5 records sex-wise distribution of the proportion of survival to ages 60 and 80, and median age at death in the state during the past century. It can be observed in Table 5 that the median age

Table 5. Median age at death and proportion of survival to age at 60 and 80, by sex in Kerala during 1911–20 to 2001–08

	1911–20	1921–30	1931–40	1941–50	1951–60	1961–70	1971–80	1981–90	1991–00	2001–08
Median age at death										
Male	20.82	30.67	39.55	45.46	50.92	63.23	68.96	71.29	72.30	73.85
Female	26.03	35.14	39.19	47.47	55.31	67.32	72.40	77.21	78.38	80.03
Proportion surviving to age 60										
Male	9.67	15.08	20.86	28.27	36.13	55.46	68.46	73.27	77.66	80.39
Female	11.75	17.77	27.60	35.73	44.16	61.59	75.53	84.85	88.27	90.42
Proportion surviving to age 80										
Male	0.22	0.97	1.80	3.21	5.25	14.44	19.26	26.27	28.02	31.56
Female	0.73	1.36	3.95	6.41	9.65	20.44	25.49	40.75	44.61	50.26

Source: Calculated from Namboodiri (18); Coale, Demeny, and Vaughan (22); Bhat (23); SRS various years.

of death is being pushed toward the older age groups in Kerala. Notably, the median age of death was 20.82 for males and 26.03 for females in the early decades of the century but rose to 73.85 and 80.03 for males and females, respectively, in the recent decades.

Similarly, the proportion of survival to the age of 60 and 80 from birth has also increased considerably over the decades in the past century. Table 5 shows that the proportion of survivors into the age of 60 was 9.67 for males and 11.75 for females in 1911–20 and increased to 80.39 and 90.42, respectively, in 2001–08 – approximately an eight-fold increase for both sexes. A similar increase is also seen in the proportion of people who survived to the age of 80 (the oldest age group). It may be noted that the survival to the oldest age group from birth cohorts in 1911–20 was less than 1%, but it increased to more than 30 percent in 2001–08 for both the sexes. The proportion of population who live for more than 60 years has dramatically increased over the century, thereby pushing mortality to the older age groups in the recent decades.

However, the pace of change in median age of death and the survival ratios vary across decades as well as by sex group. It can be observed that the rate of increase in both the median age of death and the survival ratio were relatively lower in the recent decades than in the early decades for both sexes. It could be due to low contribution from older age groups, whereas the expectations from the younger and adult ages are low because they are already close to maximum reduction. However, males have recorded relatively lower median age of death, lower survival to old ages, and comparatively lower rate of growth compared to females in Kerala. In a nutshell, the changes in proportion of survival to the older ages and the median age of deaths in the state indicates an increasing relevance of elders and their mortality to the ongoing improvement in life expectancy in Kerala. However, these indicators are incapable of revealing the exact contribution of the different age groups to the decline in mortality, which in turn bring about changes to life expectancy, limiting the possibility of understanding the advanced stage of mortality decline.

Contribution from each age group to gains in longevity

Decomposition of contribution of mortality decline from different age groups to the gain in life expectancy is one of the best measurements for identifying the advanced stage of mortality decline and human longevity in Kerala. According to the theory of epidemiological transition, at the fourth stage, we expect a higher contribution of life expectancy from the older age groups when the life expectancy increases to its maximum. Therefore, it is important to assess the relative contribution of age groups at different periods in order to ascertain the pattern of mortality change. The percentage contribution of each age group in gain in life expectancy at different periods is shown in Table 6.

Kerala experienced high improvement in life expectancy in the decades of 1951–60 to 1971–80 as shown in the table. The positive values indicate the percentage gains in life expectancy due to decline in mortality. Similarly, the negative values indicate reduction in life expectancy due to increase in mortality. However, we have ignored the negative figures from the mainstream interpretation because they are very few, of small magnitude, and refer to the oldest age group where misreporting of age at death is relatively high. The improvement peaked during 1951–70 gaining more than 9.5 years of life expectancy. During this period, all age groups contributed significantly. However, the contribution of the younger age groups (0–15) was the highest. It is visible that in all periods, except for the recent one, the younger age group (0–15) is the most important contributor to life expectancy improvement. This contribution of this group ranges from 28.3 to 69.6% for males and 16.2 to 64.1% for females. Nevertheless, their contribution was comparatively higher in the 1981–90 to 1991–2000 period but declined in the 1991–2000 to 2001–08 period.

Citation: Glob Health Action 2014, **7**: 22938 - http://dx.doi.org/10.3402/gha.v7.22938

Table 6. Contribution of mortality change at selected ages to total change in life expectancy by sex in Kerala in between different decades, 1931–40 to 2001–08 (%)

	1931–40| 1941–50	1941–50| 1951–60	1951–60| 1961–70	1961–70| 1971–80	1971–80| 1981–90	1981–90| 1991–00	1991–00| 2001–08
Male							
LE +	4.58	4.63	9.90	7.83	4.53	2.68	1.56
0–15	54.7	53.0	41.5	54.6	63.8	69.6	28.3
15–40	21.5	21.1	20.9	28.0	6.4	9.1	7.0
40–60	19.6	20.2	27.2	13.1	5.1	18.9	29.6
60–80	4.1	5.6	10.0	3.9	19.1	−0.5	35.9
80 +	0.1	0.2	0.4	0.3	5.6	3.0	−0.9
All	100	100	100	100	100	100	100
Female							
LE +	5.06	5.17	9.41	7.10	7.64	2.76	2.07
0–15	50.6	50.1	30.6	43.3	49.5	64.1	16.2
15–40	31.0	27.2	35.9	33.7	9.0	8.4	9.0
40–60	13.0	15.9	21.6	20.5	12.2	7.9	20.6
60–80	5.2	6.4	11.1	3.4	24.0	14.1	39.3
80 +	0.2	0.4	0.8	−0.8	5.3	5.4	14.8
All	100	100	100	100	100	100	100

Source: Calculated from Namboodiri (18); Coale, Demeny, and Vaughan (22); Bhat (23); SRS various years using the formulae given by Preston et al. (28).

However, the contribution of the age group 15–40 varies from period to period. It is seen that the contribution was high (between 20.9 and 35.9%) until 1971–80 but significantly came down (between 6.0 and 9.1%) after this period for both the sexes. The drastic change could be due to the fact that this age group had attained almost the maximum possible reduction during those decades. However, the age group of 40–60 years shows an increasing pattern in the contribution to life expectancy over the decades for both the sexes. Their relative contribution was below 20 percent in 1931–40 to 1941–50 but was significantly higher during the 1991–2000 to 2001–08 period. Interestingly, the contribution of the oldest age group (60–80) shot up from a nominal share (below 5%) to $>35\%$ during the same period.

It must be noted that the data for 2001–08 present a different picture. Though the values are small and the period does not represent a decade, the figures show a reversal from the previous trend. It indicates a high contribution of improvement in life expectancy from the oldest age group than from the youngsters as a result of greater reduction in mortality among adults and the older age groups when compared with the youngsters. Therefore, the emerging trends may indicate the beginning of a fourth stage of mortality reduction in Kerala as it has happened in developed countries. Moreover, the new trend is more evident in females than in males; the females are in advanced momentum of changes in life expectancy, whereas males are slow in decline in mortality during their adulthood and in the older age groups.

Contribution to total changes in life expectancy by causes of death

The total changes in the life expectancy, absolute as well as its percentage contributions of years by causes of deaths to total changes in the life expectancies at selected exact ages in urban areas in Kerala are given in Tables 7 and 8. Table 7 shows that life expectancy in Kerala has been increased by 7.92 years for males and 11.13 years for females during the period 1976 to 2000–04 in urban areas. Notably, all the groups of causes of death have contributed positively to these changes. However, higher share of the contribution is from infectious and parasitic diseases which was contributed 32 and 24 for percentage respectively for both males and females during this period. Also, contribution by non-communicable diseases, accidents, and injuries is commendably low in the same period.

The mortality change of a specific cause may differ from one period to another and therefore, the contribution of this cause may also vary with time as shown in Table 8. The table shows the changes in life expectancy by the contribution of different groups of diseases in the urban areas of Kerala for two time-periods (from 1976 to 1990–94 and from 2000 to 2004). It is reflected that the life expectancy in Kerala has improved by 5.95 years for males and by 8.59 years for females in 1976 to 1990–94

Table 7. Contribution of mortality changes by causes of deaths to the total increment in life expectancy in 1976 to 2000–04, Kerala* (Urban)

	Gain in LE	Infectious and parasites	Circulatory diseases	Neoplasm	Respiratory diseases	Digestive diseases	Endocrine, nutritional, metabolic diseases	Accidents, poisonings and injuries	Others ^
Males (1976 to 2000–04)									
0–1	1.66	0.52	0.00	0.01	0.26	0.08	0.10	0.01	0.43
%	100.00	31.52	0.14	0.32	15.68	4.85	5.92	0.31	25.68
1–4	1.08	0.34	0.01	0.01	0.15	0.07	0.05	0.01	0.22
%	100.00	31.17	0.82	1.05	14.19	6.10	4.78	0.81	20.62
5–14	0.39	0.18	−0.01	0.02	−0.01	0.06	0.01	0.00	0.06
%	100.00	46.38	−3.82	5.89	−1.60	14.86	3.02	−0.12	14.40
15–44	1.19	0.37	−0.01	0.11	−0.03	0.16	0.05	0.03	0.16
%	100.00	30.90	−0.64	9.60	−2.34	13.03	4.21	2.84	13.41
45–64	2.28	0.67	0.48	0.15	−0.01	0.19	0.06	0.02	0.16
%	100.00	29.55	21.05	6.65	−0.64	8.33	2.84	0.95	6.82
65 +	1.30	0.46	0.01	0.18	−0.21	−0.03	−0.15	0.12	0.04
%	100.00	35.26	0.92	13.70	−16.39	−2.16	−11.27	8.93	2.79
All	7.91	2.54	0.48	0.48	0.15	0.52	0.13	0.19	1.06
%	100.00	32.16	6.08	6.12	1.91	6.59	1.64	2.34	13.37
Females (1976 to 2000–04)									
0–1	2.26	0.70	0.03	0.01	0.35	0.11	0.09	0.01	0.60
%	100.00	30.92	1.40	0.27	15.57	4.74	3.87	0.64	26.67
1–4	1.92	0.69	0.03	0.02	0.25	0.10	0.07	0.02	0.40
%	100.00	35.68	1.82	0.92	13.17	5.00	3.56	1.04	20.92
5–14	0.53	0.16	0.01	0.02	0.03	0.05	0.01	0.03	0.11
%	100.00	29.67	2.63	3.66	5.40	9.87	2.30	5.14	21.00
15–44	1.21	0.18	0.11	0.07	0.01	0.19	0.01	−0.06	0.21
%	100.00	14.57	8.83	5.94	0.73	15.40	0.71	−5.05	17.42
45–64	2.28	0.40	0.66	0.35	0.05	0.12	0.08	0.03	0.21
%	100.00	17.55	29.13	15.53	2.22	5.17	3.51	1.33	9.23
65 +	2.92	0.57	0.35	0.17	−0.18	−0.15	−0.18	0.36	0.64
%	100.00	19.60	12.09	5.81	−6.08	−5.07	−6.18	12.43	21.99
All	11.13	2.69	1.21	0.64	0.52	0.41	0.08	0.39	2.18
%	100.00	24.19	10.83	5.74	4.63	3.70	0.68	3.55	19.60

Source: Calculated from the compiled figures of medically certified causes of deaths in 1976 and 2000–04 of Trivandrum, Kochi, Kollam, and Kozhikode Corporations and Alappuzha municipality.
*Figures are unadjusted for non-classification of causes of deaths, ^comprise other diagnosed causes of death.

periods in Kerala while it is only 1.97 years and 2.54 years, respectively, for males and females between 1990–94 and 2000–04. Between 1972 and 1990–94, the contribution by the infectious and parasites groups was comparatively higher for both males (1.86 years) and females (2.06) in the urban areas. However, between 1990–94 and 2000–04, the contribution from this group dwindled for both males (0.74 years) and females (0.84 years).

The contribution from circulatory and digestive diseases, and accidents and injuries increased in the latest period. The circulatory diseases, which were making a negative contribution to the life expectancy in 1976 to 1990–94, became a positive contributor in 1990–94 to 2000–04, with a contribution of 1.32 years for males and 1.94 for females. On the contrary, in 1990–2004, the contributions of neoplasm (only females), endocrine, and nutritional and metabolic diseases were negative, whereas they were contributed positively between 1976 and 1990–94. Though the contribution of males and females are roughly similar, females have higher contribution in most of the groups, especially from circulatory diseases. Also, the low contribution of females toward neoplasm, and respiratory and digestive diseases are a matter of concern.

The table 8 also indicates contributions made by each cause to changes in life expectancies at different age groups that greatly vary among both males and females. The contributions by infectious diseases and parasites to

Citation: Glob Health Action 2014, **7**: 22938 - http://dx.doi.org/10.3402/gha.v7.22938

Table 8. Absolute contributions of mortality changes by causes of deaths to the total increment in life expectancy in 1976 to 1990–94 and 1990–94 to 2000–04, Kerala* (Urban)

Cause of death	Gain in LE	Infectious and parasites	Circulatory diseases	Neoplasm	Respiratory diseases	Digestive diseases	Endocrine, nutritional, metabolic diseases	Accidents, poisonings and injuries	Others ^
Male (1976 to 92)									
0–1	1.33	0.44	0.02	0.01	0.23	0.07	0.09	0.01	0.21
1–4	0.90	0.24	0.02	0.01	0.14	0.05	0.05	0.00	0.18
5–14	0.50	0.17	0.04	0.03	0.01	0.05	0.01	0.02	0.08
15–44	0.51	0.22	−0.05	0.08	−0.06	0.03	0.03	−0.26	0.14
45–64	1.28	0.46	−0.04	0.12	−0.04	−0.03	0.07	−0.12	0.31
65 +	1.42	0.33	−0.60	0.46	−0.01	−0.14	0.00	−0.02	0.48
Total	5.95	1.86	−0.61	0.70	0.27	0.04	0.25	−0.37	1.41
Male (1992 to 2000–04)									
0–1	0.30	0.08	−0.01	−0.01	0.02	0.01	0.00	0.00	0.22
1–4	0.16	0.10	−0.02	0.01	0.01	0.01	0.00	0.01	0.04
5–14	−0.13	0.01	−0.05	−0.01	−0.02	0.00	0.00	−0.02	−0.03
15–44	0.69	0.16	0.04	0.04	0.03	0.13	0.02	0.27	0.03
45–64	1.07	0.22	0.57	0.04	0.03	0.24	−0.01	0.16	−0.18
65 +	−0.13	0.17	0.80	−0.36	−0.28	0.15	−0.20	0.18	−0.59
Total	1.97	0.74	1.32	−0.30	−0.20	0.53	−0.18	0.58	−0.51
Female (1976–92)									
0–1	1.85	0.61	0.04	0.01	0.31	0.09	0.08	0.01	0.35
1–4	1.49	0.51	0.03	0.02	0.17	0.07	0.06	0.00	0.31
5–14	0.47	0.12	0.04	0.02	0.03	0.04	0.01	0.02	0.10
15–44	1.00	0.25	0.09	0.04	0.07	0.08	0.02	0.04	0.21
45–64	1.06	0.16	0.11	0.28	0.03	−0.01	−0.03	−0.04	0.19
65 +	2.71	0.41	−0.73	0.38	0.17	−0.13	0.02	0.19	1.09
Total	8.59	2.06	−0.42	0.74	0.78	0.14	0.15	0.20	2.25
Female (1992 to 2000–04)									
0–1	0.35	0.07	−0.01	0.00	0.03	0.01	0.01	0.00	0.25
1–4	0.41	0.17	0.01	0.00	0.08	0.02	0.01	0.02	0.09
5–14	0.06	−4.91	7.07	0.25	1.49	−1.89	−0.63	−1.32	−0.16
15–44	0.18	0.08	−0.09	−0.01	0.00	0.04	0.01	0.01	0.14
45–64	1.30	0.26	0.62	0.07	0.02	0.14	0.13	0.08	0.01
65 +	0.24	0.21	1.45	−0.29	−0.46	−0.02	−0.27	0.23	−0.61
Total	2.54	0.84	1.93	−0.23	−0.34	0.22	−0.12	0.37	−0.12

Source: Calculated from the compiled figures of medically certified causes of deaths in 1976 and 2000–04 of Trivandrum, Kochi, Kollam, and Kozhikode Corporations and Alappuzha municipality.

*Figures are unadjusted for non-classification of causes of deaths, ^comprise other diagnosed causes of death.

life expectancy improvement were slightly higher among the younger age groups between 1976 and 1990–94. But this difference disappeared in the recent period (1990–94 to 2000–04) due to miniscule contributions from all age groups. Also, contributions of major causes of deaths such as circulatory diseases increased considerably in the recent period. The overall contribution of circulatory diseases was negative in the 1976 to 1990–94 period, mainly due to low or even negative contributions from the adult and the older age groups. However, in contrast, the recent period witnessed a positive turn in their contribution to life expectancy, which is considerably increasing with ages. Almost similar change is also visible in the accidents, poisoning, and the injuries group. In contrast, the contribution of neoplasm and respiratory diseases came down to a negative level in the older age groups between 1992 and 2000–04.

In short, it is understandable that the improvement in life expectancy slowed down in the recent decade in Kerala. Though the contribution of major cause of deaths like infectious and parasites, circulatory diseases contributed positively, it was not significant enough to move up

the entire life expectancy to the highest ages of life span. Also, contributions of other major groups such as neoplasm and respiratory diseases worsened. However, a tendency of increasing contributions to the life expectancy from the oldest age groups is visible reflecting a movement toward the advanced stages mainly among females.

Discussion

A major focus of the paper is to identify the period during which mortality transition and improvement in human longevity significantly happened in Kerala. Our analysis found that, though Kerala has experienced a drastic decline in mortality and a resultant impressive growth in life expectancy throughout the past century, the major reduction occurred between 1951 and 1970. Decline in child and infant mortality rates has played an important role during the period, resulting in drastic reduction in mortality. These findings corroborate with that of others, especially Caldwell, who pointed out the historical time period when the state had experienced the highest mortality reduction by using macro level indicators. Moreover, such corroboration also nullifies the arguable inconsistency in the core trend in mortality decline that stems from usage of different data sets for estimation over a centurial period. Also, the higher reduction of mortality achieved through a reduction in infant and child mortality was the result of healthcare intervention by the state through effective primary healthcare programs such as building awareness, vaccination, and so on, as pointed by the scholars. This draws parallels between Kerala's experience and that of other developing countries where primary healthcare has been more dominant than the socioeconomic improvements in the early stages of mortality reduction. Moreover, it upholds the fruitfulness of state intervention in mortality decline – a concept that can be adopted in the developing countries where the socioeconomic factors slow down the pace of mortality changes.

The second aim of the analysis was to understand the changes in the mortality decline and improvement in life expectancy in different age and sex groups to figure out the possibility of mortality transition reaching an advanced stage in the state. Our speculation was that mortality reduction would be higher in the older age groups denoting the onset of advanced stages in the state. The analysis of TLE reveals that the age groups below 60 (except males at 40–60) have already reached their limit of mortality reduction in contributing to life expectancy. Thus, the possibility of a further mortality decline that can contribute to an improvement in life expectancy is confined to the older age groups. However, the analysis of Index of ARC in TLE found that the pace of TLE in the older age groups (40–60 and 60–80) is relatively lower than in youngsters, implying low mortality decline in the older age groups.

Median age at death and survival proportional to age 60 and 80 reflect the relative importance of reduction in old age mortality for further increasing life expectancy in the state. However, the contribution of different age groups to life expectancy until 1991–2000 shows that reduction in mortality in adults and the older age groups contributed <50 percent to increments in life expectancy. It indicates that the state has not yet entered into the advanced stage of mortality transition. Nevertheless, there is a recent overall tendency of increasing contribution from older age groups toward life expectancy. This could point to an emergence of this advanced stage, particularly among females.

The slow reduction in the mortality rates among adults and elders are associated with the ongoing epidemiological transition and mismatches in health policies in Kerala. The third aim of our analysis, the decomposition analysis on changes in life expectancy by the major causes of death is corroborating with this finding. Notably, the state is experiencing more deaths due to non-communicable diseases such as cardiovascular diseases, neoplasm, accidents, and injuries by shifting the dominance of deaths from infectious diseases and maternal and child deaths in the recent decades as part of ongoing epidemiological transition. These causes affected the adults and elders to a large extent. The shift in the causes of death necessitates a different level of healthcare mechanism. In other words, the post-1971 epidemiological stage in Kerala requires promotive and curative intervention in addition to maintaining the existing primary healthcare facilities. In view of this, any delay in the onset of the advanced stage of mortality transition is caused by lower mortality decline of adults and elders and raises concerns over the efficiency of the existing healthcare system in tackling the additional challenges.

At this juncture, the experience of developed countries such as the United States and western European countries can shed some light. The basic feature of advanced mortality reduction in these countries was the reduction in morbidity mainly as a result of a combination of preventive and health-promotive measures. Moreover, supplementing their successful primary healthcare model with a supportive intervention by providing accessible healthcare facilities and financial support through health insurance also contributed to their success. However, such focused strategies are not yet widely implemented in Kerala. Since Kerala already has a good network of primary healthcare services, the focus should shift to the health problems of adults and address concerns related to their morbidity and mortality. The current trends in disease pattern suggest that different strategies other than the typical primary healthcare intervention followed are necessary for addressing the emerging changes in the disease pattern in the state.

Citation: Glob Health Action 2014, **7**: 22938 - http://dx.doi.org/10.3402/gha.v7.22938

Main findings

- Health issues of infants, children, and mothers of reproductive age have been effectively addressed by various policies in Kerala, leading to dramatic decreases in overall mortality and increase in life expectancy during the last century.
- Further mortality reduction is possible in adult and early old age groups, which could further increase life expectancy in Kerala.
- The limited reduction in mortality in the adult and early old age groups in Kerala during the last century is mainly due to the effects of highly prevalent non-communicable diseases. This also indicates a slow progress towards the advanced stages of epidemiological transition.

Key messages for action

- Health status of adult and elderly men is unsatisfactory, particularly given the large differences in mortality observed between men and women in these age groups; future government policy should prioritise their health needs.
- High mortality and morbidity in Kerala are now mainly due to non-communicable diseases; planning and policy in healthcare should reflect these changes in disease patterns.
- Priority should be given to strategies for disease prevention and health promotion, in addition to existing curative healthcare services, in order to curb the challenges from non-communicable diseases.

Conflict of interest and funding

The authors have not received any funding or benefits from industry or elsewhere to conduct this study.

References

1. Franke RW, Chasin BH. Is the Kerala model sustainable? Lessons from the past, prospects for the future. In: Parayil G, ed. Kerala: The development experience, reflections on sustainability and replicability. London: Zed Books; 2000, pp. 16–39.
2. Parameswaran MP. What does the Kerala model signify? Towards a possible 'fourth world'. In: Parayil G, ed. Kerala: the development experience, reflections on sustainability and replicability. London: Zed Books; 2000, pp. 232–48.
3. Centre for Development Studies (CDS) (1975). Poverty, unemployment and development policy: a case study of selected issues with reference to Kerala. New York: United Nations.
4. Panikar PGK. Health transition in Kerala. Discussion paper series. Trivandrum: Kerala Research Program on Local Level Development, Centre for Development Studies; 1999.
5. Sen A. Radical needs and moderate reforms. In: Dreaze J, Sen A, eds. Indian development: selected regional perspective. New Delhi: Oxford University Press; 1997, pp. 1–32.
6. Parayail G. Introduction: is Kerala's development experience a 'model'? In: Parayil G, ed. Kerala: the development experience, reflections on sustainability and replicability. London: Zed Books; 2000, pp. 1–15.
7. Panikar PGK, Soman CR. Health status of Kerala: paradox of economic backwardness and health development. Trivandrum: Center for Development Studies; 1984.
8. Caldwell JC. Routes to low mortality in poor countries. Popul Dev Rev 1986; 12: 171–220.
9. Krishnan TN. Health statistics in Kerala state, India. In: Halstead SB, Walsh JA, Warren KS, eds. Good health at low cost. New York: Rockefeller Foundation; 1985, pp. 39–46.
10. Bhat PN, Rajan SI. Demographic transition in Kerala revisited. Econ Polit Wkly 1990; 25: 1957–80.
11. James KS. The historical nature of demographic transition in Kerala: estimates of vital rates using parish records. Demography India 2001; 30: 13–30.
12. Omran AR. The epidemiological transition: a theory of epidemiology of population change. Milbank Q 1971; 49: 509–38.
13. Olshansky A, Ault AB. The fourth stage of epidemiological transition; the age of delayed degenerative diseases. Milbank Q 1986; 64: 355–91.
14. Navaneetham K, Kabeer M, Santhakumar V. Emerging morbidity patterns in Kerala. In: Ashokan A, ed. Perspective of health economics. New Delhi: Serial Publication; 2009, pp. 34–67.
15. Ashokan P. Acute morbidity in rural Kerala: levels differentials and determinants. In: Ashokan P, ed. Perspective of health economics. New Delhi: Serial Publication; 2009, pp. 96–107.
16. Hameed SS, Kutty VR, Vijayakumar K, Kamalasanan A. Migration status and prevalence of chronic diseases in Kerala State, India. Int J Chron Dis 2013; 2013: 1–6.
17. Thomas MB. Changes in mortality in Kerala, India: some emerging concerns. Univers J Educ Gen Stud 2012; 1: 234–41.
18. Namboodiri K. The changing population of Kerala. Census Monograph Series, No.7. New Delhi: Registrar General; 1968.
19. Namboodiri K. A primer of population dynamics. New York: Plenum Press; 1961.
20. Bhattacharjee PJ, Shastri GN. Population in India: a study of interstate variations. New Delhi: Vikas Publications House; 1976.
21. Nag M. Impact of social and economic development on mortality: comparative study of Kerala and West Bengal. Econ Polit Wkly 1983; 18: 877–900.
22. Coale AJ, Demeny P, Vaughan B. Regional model life table and stable populations. New York: Academic Press; 1983.
23. Bhat PNM. Mortality in India: levels, trends and patterns. PhD dissertation. Unpublished: University of Pennsylvania; 1987.
24. Arriaga EE. Measuring and explaining the change in life expectancies. Demography 1984; 21: 83–96.
25. Feachem RGA, Philips MA, Bulatao RA. Introducing adult health. In: Feachem RGA, Kjellstrom T, Murray CJL, Over M, Philips MA, eds. The health of adults in developing world. London: Oxford University Press; 1992.
26. Murray CJL, Yang G, Qiao X. Adult mortality: levels, patterns and causes. In: Feachem RGA, Kjellstrom T, Murray CJL, Over M, Philips MA, eds. The health of adults in developing world. London: Oxford University Press; 1992.
27. Dutton J. Changes in Soviet mortality pattern 1959–1977. Popul Dev Rev 1979; 5: 267–91.
28. Preston SH, Hueline P, Guillot M. Demography: Meassuring and Modelling Population Process. Oxford: Blackwell Publishers; 2013.

Global Health Action

PART III

The epidemiological transition in Antananarivo, Madagascar: an assessment based on death registers (1900–2012)

Bruno Masquelier[1]*, Dominique Waltisperger[2], Osée Ralijaona[3], Gilles Pison[2] and Arsène Ravélo[4]

[1]Centre de recherches en démographie et sociétés, Université catholique de Louvain (UCL), Louvain-la-Neuve, Belgium; [2]Institut National d'Etudes Démographiques (INED), Paris, France; [3]Health Statistics Division, Ministry of Health, Antananarivo, Madagascar; [4]Office of demography and social statistics, Institut National de la Statistique de Madagascar (INSTAT), Antananarivo, Madagascar

Background: Madagascar today has one of the highest life expectancies in sub-Saharan Africa, despite being among the poorest countries in the continent. There are relatively few detailed accounts of the epidemiological transition in this country due to the lack of a comprehensive death registration system at the national level. However, in Madagascar's capital city, death registration was established around the start of the 20th century and is now considered virtually complete.

Objective: We provide an overview of trends in all-cause and cause-specific mortality in Antananarivo to document the timing and pace of the mortality decline and the changes in the cause-of-death structure.

Design: Death registers covering the period 1976–2012 were digitized and the population at risk of dying was estimated from available censuses and surveys. Trends for the period 1900–1976 were partly reconstructed from published sources.

Results: The crude death rate stagnated around 30‰ until the 1940s in Antananarivo. Mortality declined rapidly after the World War II and then resurged again in the 1980s as a result of the re-emergence of malaria and the collapse of Madagascar's economy. Over the past 30 years, impressive gains in life expectancy have been registered thanks to the unabated decline in child mortality, despite political instability, a lasting economic crisis and the persistence of high rates of chronic malnutrition. Progress in adult survival has been more modest because reductions in infectious diseases and diseases of the respiratory system have been partly offset by increases in cardiovascular diseases, neoplasms, and other diseases, particularly at age 50 years and over.

Conclusions: The transition in Antananarivo has been protracted and largely dependent on anti-microbial and anti-parasitic medicine. The capital city now faces a double burden of communicable and non-communicable diseases. The ongoing registration of deaths in the capital generates a unique database to evaluate the performance of the health system and measure intervention impacts.

Keywords: *epidemiological transition*; *Antananarivo*; *death registers*; *mortality*; *Madagascar*; *vital statistics*; *Africa*

Responsible Editors: Nawi Ng, Umeå University, Sweden; Barthélémy Kuate Defo, University of Montreal, Canada.

*Correspondence to: Bruno Masquelier, Centre de recherches en démographie et sociétés, Université catholique de Louvain, SSH/IACS/DEMO, Collège Jacques Leclercq, Place Montesquieu 1, bte L2.08.03, 1348 Louvain-la-Neuve, Belgium, Email: bruno.masquelier@uclouvain.be

This paper is part of the Special Issue: *Epidemiological Transitions – Beyond Omran's Theory*. More papers from this issue can be found at http://www.globalhealthaction.net

Received: 31 October 2013; Revised: 11 January 2014; Accepted: 2 February 2014; Published: 15 May 2014

The theoretical model of the epidemiological transition describes a long-term shift from a regime of high and fluctuant mortality, dominated by infectious diseases, to a regime of low mortality where deaths are predominantly due to non-communicable diseases (NCDs) linked with population aging and changes in lifestyles (such as increasing fat and calorie consumption, decreasing exercise, and tobacco use) (1). In western countries, there has been considerable debate over the chronology of this transition and the contribution of better nutrition, rises in living standards and advances in medical science to the decline in mortality (2, 3). The relevance of this model for Latin America and Asia has also been widely debated (4–7). Less attention has been

Citation: Glob Health Action 2014, **7**: 23237 - http://dx.doi.org/10.3402/gha.v7.23237

devoted to sub-Saharan Africa (SSA), apart from a few case studies on Accra (8), South Africa (9), and Mauritius (10). In SSA, the epidemiological transition started only after the World War II, largely thanks to public health measures that were internationally supported (11). It has progressed at a much slower pace than in other developing regions. For instance, although life expectancy was around 36 years in SSA and India in 1950–1955, it had reached 65 years in India in 2005–2010, against only 53 in SSA (12). Over the past 30–40 years, child mortality has declined dramatically in most African countries, but conspicuous stalls were observed in the 1990s and early 2000s. These setbacks were related to the HIV epidemic, temporary declines in immunization rates and upsurges in malaria-related mortality (13). Declines in child mortality have not been matched by sustained improvements in adult survival either. Since the 1990s, most countries in eastern and southern Africa have witnessed enormous surges in adult mortality because of AIDS. Progress in adult survival has been limited even in countries whose HIV epidemics have been small. Madagascar is an exception in this regard, because adult and child mortality rates have evolved in concert toward a regime of low mortality, as they have in Ethiopia or Senegal. According to estimates from the Demographic and Health Surveys (DHS), the probability that a newborn would die before reaching age 5 ($_5q_0$) in Madagascar was 0.20 in 1985 and had declined to 0.08 by 2005. The probability that a person aged 15 would die before reaching age 60 ($_{45}q_{15}$) was around 0.36 in 1985 (among males) and had declined to 0.28 by 2005 (13). As a result, no country in SSA had a higher life expectancy than Madagascar during the period 2005–2010 (62 years), except for small islands such as Cape Verde and Mauritius (12). And yet Madagascar is among the poorest countries in the region, with 93% of the population living on less than $2 a day (14).

This disconnection between health and economic progress calls for a detailed examination of the epidemiological transition occurring in Madagascar. To date, this has been hindered by the lack of long-term series of data on mortality by cause, due to the low coverage of death registration, estimated at less than 50% at the national level (15). Although this low coverage is a feature of most countries in SSA, the vital registration system is atypical in Madagascar because it was established relatively early, in 1878, by the Merina Kingdom, which ruled over most of the island at that time. The system only covered areas controlled by this Kingdom, however, and also did not include the slave population. The French colonial administration officially abolished slavery in 1896 and enforced vital registration throughout the island. Vital registration offices were set up in all districts, and Antananarivo, the capital city, was considered an administrative unit on its own, with its own offices (16). Since then, the coverage of death registers has always been very high in the city, in contrast to other parts of the country. In this paper, we take advantage of this wealth of data and describe the epidemiological transition that occurred in city over the period 1900–2012. We first provide some historical background and introduce our data sources and our methods, and then we explore the changes in patterns of mortality by age groups and cause.

Historical background

Antananarivo, formerly known as Tananarive, is located in the central highlands of Madagascar at an altitude of 1,250–1,470 m. It is spread across 90 km^2. Its climate is subtropical with a cold and dry season from May to October (with an average minimum temperature of around 10°C) and a hot and rainy season from November to April (with an average maximum temperature of around 27°C).

At the end of the 19th century, the mortality regime of the city was characterized by high death rates from malaria (causing about 25% of deaths in 1895), smallpox, syphilis, and tuberculosis (17). Epidemics were frequent; Antananarivo was hit by smallpox epidemics in 1875–1881 and 1884–1889, a flu epidemic in 1890 and a typhoid epidemic in 1894 (17, 18). The incidence of infectious diseases apparently increased when the Merina Kingdom reinforced its coercive policies from the 1870s. Massive movements of units of forced labor contributed to the spread of malaria over the central highlands in 1878, in conjunction with the resettlement in the valleys of populations that previously lived in fortified centers located over the hills.

After the French takeover, the colonial administration acted quickly to implement measures to increase the health of the indigenous workforce and promote population growth. By 1896, the city had a hospital and a medical school to train local doctors and midwives. The Indigenous Medical Assistance was established the following year and provided health care to the population free of charge. A Pasteur Institute was founded in 1899 to produce locally the treatment for rabies and the smallpox vaccine. Routine immunization campaigns were organized and smallpox was considered well controlled in 1908.

Unfortunately, these measures were partly offset by the expansion of malaria. This is apparent in the mortality data routinely collected in Antananarivo that we describe below. In February 1906, during the rainy season, 57% of deaths were due to malaria, against only 10% in February 1903; it went back down to 35% in February 1907. This expansion of malaria was linked to a broad set of factors, including large movements of population induced by infrastructure projects, changes in the methods used for rice cultivation (rice plants were no longer dried out after the harvest), and the settlement of some populations

closer to the rice fields (19). The French administration established quinine depots, but malaria nonetheless became endemic in the central highlands.

The colonial administration also developed the city by clearing some marshland, setting up a public lighting system in 1910 and installing the first standpipes in 1911 (20). In 1916, a Municipal Office of Hygiene (*Bureau municipal d'hygiène* – BMH) was created. The BMH worked under the supervision of the mayor to deliver free consultations, conduct immunization campaigns, provide malaria prophylaxis, and isolate patients affected by highly infectious diseases such as the plague. The plague reached Antananarivo in 1921, and up to 140 deaths were reported in 1927 (when the city had 73,000 inhabitants). The BMH responded to this outbreak with very drastic measures including isolation of infected patients into plague houses and the disinfection or destruction by fire of the patients' homes. The introduction of sulfonamides and streptomycin helped reduce the incidence of the plague and no human cases were notified between 1949 and 1978, when it reemerged in the city.

Penicillin and other antibiotics were introduced in Antananarivo in the late 1940s and early 1950s, but little is known on the timing of their dissemination and their use in medical facilities. The importance of chloroquine in the fight against malaria is more documented (21). Beginning in 1949, the BMH launched a massive anti-malaria program, based on the indoor spraying of DDT and chemoprophylaxis administered on a weekly basis to school-age children and preschoolers. Malaria was considered under control in the central highlands in 1960, when Madagascar gained its independence, and the DDT spraying was ceased. The Expanded Program of Immunization (EPI) was launched in 1976, starting with diphtheria, tetanus, pertussis, and tuberculosis. The polio vaccine was included in the immunization schedule in 1982 and the measles vaccine in 1985.

The beneficial effects of these programs were stymied by a major economic crisis that began as early as 1972. This year marked the end of Madagascar's First Republic (1960–1972); the government was overthrown by a popular revolt after growing criticism for maintaining too strong ties with France. GNI per capita then began a steep decline that lasted until the mid-1990s (Fig. 1); it dropped from $497 in 1971 to $264 in 1996 (14). The shift toward economic insularity under the socialist-Marxist regime of Didier Ratsiraka (1975–1993) contributed to this disastrous situation. The regime took control over the agricultural production and marketing system and nationalized foreign-owned trading and industrial companies. The import trade was also put under state control. The price paid to rice producers fell and many smallholders withdrew into subsistence farming. In 1977, rice had to be imported to feed the population, despite Madagascar having been a net exporter of rice 5 years earlier. In 1981, the country removed subsidies on rice and started to implement a series of structural adjustment programs. Rice imports remained controlled by the state until 1986, even though controls over the prices paid by the consumers had been lifted. A parallel market blossomed, where rice was sold at about twice the price of the state-controlled market (22).

To make things worse, malaria resurged in Antananarivo in 1984, at a time when many treatment and prophylaxis centers had closed after their funding was cutoff by the central administration (23). The resumption of chemoprophylaxis with chloroquine in 1988 and the reintroduction of indoor DDT spraying in 1993–95 brought malaria back under control. Today, cases of autochthonous malaria are very rare in Antananarivo, but seroprevalence remains high because inhabitants come into contact with the parasite while travelling outside the city (24).

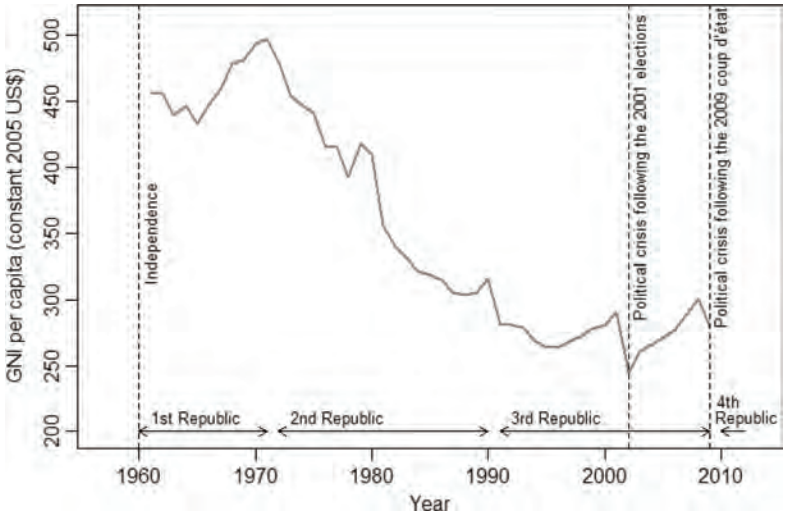

Fig. 1. Trends in GNI per capita (1961–2009), according to data from the World Bank (constant 2005 US$) (14).

Citation: Glob Health Action 2014, **7**: 23237 - http://dx.doi.org/10.3402/gha.v7.23237

Madagascar was forced to introduce a series of structural reforms and move away from state intervention in the late 1980s. Didier Ratsiraka lost the elections in 1993 but he was voted back into power in 1997 and he moved further toward economic and financial liberalization. The economic recovery was short lived as a controversy over the 2001 presidential election gave rise to violent clashes in 2002 and resulted in a 6-month standoff. GNI per capita plunged, rose again but dropped in 2009 when the president Ravalomanana was ousted from power during a coup (Fig. 1). Restrictions on aid from foreign governments and other donors, which resulted in severe cuts in public spending, followed this unconstitutional change of power. A presidential election was organized in 2013 to put an end to Madagascar's political impasse.

The economic decline and the recurrent political instability have not prevented Madagascar from achieving significant health improvements. Empirical evidence from four DHS surveys indicate both the substantial decline in under-five mortality in Madagascar over the past three decades, and a consistent advantage of Antananarivo over rural areas and other cities (Table 1) (25–28). This urban advantage has been observed in several African countries from the 19th century, because of higher incomes, higher education levels, and a disproportionate provision of modern sanitation and health care (29, 30).

In the case of Antananarivo, the lower mortality rates can be attributed to a variety of advantages. The capital has benefitted from a favorable location with a lower prevalence of malaria than in endemic coastal areas, although the region has been prone to epidemics. The city is also responsible for approximately 40% of Madagascar's GNI and a higher proportion of households have access to basic amenities than in the rest of the country. As an example, the proportion of the population using an improved drinking-water source was 98% in the capital in 2008–2009 against only 32% in rural areas and 82% in the other cities (28). In the capital, 82% of households had access to electricity against 8% in rural areas and 63% in other cities. The city is better equipped with health facilities compared to the rest of the country; it has three university teaching hospitals that correspond to the highest level of the health delivery system. In addition, large differences are found in educational attainment: above age 6, only 3% of women do not have any education in the capital, compared with 20% in Madagascar (28). These differences are a remnant of the early establishment of the school system in the central highlands by the Merina Kingdom. Possibly 7% of the adult population of the region were literate in the mid-19th century (31). In 1960, 16% of the inhabitants of Antananarivo aged 15 and over had already attended school beyond primary level (32); this proportion rose to 64% in 1995 and to 71% in 2008/2009 (28).

Table 1. Under-five mortality rates for the 10-year period preceding each DHS survey conducted in Madagascar, according to type of place of residence. Sc: (25–28)

	1992 DHS	1997 DHS	2003/04 DHS	2008/09 DHS
Antananarivo-city	152	110	43	51
Other cities	137	132	81	70
Rural areas	183	174	120	84

Materials and methods

Our assessment of the epidemiological transition in Antananarivo is based on three types of data sources: 1) monthly reports of deaths by cause for the period 1900–1907, 2) some published estimates for the period 1931–1951, and 3) individual-level data from death registers for the period 1976–2012.

First, monthly reports of deaths by cause started to be published in 1900 by the municipal administration in the *Journal Officiel de Madagascar et dépendances* (33). We accessed the reports for the period 1900–1907. They provide a breakdown of deaths by age, sex and cause, but without cross-tabulation. From this first type of data source, we will only extract the ranking of categories of causes of deaths and estimates of the crude mortality rate (CDR).

Second, starting in 1916, the BMH took over responsibility for the registration of deaths in the city and for regular reporting on mortality. Some published estimates of the CDR are available for the years 1931–1951 (21).

Third, all death registers maintained at the BMH for the period 1976–2012 were digitized, that is, about 298,000 records. When deaths occur at home (about 60% of deaths since 1976), relatives or caretakers of the deceased contact the BMH, and a physician is sent to the house to assign a cause of death, based on available medical documents and on post-mortem interviews with the family concerning the symptoms and circumstances preceding the death. For deaths occurring in hospitals or clinics, the reports are filled in by the medical personnel and transmitted to the BMH by the relatives. Since 1976, more than 80% of deaths have been reported on the day of death or the day after. A high coverage of deaths is maintained because cemeteries are guarded and reporting the death at the BMH is required to obtain a burial permit or to move the corpse. Based on data for the period 1984–1995, Waltisperger et al. showed that age-specific mortality rates based on the BMH conformed to standard age patterns of mortality, and concluded that the data presented no indication of underreporting of deaths (34).

For this analysis, we retained only the deaths of residents of Antananarivo. Causes of death were recorded according to the Ninth Revision of International Statistical

Classification of Diseases (ICD-9). The same physician was in charge of recoding all deaths that had not been previously coded in hospitals or clinics. Causes were later consolidated into seven broad categories: 1) infectious diseases (classified in the first chapter of the ICD-9), 2) neoplasms, 3) nutritional deficiencies, 4) cardiovascular diseases (CVDs), 5) diseases of the respiratory system (including influenza and bronchitis), 6) injuries and accidents, and 7) other diseases. The complete list of corresponding ICD-9 codes is provided elsewhere (22). Since 1976, the percentage of deaths whose cause was unknown or ill-defined (including senility) has ranged from 7.5 to 18%, the maximum having been attained in 1988, shortly after a mortality crisis that led to a higher-than-average proportion of deaths occurring at home. The percentage of unknown or ill-defined causes then declined to below 10%, but it has risen again in recent years due to the increasing share of deaths occurring at ages 70 and above. For the present analysis, deaths of unknown or ill-defined cause were redistributed by age group and sex among the other categories of causes, assuming that imprecise statements concern all causes in the same proportion.

To estimate population exposure, we used a logistic curve to interpolate between population counts from administrative censuses taken annually from 1901 to 1963 (16, 35), counts from the 1975 and 1993 national censuses, and the provisional count that was done in 2009 as part of the mapping of the forthcoming census (which has been delayed because of the political crisis). Our estimates refer to the 'Commune urbaine d'Antananarivo Renivohitra' (CUA), corresponding to the six districts for which we have data from the BMH. Antananarivo had 53,600 inhabitants in 1901. The population increased rapidly from 1925 to attain 200,000 inhabitants around 1957. Over the past 60 years, it has increased more than five-fold, reaching 1.03 million in 2009. Yet its annual growth rate has been relatively modest in recent years (2.4% over the period 1975–2009), compared to other African capitals such as Kinshasa (5% over the period 1975–2010) or Ouagadougou (7%). Madagascar remains weakly urbanized and migration does not play a key role in shaping its capital. In 1995, only 27% of the residents of the city were migrants, and half of them came from the province of Antananarivo (32).

In order to compute age-specific mortality rates from registered deaths for the period 1976–2012, we had to distribute this population by age group and sex, which is challenging because the last two censuses carried out were in 1975 and 1993. Our strategy was to compute by sex the share of each age group in the population enumerated in the these two censuses, and in four DHS, conducted in 1992, 1997, 2003–2004 and 2008–2009, after selecting only the households that had been sampled in the city (25–28). We used penalized smoothing splines to interpolate the share of each age group between the surveys and the censuses (and extrapolated to 2012), and multiplied these percentages by the total population, estimated as explained above, to obtain a population count by age group, sex, and calendar year.

The level of detail of the mortality data thus varies over time. Age- and cause-specific mortality rates can only be computed from 1976 onwards, but summary indicators can be obtained from 1900. In the next section, we first present trends in the CDR in Antananarivo for the period 1901–2012 alongside trends in the CDR obtained from incomplete death registration for Madagascar from 1906 to 1972 (36–38). Second, using the monthly reports published in the beginning of the 20th century and the individual-level data from 1976, we compare the rankings of the leading causes of deaths, for both sexes combined. Third, we provide estimates of the life expectancy at birth by sex and under-five mortality rates from 1976. Finally, we detail changes in the standardized mortality rates by causes (1976–2012) and estimate the contribution of age groups and causes of deaths to changes in life expectancy, based on the algorithm developed by Andreev et al. (39).

Results

Trends in the crude mortality rate (1901–2012)

Caution is obviously required in interpreting long-term trends in CDR, because they can be distorted by changes in the age structure of the population over the period and variations in the completeness of death reporting. Based on the age distribution of registered deaths in 1965–67, Courbage and Fargues (40) estimated that about half of the male deaths and over one-third of the female deaths were not registered at the national scale. According to the demographic survey conducted in 1966, the CDR was around 25‰ at that time (41), much higher than the estimate based on vital registration (13‰ in 1966). Despite this massive underreporting of deaths, we can venture to say that the CDR was at least 30‰ at the start of the 20th century in Madagascar, and could have slightly declined until 1918, when mortality peaked above 40‰ following an outbreak of the Spanish flu (Fig. 2). It remained stable between 1920 and 1948, and declined sharply only after the World War II. The apparent stall between 1955 and 1972 might reflect an increase in the completeness of death reporting, because the 1975 census estimated the CDR at 18‰, compared with 25‰ in the 1966 survey. According to the 1993 census, the CDR had further declined to 14‰ by the beginning of the 1990s.

In Antananarivo, estimates from death registration have always been considered much more reliable than those based on nationwide data. In 1901, the CDR was estimated at 28.2‰ in the city (16). Little progress was registered prior to the end of the World War II; the CDR was still around 25‰ in 1931. It reached 32‰ in 1944 and

Citation: Glob Health Action 2014, **7**: 23237 - http://dx.doi.org/10.3402/gha.v7.23237

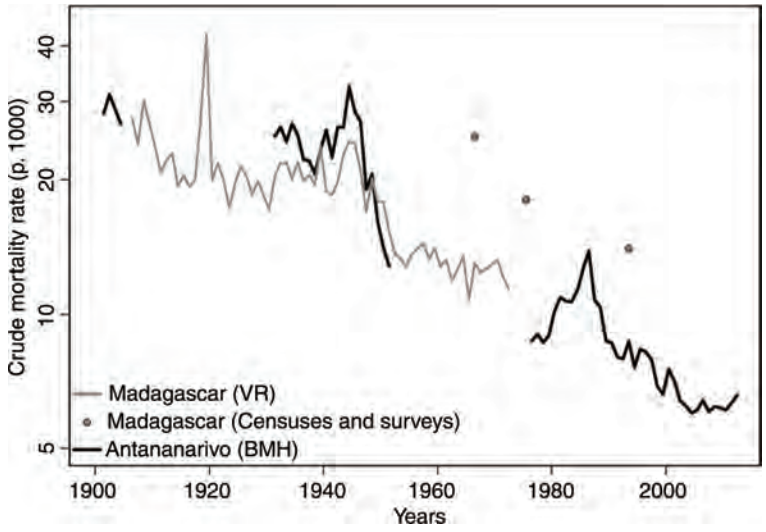

Fig. 2. Trends in the crude death rate for Madagascar and Antananarivo city (1901–2012) (logarithmic scale). Sources: Refs. (21, 36–38) and data from the BMH for the period 1976–2012.

then dropped to 13‰ in 1951 (21). This decline continued until the end of the 1970s, with the CDR reaching 8.8‰ during the period 1976–1979. From 1980, it rose again and peaked at 13.8‰ in 1986, when malaria resurged in the central highlands and increases in the price of rice resulted in acute food shortages. The CDR went down to 8.7‰ in 1989 and reached its lowest level in 2004 (6‰).

Changes in the rankings of the leading causes of deaths (1900–2012)

To investigate changes in the cause-of-death structure, we distinguish three periods: 1900–1903 (the first uninterrupted series of monthly mortality reports by cause), 1976–1980 (the first 5 years covered by the database of individual death records), and 2008–2012 (the most recent years). Because the classifications of diseases changed between 1900 and 1976, as we discuss below, we will only examine the rankings of broad categories of causes of deaths. Problems remain, however, because some categories of causes such as nutritional deficiencies (corresponding to ICD-9 codes 260–269) were not singled out in 1900–1903. Measles was not mentioned among causes of deaths either, whereas it accounted for about 4% of deaths in 1976–1980. Despite this, this comparison reveals three general features (Table 2). First, there was a stark decline in the contribution of diseases of the respiratory system (pneumonia in particular), from 38% in 1900–1903 to 12.5% in 1976–1980 (and 10% in the last period). Second, apart from pneumonia, infectious and parasitic diseases continued to occupy a major position in the cause-of-death profile in 1976–1980. If all deaths from nutritional deficiencies had been reallocated to infectious causes, the proportion of deaths due to infections would have only declined from 34% in 1900–1903 to 29% in 1976–1980. This proportion plunged to 11.9% in 2008–2012. Most of the decline that happened before 1980 was due to reductions in malaria- and tuberculosis-related mortality, with intestinal infections still accounting for about 15% of all deaths in 1976–1980. Third, CVDs, neoplasms and trauma gained in importance, accounting for 32% of deaths in 1976–1980 and 51% in 2008–2012, against only 9% at the start of the 20th century.

Trends in life expectancy and under-five mortality (1976–2012)

In 1976, the life expectancy at birth had reached 57 years for males but it dropped to 46 in 1986 (Fig. 3). Over the same period, females lost 5 years of life expectancy, from 59 to 54. The under-five mortality rate, which was estimated at 114‰ in 1976, reached 190‰ in 1983. This peak is also evident in estimates from DHS, obtained by pooling together all birth histories of women living in the capital and interviewed in 1992, 1997, 2004 and 2009. It is worth noting that DHS estimates follow quite closely those of the BMH, in support of the hypothesis that the registration of deaths is almost complete in Antananarivo. Adults were also severely affected by this crisis; the probability of dying between ages 15 and 60 peaked at 0.51 among males in 1986, against 0.29 in the period 1976–1979.

Since 1986, Antananarivo has benefitted from an unabated decline in mortality. Life expectancy has risen steadily to reach 62.4 years for males and 68.6 for females in 2012. The under-five mortality rate has decreased fivefold from its level in 1983. The decline in mortality in the city has been slightly more rapid than observed in the rest of the country. In Fig. 3, estimates for the rest of the country are again obtained by pooling together the birth

Table 2. Rankings of the leading causes of deaths in 1900–1903, 1976–1980, and 2008–2012 in Antananarivo

Rank	Causes of death (1900–1903, all ages and both sexes)	Prop. (%)	Rank	Causes of death (1976–1980, all ages and both sexes)	Prop. (%)	Rank	Causes of death (2008–2012, all ages and both sexes)	Prop. (%)
1	(Broncho)-pneumonia, influenza, ARI	30.6	1	Cardiovascular diseases	21.6	1	Cardiovascular diseases	35.8
2	Intestinal infections	15.9	2	Intestinal infections	14.8	2	Other causes	10.6
3	Malaria	7.9	3	Other causes	10.5	3	Neoplasms	8.5
4	Tuberculosis	7.5	4	(Broncho)-pneumonia, influenza, ARI	9.6	4	Injury and accidents	6.7
5	Other diseases of respiratory system (including bronchitis)	7.2	5	Other infectious and parasitic diseases (incl. measles)	5.6	5	Diseases of the digestive system	5.5
6	Other causes	6.7	6	Neoplasms	5.5	6	Other diseases of respiratory system (including bronchitis)	5.2
7	Congenital anomalies and other perinatal conditions	6.5	7	Nutritional deficiencies	5.1	7	(Broncho)-pneumonia, influenza, ARI	4.8
8	Cardiovascular diseases	6.5	8	Injury and accidents	5.0	8	Congenital anomalies and other perinatal conditions	4.3
9	Whooping cough	2.1	9	Diseases of the digestive system	5.0	9	Tuberculosis	4.1
10	Diseases of genitourinary organs	1.9	10	Congenital anomalies and other perinatal conditions	4.5	10	Diseases of genitourinary organs	3.3
11	Neoplasms	1.9	11	Diseases of genitourinary organs	3.0	11	Diseases of the nervous system	2.8
12	Diseases of the nervous system	1.8	12	Other diseases of respiratory system (including bronchitis)	2.9	12	Other infectious and parasitic diseases (incl. measles)	2.6
13	Complications related to pregnancy, labor and delivery	1.2	13	Diseases of the nervous system	2.6	13	Nutritional deficiencies	2.0
14	Diseases of the digestive system	1.2	14	Tuberculosis	1.4	14	Intestinal infections	1.7
15	Injury and accidents	0.8	15	Whooping cough	1.4	15	Malaria	1.5
16	Diphtheria	0.4	16	Complications related to pregnancy, labor and delivery	0.9	16	Complications related to pregnancy, labor and delivery	0.6
			17	Malaria	0.4	17		
			18	Diphtheria	0.3	18		

histories collected in the four DHS and excluding reports from mothers living in Antananarivo.

Changes in cause-specific mortality rates

Figure 4 presents, on a log scale, the standardized mortality rates for the main categories of causes of death (using the age structure of 1993 as a standard). The peak in 1984–1988 was predominantly due to a sheer increase in death rates from infectious diseases (malaria in particular) and nutritional deficiencies.

After 1986, the rapid decline in mortality from nutritional deficiencies is particularly salient. A negligible fraction of deaths are now due to these deficiencies. Infectious diseases still remain among the major killers, but have declined markedly, especially diseases targeted by the EPI. Diseases of the respiratory system have also declined since 1986 but to a lesser extent. By contrast, little progress has been made in the reduction of injuries, accidents and CVDs since the end of the 1980s.

In Fig. 5, we show the specific contribution of each category of causes in the change in life expectancy between two periods: 1990–1994 (when mortality had stabilized at levels observed prior to the crisis) and 2008–2012 (the past few years). For both sexes, the reduction of under-five mortality explains more than 80% of the gains in life expectancy, and about half of these gains in childhood are

Citation: Glob Health Action 2014, **7**: 23237 - http://dx.doi.org/10.3402/gha.v7.23237

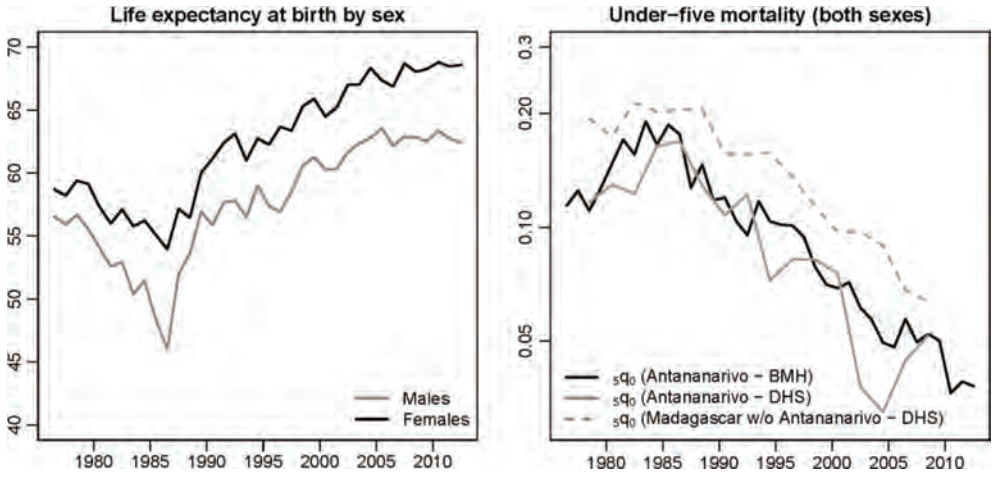

Fig. 3. Trends in life expectancy at birth by sex and the under-five mortality rate (on a logarithmic scale) in Antananarivo city (1976–2012), according to data from the BMH and DHS.

related to reductions in infectious diseases, the other half being mostly attributable to reductions in nutritional deficiencies and diseases of the respiratory system. Gains in life expectancy have been much more modest among the population aged 5 and over; reductions in infectious and respiratory diseases have been largely offset by increases in mortality from CVDs, neoplasms and other diseases, particularly at ages 50 and over.

Discussion

Death registers of Antananarivo represent a valuable population-based mortality database for Madagascar, but they nonetheless suffer from important shortcomings. First, although deemed very high, the completeness of death reporting in the capital remains unknown and it certainly changed over time. For the period prior to 1960, there is no estimate of the completeness; we can only count on the fact that the health situation was more closely monitored in the capital than anywhere else because it was the seat of the General Government and had a large European population (16). After 1960, there are some signs that most deaths were captured in the vital registration system. In 1965, the CDR for the larger area of the province in which the city is located (10–11‰) was very close to the estimate based on the 1966 demographic survey (13–19‰). Virtually all infant deaths were reported in the province of Tananarive in 1965, compared to only 40% of infant deaths in the provinces of Diego-Suarez and Tulear (42). For the more recent period, we showed that trends in under-five mortality derived from death registers largely conform to those obtained from DHS. To go beyond these encouraging signs, however,

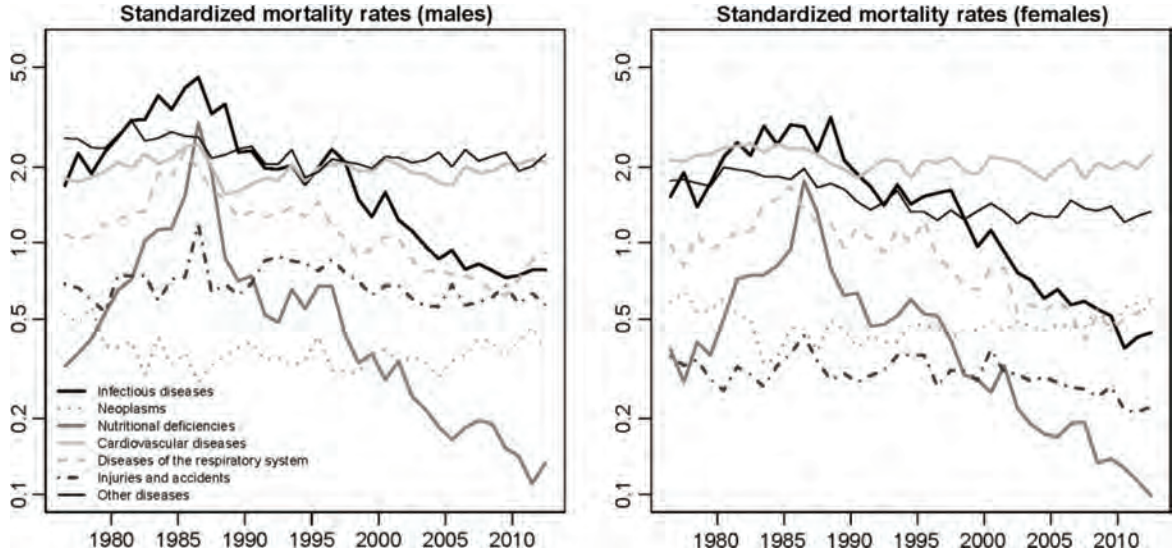

Fig. 4. Standardized mortality rates for the main categories of causes of death (using the age structure of 1993 as a standard), based on data from the BMH.

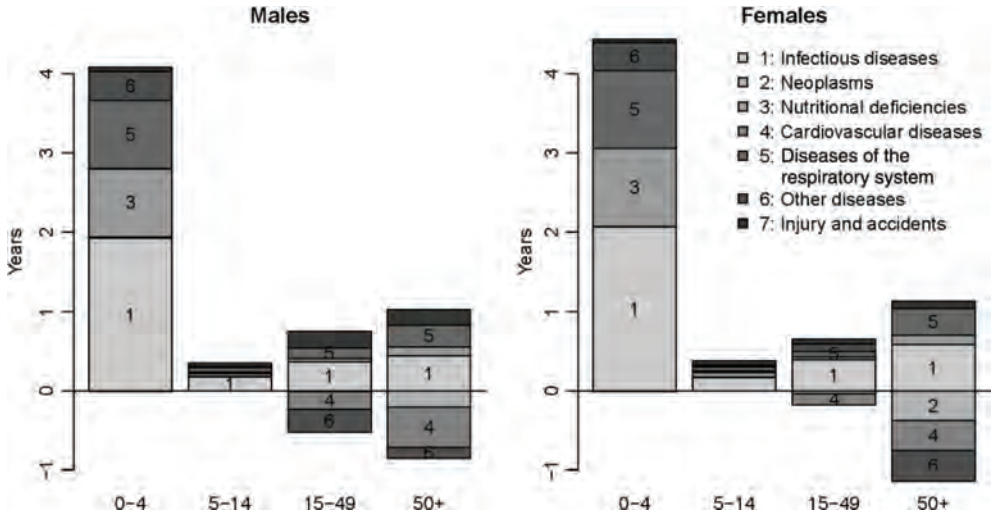

Fig. 5. Contribution of the main categories of causes of deaths to changes in life expectancy (by sex) between 1990–1994 and 2008–2012 in Antananarivo.

it is necessary that we wait for the next census to conduct a proper analysis of the completeness of death reporting through standard demographic techniques (43).

A second limitation of death registers is related to the assignment and coding of causes of deaths. One must remain particularly circumspect in interpreting cause-of-death statistics of the start of the 20th century. Syphilis is a good example of this. It was considered by the French administration to be one of the main causes of depopulation in Madagascar, with an estimated prevalence of 60–75% in adults (44). In the monthly reports, it was singled out as accounting for as much as 9.5% of the reported deaths in 1901 (excluding stillbirths). This is probably exaggerated because of the difficulty in distinguishing between syphilis and what was called *farasisa* by midwives and native doctors, a catch-all term used to describe all chronic diseases with cutaneous symptoms (16). In 1902, the shortest list of the First Revision (ICD-1) of the International Classification of Diseases started to be used for the monthly reports and syphilis was subsumed into the 'other causes'; altogether these accounted for only 3.7% of deaths.[1] These variations highlight the need for caution and indicate that the figures presented here should be considered as broad indications rather than precise estimates. Even for the most recent period, the assignment of causes of deaths in hospital records or by the physicians of the BMH, and the coding which was done for this study, might be affected by errors.

Despite these limitations, death registers yield a coherent picture of the transition that took place in Antananarivo since 1900. Substantial changes in health and disease patterns occurred only from the 1940s. This decline coincided with the start of the indoor spraying of DDT and the introduction of chloroquine; the percentage of deaths attributable to malaria declined from 20% in 1946 to 5% in 1951 in the city (21). Further health improvements were achieved between 1951 and 1976, probably thanks to the diffusion of antibiotics. The chronology of the introduction of antibiotics and their impact on mortality trends awaits further study. However, the steep decline in the proportion of deaths due to pneumonia, bronchitis, influenza and tuberculosis between the years 1900–1903 and 1976–1980 indicates that they played a significant role. As for immunizations, the impact of the EPI was first obfuscated by the major mortality crisis of the mid-1980s but it became clearly apparent in the following decades. Not one death from measles has been reported to the BMH since 2005, even though measles accounted for 8% of deaths under age 5 in the period 1976–1980 and as much as 14% during the 1985 epidemic. Likewise, tetanus, polio, diphtheria and pertussis together caused 3% of under-five deaths in 1976–1980, and not one death from these four causes has been reported since 2005.

Gains in education undoubtedly contributed to the increased use of modern medical technology and the earlier recognition and treatment of diseases. By contrast, there is enough evidence to suggest that the mortality decline was not supported by rises in standards of living or improvements in nutrition. Within the capital city, it is estimated that the per capita consumption declined by 45% between 1961 and 1995, with a decline of 34% in family expenditures allocated to food and a decline of 44% in health care spending (32). Between the DHS conducted in 1992 and 2008/9, there has been little, if any progress in the reduction of stunting (growth retardation associated with chronic malnutrition). The percentage of children aged less than 5 years with moderate and severe

[1]Prior to 1902, the diseases were probably classified according to the classification used in Paris and other French cities of more than 10,000 inhabitants starting in 1887.

Citation: Glob Health Action 2014, **7**: 23237 - http://dx.doi.org/10.3402/gha.v7.23237

stunting was 47.7% in the capital city in 1992, and 46.8% in 2008/9 (25, 28).

The experience of Antananarivo thus supports the view that changes in living standards and nutrition in developing countries were largely insufficient to explain the rapid gains in life expectancy after the World War II (45). The fact that the mortality decline was so dependent on anti-parasitic and anti-microbial medicine raises concerns in the context of increasing threads from drug-resistant infections (46, 47). It also bears implications for the future of the transition, as it means that the health care system has a crucial role to play despite persistent political turmoil and scarcity of public resources. From a theoretical perspective, the case of Antananarivo also raises a more general question, which is whether the likelihood of experiencing mortality reversals or stalls in the transition varies according to the type of factors that are driving the mortality declines (changes in income, nutrition, education or medical advances). Not only did the epidemiological transition start late in Antananarivo, it also proceeded by fits and starts, with important setbacks, particularly in the mid-1980s. In this sense, this transition bears some resemblance to the *protracted polarized model* conceptualized by Frenk et al. (48) and used recently to describe the changes occurring in Accra (8) and South Africa (9). In this model, counter-transitions can occur. Frenk et al. (48) also showed that the different stages of the epidemiological transition identified by Omran (1) are not necessarily sequential but can overlap for a considerable amount of time, resulting in the coexistence of infectious diseases and nutritional deficiencies with NCDs. In Antananarivo, it is only after 1990 that mortality rates from infectious diseases started to decline clearly below those attributable to CVDs. This relates to the notion of 'double burden of disease' experienced in many other African countries (49, 50). Another important dimension of this model, that we could not investigate here, is that transitions are thought to progress at varying speeds in the different socio-economic groups, leading to a widening of inequalities in health between various segments of the population. It is likely that there has been substantial intra-urban variation in the pace of the transition in Antananarivo and this should be further explored. More attention should also be devoted to the mortality differentials between the capital city, other urban areas and the rest of the country. Other cities have a BMH in place, and some of them are known to maintain a virtually complete registration of deaths, such as Antsirabe, the third largest city of Madagascar (51). A comparative analysis could reveal whether an 'epidemiological polarization' is occurring across regions.

The diversity within SSA makes it difficult to generalize our findings to a broader context, especially because the HIV epidemic has dramatically distorted the epidemiological profile of several countries of mainland SSA. Madagascar has succeeded in keeping its HIV prevalence among the adult population below 1%, despite high rates of sexually transmitted infections such as syphilis. Mortality rates in adults and children have evolved in tandem. Progress in adult mortality has remained limited, however, because the decline in mortality rates caused by diseases predominant in the pre-transitional stage has been partly offset by the increase in mortality due to NCDs. Few data exist on risk factors for NCDs in Antananarivo. Obesity levels remain low in the city [3% of women aged 15–49 in 2008–2009 (28)], but prevalence of hypertension (blood pressure level of 140/90 mmHg or higher), estimated at 23% in 1997 and 28% in 2009 among the adult population (52, 53), is comparable to other urban areas in Africa. According to death registers, neoplasms account for a smaller proportion of deaths (7% in 2012), but standardized mortality rates from neoplasms have been on the rise over the past two decades, from 0.42 per thousand in 1990 to 0.61 in 2012.

Finally, this overview illustrates that death registration in major cities can provide valuable insights into the changes in patterns of mortality by cause. This had been previously demonstrated in other cities such as Abidjan (1986–1992), Saint-Louis (1930–1988), and Bamako (1974–1985) (54–56), but there has been little update since the 1990s. When high coverage is maintained, the registration of deaths offers a basis for monitoring the health system performance and evaluating the impact of interventions. In Madagascar in 2006, the Ministry of Health and Family Planning introduced the *National Health Sector Strategy and Development Plan* (PDSS) that expired in 2011 and was followed by an interim short-term plan. The performance of these plans could be reviewed based on data from the death registers, and new targets could be set for specific cause-specific mortality rates, beyond the focus on child mortality and on maternal deaths. In the future, achieving further health improvements will require a more effective fight against CVDs and increased efforts to reverse the upward trends observed in neoplasms, while at the same time addressing the unfinished agenda of communicable diseases.

Conclusion

Our findings demonstrate that the epidemiological transition in Antananarivo has been protracted, with a pronounced mortality crisis in the mid-1980s followed by a very rapid decline in child mortality that was not supported by sustained economic growth or improved nutrition. Progress in adult survival have been more limited because gains in mortality due to infectious diseases and diseases of the respiratory system have been partly offset by increases in CVDs, neoplasms and other diseases. Antananarivo now clearly faces a double burden of communicable diseases and NCDs.

Main findings

- Epidemiological transition has been delayed in Antananarivo, Madagascar's capital city. Mortality rates first declined after the 1940s and then resurged again in the mid-1980s, due to a temporary re-emergence of malaria and the collapse of Madagascar's economy.
- Between 1986 and 2012, life expectancy at birth increased steadily from 54 to 69 among females and from 46 to 62 among males. Under-five mortality rates declined dramatically, despite persistent political instability and long-lasting economic crisis.
- Antananarivo now faces a double burden of disease. Cardiovascular diseases (CVDs), neoplasms and trauma account for more than 50% of deaths (compared to 9% around 1900). Infections and nutritional deficiencies account for approximately 12% of deaths, partly due to moderate and severe stunting in more than 40% of children under age 5.

Key messages for action

- Further gains in life expectancy require more effective interventions for CVDs and neoplasms, while concurrently addressing the unmet goals with regards to communicable diseases. Adult health needs require more attention, due to the limited progress in adult survival in recent years.
- Death registers maintained in major cities, though not able to compensate for the lack of a comprehensive vital registration system at the national level, could be further exploited to evaluate the performance of the health system and monitor intervention impacts in the urban population.
- Other capital cities in Sub-Saharan Africa have also set up a system of routine registration of deaths; however, the completeness of death reporting is often unknown and the data are rarely analysed. More research is required to evaluate whether such systems provide valuable cause-specific mortality statistics.

Acknowledgements

We are grateful to Dr. Brigitte Rakotosolofo Bakololao (Direction de l'Assistance sociale et de la santé publique) for providing access to death registers. Pierre Cantrelle initiated the systematic analysis of data from the BMH. The digitization of the registers was funded by INED (Institut National d'Etudes Démographiques) and the present analysis was carried out thanks to a fellowship from the Belgian National Fund for Scientific Research (FNRS). We also thank the reviewers for their suggestions and comments.

Conflict of interest and funding

The authors have not received any funding or benefits from industry or elsewhere to conduct this study.

References

1. Omran A. The epidemiologic transition: a theory of the epidemiology of population change. Milbank Mem Fund Q 1971; 49: 509–38.
2. Mackenbach J. The epidemiologic transition theory. J Epidemiol Community Health 1994; 48: 329–31.
3. Caldwell J. Population health in transition. Bull World Health Organ 2001; 79: 159–70.
4. Albala S, Vio F. Epidemiological transition in Latin America: the case of Chile. Publ Health 1995; 109: 431–42.
5. Frenk J, Bobadilla J, Lozano R. The epidemiological transition in Latin America. In: I Timæus, J Chackiel, L Ruzicka (eds.), Adult mortality in Latin America. Oxford: Clarendon Press; 1996. pp. 123–40.
6. Cook IG, Dummer TJ. Changing health in China: re-evaluating the epidemiological transition model. Health Pol 2004; 67: 329–43.
7. Phillips DR. Problems and potential of researching epidemiological transition: examples from Southeast Asia. Soc Sci Med 1991; 33: 395–404.
8. Agyei-Mensah S, de Graft Aikins A. Epidemiological transition and the double burden of disease in Accra, Ghana. J Urban Health 2010; 87: 879–97.
9. Kahn K, Garenne ML, Collinson MA, Tollman SM. Mortality trends in a new South Africa: hard to make a fresh start. Scand J Public Health Suppl 2007; 69: 26–34.
10. Kalla A. Health transition in Mauritius: characteristics and trends. Health Place 1995; 1: 227–34.
11. Riley J. The timing and pace of health transitions around the world. Popul Dev Rev 2005; 31: 741–64.
12. United Nations. World population prospects: the 2012 revision. Available from: http://esa.un.org/wpp/ [cited 6 December 2013].
13. Masquelier B, Reniers G, Pison G. Divergences in mortality trends in sub-Saharan Africa: survey evidence on the survival of children and siblings. Population Stud. DOI: 10.1080/00324728.2013.856458
14. World Bank. World Bank data. Available from: http://databank.worldbank.org/data/home.aspx [cited 6 December 2013].
15. United Nations Statistics Division. Coverage of birth and death registration. Available from: http://unstats.un.org/unsd/demographic/CRVS/CR_coverage.htm [cited 6 December 2013].
16. Paillard YG. Les recherches démographiques sur Madagascar au début de l'époque coloniale et les documents de l'AMI [Demographic research on Madagascar at the beginning of the colonial period and AMI documents]. Cahiers d'Etudes Africaines 1987; 27: 17–42.
17. Campbell G. The state and pre-colonial demographic history: the case of nineteenth-century Madagascar. J Afr Hist 1991; 32: 415–45.
18. Gaüzère B, Aubry P. Histoire des épidémies et des endémo épidémies humaines dans le sud-ouest de l'océan Indien. Médecine et Santé Tropicales 2013; 23: 145–57.
19. Blanchy S, Rakotonjanabelo A, Ranaivoson G, Rajaonarivelo Epidémiologie du paludisme sur les hautes terres malgaches depuis 1878. Cahiers d'tudes et de recherches francophones/Santé 1993; 3: 155–61.
20. Robequain C. Une capitale montagnarde en pays tropical: Tananarive. Revue de géographie alpine 1949; 37: 273–330.

Citation: Glob Health Action 2014, **7**: 23237 - http://dx.doi.org/10.3402/gha.v7.23237

21. Mercier S, Razafindrakoto JB. Bilan de trois années de campagnes de désinsectisation domestique à Tananarive. Bulletin de la Société de pathologie exotique et de ses filiales 1953; 46: 463–73.
22. Waltisperger D, Meslé F. Crise économique et mortalité: Le cas d'Antananarivo 1976–2000 [Economic crisis and mortality: the case of Antananarivo, 1976–2000]. Population (French Edition) 2005; 60: 243–75.
23. Mouchet J, Laventure S, Blanchy S, Fioramonti R, Rakotonjanabelo A, Rabarison P, et al. La reconquête des Hautes Terres de Madagascar par le paludisme [The reconquest of the Madagascar highlands by malaria]. Bull Soc Pathol Exot 1997; 90: 162–8.
24. Domarle O, Razakandrainibe R, Rakotomalala E, Jolivet L, Randremanana RV, Rakotomanana F, et al. Seroprevalence of malaria in inhabitants of the urban zone of Antananarivo, Madagascar. Malar J 2006; 5: 106.
25. CNRE (1994). Enquête Nationale Démographique et Sanitaire 1992, Madagascar. Claverton: Centre National de Recherches sur l'Environnement (CNRE) et Demographic and Health Surveys, Macro International.
26. Direction de la Démographie et des Statistiques Sociales, Institut National de la Statistique (INSTAT) [Madagascar] et Macro International (1998). Enquête Démographique et de Santé, Madagascar 1997. Claverton, UK: INSTAT and Macro International.
27. INSTAT (2005). Enquête démographique et de santé de Madagascar 2003–2004. Claverton: ORC Macro.
28. Macro International (2010). Enquête Démographique et de Santé de Madagascar 2008–2009. Claverton, Maryland: INSTAT and ICF Macro.
29. Gould W. African mortality and the new urban penalty. Health Place 1998; 4: 171–81.
30. Bocquier P, Madise N, Zulu E. Is there an urban advantage in child survival in sub-Saharan Africa? Evidence from 18 countries in the 1990s. Demography 2011; 48: 531–58.
31. Campbell G. An industrial experiment in pre-colonial Africa: the case of imperial Madagascar, 1825–1861. J South Afr Stud 1991; 17: 525–59.
32. Ravelosoa R, Roubaud F. La dynamique de la consommation des ménages dans l'agglomération d'Antananarivo, 1965–1995 (Madagascar) [The dynamics of household consumption in the urban agglomeration of Antananarivo, 1965–1995 (Madagascar)]. Autrepart 1998; 7: 63–87.
33. Journal Officiel de Madagascar et dépendances, Tananarive: Imprimerie Nationale 1896–1958. Available from: http://gallica.bnf.fr/ark:/12148/cb34425284n/date [cited 6 December 2013].
34. Waltisperger D, Canterelle P, Ralijaona O. La mortalité à Antananarivo de 1984 à 1995. Paris: Les documents du Ceped n°7; 1998.
35. Gendreau F. Centres urbains. In: Délégation générale à la recherche scientifique et technique, ed. Afrique noire, Madagascar, Comores: Démographie Comparée. Tome 1, Volume 1; 1967.
36. Chevalier L. Madagascar: populations et ressources. Paris: Presses universitaires de France; 1952.
37. Disaine B. La transition démographique à Madagascar. In: Organisation for Economic Co-operation and Development, Development Centre, Population Unit, eds. La transition démographique en Afrique tropicale: comptes-rendus de la réunion d'un groupe d'experts, Paris, 17–19 November 1970; 1971, pp. 55–64.
38. United Nations. Demographic yearbook 1997 – historical supplement. Department of International Economic and Social Affairs, United Nations; 1997. Available from: http://unstats.un.org/unsd/demographic/products/dyb/dybhist.htm [cited 6 December 2013].
39. Andreev E, Shkolnikov V, Begun A. Algorithm for decomposition of differences between aggregate demographic measures and its application to life expectancies, healthy life expectancies, parity-progression ratios and total fertility rates. Demogr Res 2002; 7: 499–522.
40. Courbage Y, Fargues P. A method for deriving mortality estimates from incomplete vital statistics. Popul Stud 1979; 33: 165–80.
41. INSRE (1967). Enquête démographique, Madagascar 1966. Antananarivo, Madagascar: Institut National de la Statistique et de la Recherche Economique.
42. Lacombe B. L'état civil malgache et son exploitation démographique. Cahiers Orstom 1973; 10: 344–60.
43. Moultrie T, Dorrington R, Hill A, Hill K, Timaeus I, Zaba B. Tools for demographic estimation. Paris: International Union for the Scientific Study of Population; 2013.
44. Gallieni J. Instructions relatives aux mesures à prendre pour favoriser l'accroissement de la population en Emyrne [Instructions on measures to promote the growth of the population in Emyrne]. Journal officiel de Madagascar et dépendances 1898; 265: 2017–21.
45. Soares R. On the determinants of mortality reductions in the developing world. Popul Dev Rev 2007; 33: 247–87.
46. Randrianirina F, Vaillant L, Ramarokoto CE, Rakotoarijaona A, Andriamanarivo ML, Razafimahandry HC, et al. Antimicrobial resistance in pathogens causing nosocomial infections in surgery and intensive care units of two hospitals in Antananarivo, Madagascar. J Infect Dev Ctries 2010; 4: 74–82.
47. Andriantsoanirina V, Ratsimbasoa A, Bouchier C, Jahevitra M, Rabearimanana S, Radrianjafy R, et al. *Plasmodium falciparum* drug resistance in Madagascar: facing the spread of unusual *pfdhfr* and *pfmdr-1* haplotypes and the decrease of dihydroartemisinin susceptibility. Antimicrob Agents Chemother 2009; 53: 4588–97.
48. Frenk J, Bobadilla J, Sepuúlveda J, Cervantes M. Health transition in middle-income countries: new challenges for health care. Health Pol Plann 1989; 4: 29–39.
49. BeLue R, Okoror T, Iwelunmor J, Taylor K, Degboe A, Agyemang C, et al. An overview of cardiovascular risk factor burden in sub-Saharan African countries: a socio-cultural perspective. Global Health 2009; 5: 10.
50. Dalal S, Beunza J, Volmink J, Adebamowo C, Bajunirwe F, Njelekela M, et al. Non-communicable diseases in sub-Saharan Africa: what we know now. Int J Epidemiol 2011; 40: 885–901.
51. Andriamboahanguy B. Confrontation de l'état civil avec un recensement: un exemple dans la préfecture d'Antsirabe. Cahiers ORSTOM. Sér Sci Hum 1973; 10: 361–70.
52. Mauny F, Viel J, Roubaux F, Ratsimandresy R, Sellin B. Blood pressure, body mass index and socio-economic status in the urban population of Antananarivo (Madagascar). Ann Trop Med Parasitol 2003; 97: 645–54.
53. Rabarijaona L, Rakotomalala D, Rakotonirina E, Rakotoarimanana S, Randrianasolo O. Prévalence et sévérité de l'hypertension artérielle de l'adulte en milieu urbain à Antananarivo. Revue d'Anesthésie-Réanimation et de Médecine d'Urgence 2009; 1(4): 24–7.
54. Garenne M, Madison M, Tarantola D, Zanou B, Aka J, Dogore R. Conséquences démographiques du sida en Abidjan: 1986–1992. Les études du CEPED n°10, Paris: CEPED; 1995.
55. Diop I. Etude de la mortalité à Saint-Louis du Sénégal à partir des données d'état civil. PhD thesis, Institut de Démographie de Paris, 1990.
56. Fargues P, Nassour O. Douze ans de mortalité urbaine au Sahel: Niveaux, tendances, saisons et causes de mortalité à Bamako 1974–1985. Travaux et documents n°123, Paris: Cahier de l'INED; 1988.

Global Health Action

PART III

Migration and the epidemiological transition: insights from the Agincourt sub-district of northeast South Africa

Mark A. Collinson[1,2,3]*, Michael J. White[1,4], Philippe Bocquier[1,5], Stephen T. McGarvey[6], Sulaimon A. Afolabi[1], Samuel J. Clark[1,3,7,8], Kathleen Kahn[1,2,3] and Stephen M. Tollman[1,2,3]

[1]MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; [2]Centre for Global Health Research, Division of Epidemiology and Global Health, Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden; [3]INDEPTH Network, Accra, Ghana; [4]Department of Sociology, Population Studies and Training Center, Brown University, Providence, RI, USA; [5]Centre de recherche en démographie et sociétés, Université Catholique de Louvain, Louvain-la-Neuve, Belgium; [6]Brown University School of Public Health, International Health Institute, Brown University, Providence, RI, USA; [7]Department of Sociology, University of Washington, Seattle, USA; [8]Institute of Behavioral Science (IBS), University of Colorado at Boulder, Boulder, CO, USA

Background: Migration and urbanization are central to sustainable development and health, but data on temporal trends in defined populations are scarce. Healthy men and women migrate because opportunities for employment and betterment are not equally distributed geographically. The disruption can result in unhealthy exposures and environments and income returns for the origin household.

Objectives: The objectives of the paper are to describe the patterns, levels, and trends of temporary migration in rural northeast South Africa; the mortality trends by cause category over the period 2000–2011; and the associations between temporary migration and mortality by broad cause of death categories.

Method: Longitudinal, Agincourt Health and Demographic Surveillance System data are used in a continuous, survival time, competing-risk model.

Findings: In rural, northeast South Africa, temporary migration, which involves migrants relocating mainly for work purposes and remaining linked to the rural household, is more important than age and sex in explaining variations in mortality, whatever the cause. In this setting, the changing relationship between temporary migration and communicable disease mortality is primarily affected by reduced exposure of the migrant to unhealthy conditions. The study suggests that the changing relationship between temporary migration and non-communicable disease mortality is mainly affected by increased livelihood benefits of longer duration migration.

Conclusion: Since temporary migration is not associated with communicable diseases only, public health policies should account for population mobility whatever the targeted health risk. There is a need to strengthen the rural health care system, because migrants tend to return to the rural households when they need health care.

Keywords: *migration*; *temporary migration*; *mortality*; *epidemiological transition*; *South Africa*; *Agincourt*; *health and demographic surveillance*

Responsible Editors: Nawi Ng, Umeå University, Sweden; Barthélémy Kuate Defo, University of Montreal, Canada.

*Correspondence to: Mark A. Collinson, MRC/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt), Tintswalo Hospital, P.O. Box 2, Acornhoek, 1360, South Africa, Email: mark.collinson@wits.ac.za

This paper is part of the Special Issue: *Epidemiological Transitions – Beyond Omran's Theory*. More papers from this issue can be found at http://www.globalhealthaction.net

To access the supplementary material for this article, please see Supplementary files under Article Tools online

Received: 5 December 2013; Revised: 13 March 2014; Accepted: 28 March 2014; Published: 15 May 2014

Migration and urbanization are strongly implicated in the intertwined demographic and health transitions. These aspects of population distribution are associated with (even help drive) the demographic transition. At the same time the geographic redistribution of persons, chiefly from rural to urban areas,

Citation: Glob Health Action 2014, **7**: 23514 - http://dx.doi.org/10.3402/gha.v7.23514

exposes them to new health regimes, and the stress of relocation itself may introduce its own health consequences. Furthermore, in a mobile world, the continued circulation of persons between origin and destination may serve to expedite the spread of disease. At the same time our understanding of the interconnection between population redistribution and health dynamics is poorly understood. Analysts and policy-makers often know only the broad sweep of net changes across gross geography; fine-grained, longitudinal information about movement – both with respect to time and to place – is often unavailable.

In South Africa, circular labor migration was a cornerstone of apartheid policies and accompanied national economic and industrial growth, leaving legacies that still shape contemporary opportunities for rural South Africans (1–3). Mineral discoveries of the late 19th century led to rapid development of the mining industry and parallel industrialization. For political and economic reasons, a 'Bantustan' system was developed which restructured the settlement patterns and livelihood strategies of the native African population to provide necessary labor, while forcing unemployed family members to remain in densely settled, rural, and peri-urban areas (2, 4). Thus, circular labor migration played a key role in the success of the national economy, while placing a heavy burden on rural households (5, 6). Nowadays, temporary labor migration remains a prime strategy used by rural households to kick back against poverty (2, 7). Gold mining has been a key employment sector, and recent investigations have revealed a significant epidemic of silicosis in former mine workers (8). Other health impacts of the migrant labor system on migrants and the households left behind have been the largest HIV/AIDS epidemic in the world (9–11) and linked to this an intractable tuberculosis (TB) epidemic (12).

Our use of the Agincourt Health and Demographic Surveillance System in the rural northeast of South Africa, located in a former 'Bantustan' area, help us better understand the foundations of the relationship between migration – and urbanization, since the migration is often to urban areas – and mortality. Our rural district is economically dependent on opportunities in the major metropolitan areas of Johannesburg and Pretoria, and mines close to these cities, about 500 km away. Our approach enables us to take a detailed look at a small, well-defined population, where an exhaustive population register has kept a record of residents, migratory behavior, and mortality, with the sub-district boundary acting as a migration-defining boundary (13).

In this paper, we examine several key patterns of demographic dynamics with respect to this rural sub-district population, including the patterns, levels, and trends in temporary migration linking rural and urban areas, as well as the mortality trends by sex for the rural sub-district in the 12-year period, 2000–2011. We further provide regression-based analyses of the relationship between temporary migration and mortality in this population.

Literature review

The epidemiological and demographic transitions

Two paradigmatic transitions have offered considerable insight into change within populations over time: the epidemiological transition and the demographic transition. Both are the subject of extensive writing, which includes critiques of these as contemporary paradigms. We here provide a brief exposition of these paradigms (given widespread familiarity) and we concentrate our discussion on the connection between the two. Note that while the 'epidemiological transition' is the phrase that has long held sway, contemporary writing also makes reference to a broader 'health transition' (14, 15). We consider the epidemiological transition (with its focus on causes of morbidity and mortality) to be embedded within a health transition (which subsumes the former and also considers risk factors, and other characteristics). Recent writing on the global burden of disease (GBD) has expanded from concerns about mortality to risk factors related to morbidity and loss of healthy years of life (16, 17). We regard the epidemiological transition in its conventional sense, generally the transition from the acute communicable diseases (CDs) of poverty to the chronic non-communicable diseases (NCDs) of high-income societies. The classic exposition of this transition is due to Omran (18), although there have been numerous repositionings, extensions, and analyses both within and across low- and middle-income countries (19–21).

The second key transition of concern here is the demographic transition. Like the epidemiological transition, the demographic transition is paradigmatic and has experienced likewise a number of repositionings, and extensions (22–24). As stated most simply the demographic transition is a movement from a regime of high mortality and high fertility to one of low fertility and low mortality. The epidemiological transition is embedded within the demographic transition through mortality reduction. Population growth rates are low or near-nil at the beginning and end of the transition, while population growth can be substantial (as, conventionally, mortality decline precedes fertility decline) in the interim phase. While the basic demographic transition is a highly structured paradigm, variations in the relationship between mortality and fertility trends across time and space have important implications for age structure and economic development, and it is linked to a number of other phenomena, including urbanization (23).

Public health discussion once presumed an inexorable parallel between economic development and health transitions, but changes in both technology and disease regimes have challenged this paradigm. As populations move

from the diseases of poverty to those of wealth, this rise in NCDs will have profound impacts on health care systems of low- and middle-income countries (LMIC) (25, 26). Many African populations already manifest risk factors for NCDs, despite remaining classified well within the World Bank low-income category. NCDs have already reached to a level of about one-quarter of all mortality in the WHO Africa Region (27). Of particular note is the need to understand the heterogeneity in the dual burden of both CDs and NCDs, a public health problem of rising concern (28–32). Thus, the association between development indicators and the health transition is different at the national and local levels, and we cannot consider NCDs to be solely conditions of national affluence (21). The disjuncture between the conventional development and population health paradigm and the newly developing and heterogeneous health transition is particularly acute and consequential for Africa, where the burdens of poverty, HIV, and other diseases are sharply felt.

The recently updated GBD 2010 study (Global Burden of Disease) points quite clearly to the dual burden currently experienced in developing settings, linking these to leading risk factors. For sub-Saharan Africa, the GBD 2010 study finds high (and thus adverse impact) rankings for high blood pressure and tobacco use, while also finding high rankings for suboptimal breastfeeding, inadequate sanitation, and iron deficiency. More specifically and pertinently for our sub-region of interest, the GBD team writes 'In 2010, alcohol use was the leading risk factor in southern sub-Saharan Africa, followed by high blood pressure and high body-mass index . . .'(17).

The contemporary pace of the epidemiological (or health) transition argues further for consideration of its relationship with migration, urbanization, and the overall demographic transition. The demographic transition in Sweden occurred on a time scale of over a century. In more recently developing parts of the world the demographic transition has proceeded more swiftly. In the Southern African region as a whole, UN estimates indicate that the overall growth rate declined from 2.41% annually in the 1950s to 0.9% annually in the 2000–2010 interval. Even more pointedly, the demographic transition (as seen through the key driver of fertility) has been evident in the Agincourt study population, where the total fertility rate (TFR) averaged 6.0 in 1979 has declined to 2.3 by 2004 (33). Despite the rapid overall transition, significant differentials exist by African geographic sub-region and urbanization level (34).

Long-standing CDs still matter much. For contemporary African children, 75% of the burden of disease is accounted for by malaria, diarrheal disease, respiratory infections, and other CDs (17, 35). However, the dual burden of disease is already manifest: sub-Saharan Africa exhibits death rates from cardiovascular disease (CVD) parallel to LMIC South Asia or Central Asia (35). This phenomenon can be seen in all contemporary developing settings but is especially pertinent to sub-Saharan Africa (30, 36, 37).

Migration, urbanization, and health

The majority of the world's population lives in urban areas now. Even in sub-Saharan Africa, some 37% of the population resides in urban areas (38). While the urbanization associated with economic growth is likely to raise average levels of well-being, there has always been some concern about the potentially deleterious effects of urbanization on health (39, 40). This concern takes several forms, but central is the shift in exposure regime. Whereas in rural areas certain endemic vector-borne diseases drive the burden of disease, in urban areas these are reduced in incidence (not eliminated!) but are supplemented with concerns that accompany issues of sanitation in dense settings, exposure to air pollution, and the like. Added to risk in urban settings are the diseases of more sedentary lifestyles and purchased foods concomitant with rising incomes and CVDs and the low health literacy among citizens and clinicians about NCDs such as hypertension and type 2 diabetes. For instance, in a broad review article on CVD (41), Yusuf and co-authors strongly implicate urbanization in developing societies. With urbanization (or migration to Western environments), there is a marked increase in consumption of energy rich foods, a decrease in energy expenditure (through less physical activity) and a loss of the traditional social support. While these behavioral and health trends hold broadly – and while they clearly implicate migration and urbanization – the timing and features of the relationship are less clearly identified. Thus, both individual level exposures and health system unpreparedness in some urban neighborhoods can serve to exacerbate the rise of NCDs. Although, overcrowded conditions that accompany urbanization (and particularly urban poverty) may increase risk of infectious diseases, other aspects of resettlement to urban settings may increase the risk of exposure to environmental hazards, in turn linked to the potential for respiratory disease and diarrhea (42).

In addition to its role in urbanization, migration itself is further linked to health. Migration can be stressful, just as is the case with many major life transitions. Migration is also implicated in the transmission of human disease, and infected individuals carry disease from one location to another. Nowhere is this more in evidence and of concern that with the spread of HIV in southern Africa, a phenomenon very relevant to our analysis here (11, 43). Even though migration may disrupt or even sever residential patterns, often the networks, and sharing of resources – most typically in the form of remittances – persist. These flows of monetary and non-monetary resources are also quite relevant to our population. Although some public health writing on migration and urbanization discusses

Citation: Glob Health Action 2014, **7**: 23514 - http://dx.doi.org/10.3402/gha.v7.23514

the severing of ties and loss of social support that may accompany urbanward migration (41, 42), studies within the migration field often take note of the variety of strategies that migrants undertake to adapt in the destination and/or remain tied to the origin community through monetary remittances or other forms of communication (44, 45). At the same time, the urban setting may provide access to alternative social networks or more specific health services that could beneficial for health (46, 47). Thus, the net impact of migration could be positive or negative, and empirical analysis, such as we conduct in this paper, is necessary to sort out the balance.

The contribution of migration to urban growth is low in Africa compared to other developing regions, in part due to the highly circular nature of rural–urban migration on the continent (48). This means that there is highly prevalent back-and-forth mobility of young adults freeing themselves from constraints and poor opportunities in rural areas and joining the unpredictable urban setting in search of better livelihood opportunities (40). For many urban centers in sub-Saharan Africa, there is evidence of increasing levels of circular migration, which has reduced the contribution of in-migration to urban growth (48). As our discussion below indicates, we observe a population with a high degree of circulation.

Conceptual framework

The preceding literature review incites us to consider circular, that is, temporary rural–urban, migration in a different way than permanent migration to urban areas. Diagram 1 shows that although the two types of migrations are motivated by the (lack of) livelihoods opportunities, the latter leads to a permanent, often definitive change of environment, while the former exposes to health risks at both origin and destination (hence boxes placed on the border of rural and urban areas). Despite their dual residence temporary migrants' deaths usually happen in rural areas because these migrants are attached to rural households. In other words, the burden of disease does not reflect the exposure of temporary migrants, which has implications for policies and programs. Places which expose migrants to higher risks are not necessarily places that support most of the consequences on health.

The diagram is not meant to depict all possible moves (including in-migration and return migration) and corresponding health risks and the paper does not attempt to compare health outcomes of in- and out-migrants with non-migrants or temporary migrants. It will thus not address issues of selection by migration. Consequently, the diagram is meant to pinpoint the peculiar health situation of temporary migrants sitting astride urban and rural areas and depict a typical situation when migrants cross a rural–urban border. This can be extended to more complex situations encountered in South Africa where migrants go and work in (urban) mining locations or industrial farms. 'Urban areas' in the diagram could also be named 'labor markets', and 'rural areas' be named 'subsistence economies'.

The present empirical study focuses on the lower part of the diagram, that is, on rural areas. Diagram 2 is an empirical translation of diagram 1 and summarizes the variables available for analysis in Agincourt Health and Demographic Surveillance System (HDSS) that will serve to test hypotheses on the migration–mortality relationship. Temporary migration exposure is our main independent variable. The evolution of livelihood opportunities (including changing health system context) is captured through period effect and its interaction with migration exposure. The control variables are the demographic characteristics (age and sex). While CDs are usually associated

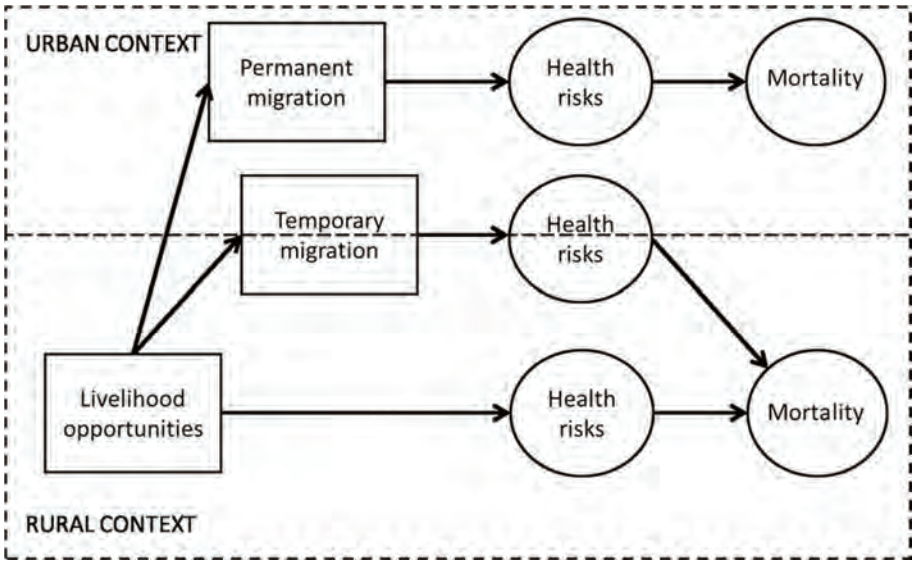

Diagram 1. Migration–health conceptual diagram.

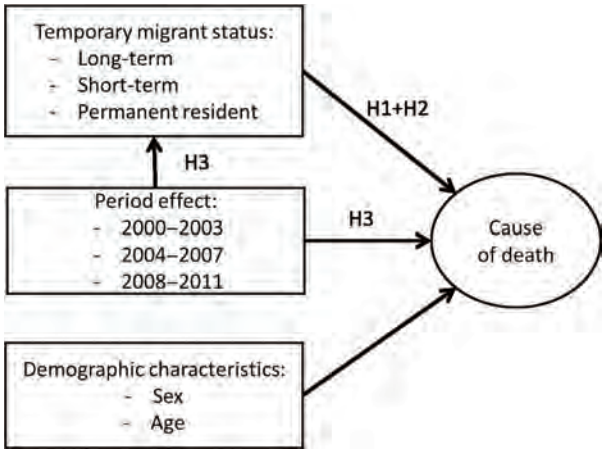

Diagram 2. Rural migration–health empirical model.

with younger adults and NCDs with older adults, the disease prevalence by age is by no means exclusively dichotomous and mortality from NCDs occurs at increasingly younger ages. To accommodate this we used age as the temporal dimension in the regressions, which creates an age standardization throughout the analysis.

Diagram 2 also indicates which causal relationships the following hypotheses are testing:

H1 Temporary migration, net of demographic characteristics, predicts higher rates of overall mortality as compared to permanent residents in rural areas. Longer exposure to health risks in destination areas should lead to higher mortality rates.

H2 The second hypothesis is embedded in the first: Temporary migration is less strongly associated with the relative risk of death from NCD. CDs should be the highest health risks that circular migrants are subjected to.

H3 Both overall mortality and the impact of temporary migration on mortality have declined over time, as temporary migrants and permanent residents have experienced better livelihood opportunities, including an increase in health outreach, especially antiretroviral treatment (ART) roll-out.

Methods

Health and demographic surveillance system

The Agincourt HDSS was established in 1992, and it was designed to capture the health, sociodemographic, and other information of all the residents of Agincourt, a sub-district of the Bushbuckridge municipality situated in rural north-eastern part of South Africa. Agincourt encompasses a geographical area of about 420 km^2 where a pre-dominantly Shangaan-speaking black South African population reside. The surveillance started with a baseline census in 1992 and was followed by annual updates of the initial records in the subsequent years. The initial population was about 70,000 individuals residing in 21 villages. In 2007, the study population was expanded by increasing the number of villages under surveillance to 27, with an expanded population of about 90,000 (49). This paper uses data collected between 2000 and 2011. In addition to collecting information about births, death, migrations, unions, and household membership, specialized census modules are employed to capture different socioeconomic data in each annual census round. Modules are repeated with a certain period to provide detail about the dynamics of key phenomena that influence health and demographic patterns (49, 50).

Migration methods

Migration is recorded in one of two ways in the HDSS depending on the report given by the household respondent. If a person enters or leaves a household with a permanent intention, it is recorded as a permanent in- or out-migration event with details captured, such as date of move, origin or destination place, and reason for the move. Permanent residents include all non-migrant residents who did not move over the 2000–2011 period as well as in-migrants from the time they start residing in the site (being left-censored) and out-migrants until they move out of the site (being right-censored). Individuals may be both in-migrating and out-migrating over the period. It is important to note that in this study, we do not attempt to analyze the differential mortality outcomes of the subcategories of permanent residents. Neither do we try to analyze the effect of in- or out-migration as an event since that would require the population at risk at both ends of the migration streams. Rather, we analyze outcomes of permanent residents against those of temporary migrants exposed to both origin and destination environments.

A person is recorded as a temporary migrant if the respondent says that they exited the household with a temporary intention and now resides elsewhere in a second household. A temporary migrant continues to be a household member and is not removed from the household roster. Temporary migration status is updated annually with reference to the year preceding the household interview by asking on how many months a person was physically present (on aggregate) out of the previous 12, with 6 months or less classified as temporary migrant. Figures 1 and 2 show data obtained from this temporary migration status question. In the migration and mortality analyses, two categories of temporary migration duration are used, viz. short duration, when the completion of a temporary migrant cycle is within 3 years, and long duration, with completion of a cycle in longer than 3 years. Three years duration is an arbitrary cut-off used to discriminate between circumstantial or exploratory temporary migration (short duration) and an entrenched pattern (long duration), with the assumption that a 3 years or longer duration requires a relatively stable incentive such as a

Citation: Glob Health Action 2014, **7**: 23514 - http://dx.doi.org/10.3402/gha.v7.23514

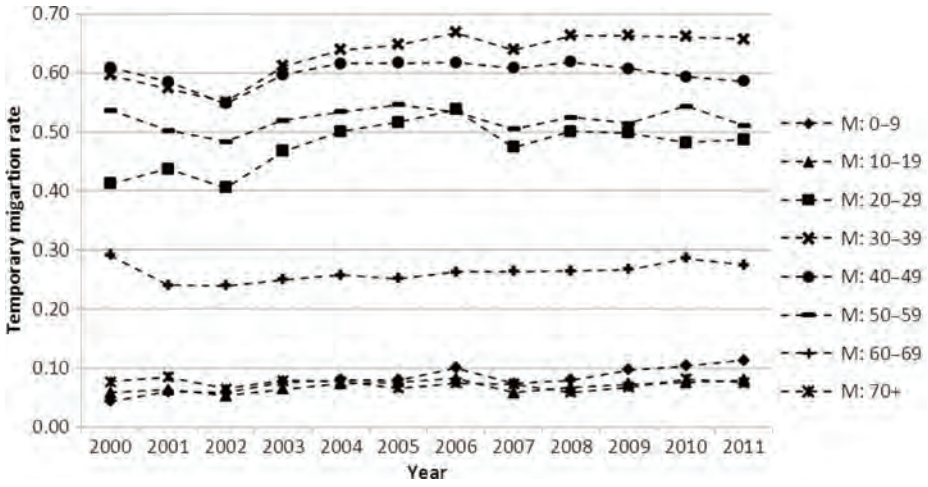

Fig. 1. Male temporary migration rate trend by calendar year and 10-year age groups, Agincourt, 2000–2011.

permanent employment contract, which can have different mortality consequences.

In addition to the annual temporary migration status update, there is a more comprehensive temporary migration module run every 5 years that was implemented in 2002, 2007, and 2012. The aim is to acquire more detail on the temporary migration, such as destination, reason for move, and the nature of links between the origin and destination households, including remittance transfers (49). Entry into this module requires that the definition of temporary migration status be met. This implies that from the perspective of the respondent, the intention of the migrant is to remain a member of the rural household while away and that the person was absent for at least 6 months out of the 12 months preceding the interview. The inferred intention of the migrant in the definition may seem unstable from a scientific perspective, because intentions and behaviors can change or be misreported, but in this context there has been much experience of labor migration and the notion of a non-resident household member is well understood. If a temporary migrant is deemed by a respondent to no longer have the intention of remaining a household member, then the fieldworker can permanently out-migrate the person and they are removed from the household roster. Data from the periodic temporary migration modules are used in Table 1 to highlight the link between the temporary migrant and rural household and also to validate the definition.

Temporary migration as a repeatable exposure is the key independent variable used in the analysis. To understand the relative prevalence of temporary migration compared to permanent migration, the Crude Migration Rates were computed in 2006, a year in the middle of the observation period. For females, the permanent in-migration rate was 26/1,000 and the permanent out-migration rate was 30/1,000, while the temporary migration rate was 177/1,000. For males, in 2006, the permanent in-migration rate was 17/1,000, the permanent out-migration rate 22/1,000 and the temporary migration rate was 322/1,000.

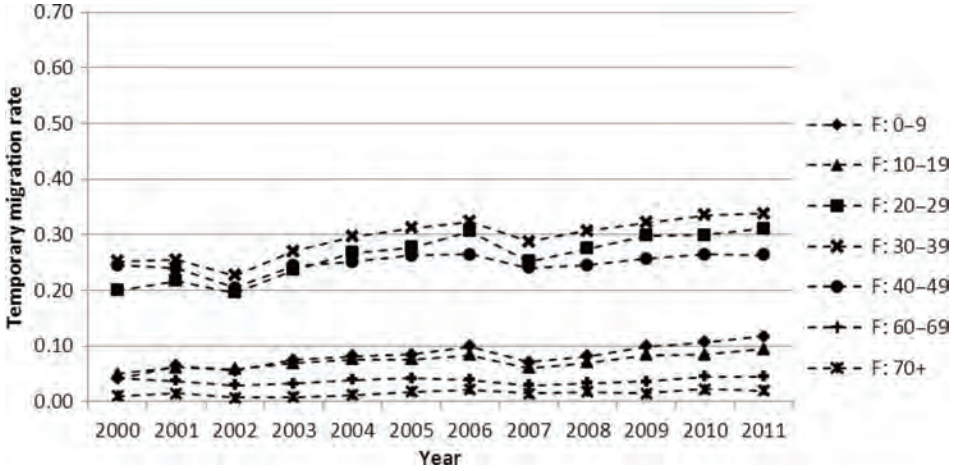

Fig. 2. Female temporary migration rate trend by calendar year and 10-year age groups, Agincourt, 2000–2011.

Table 1. By sex and observation year: percent of temporary migrants employed and percent of employed temporary migrants that remitted cash or another item back to the origin household

	2002		2007		2012		Total	
% of temporary migrants employed	*n*	%	*n*	%	*n*	%	*n*	%
Male	5,529	75	7,042	78	7,106	73	19,677	75
Female	2,021	53	2,532	52	2,745	49	7,298	51
Test sex difference		M > F		M > F		M > F		M > F
p		*p* = 0.000		*p* = 0.000		*p* = 0.000		*p* = 0.000
	2002		2007		2012		Total	
% of employed migrants that remitted	*n*	%	*n*	%	*n*	%	*n*	%
Male	3,820	69	4,437	64	4,815	68	13,072	67
Female	1,433	71	1,653	66	1,994	73	5,080	70
Test sex difference		M < F		M < F		M < F		M < F
p		*p* = 0.065		*p* = 0.01		*p* = 0.000		*p* = 0.000

Mortality methods

Each death that occurs in the intercensal period is recorded during the annual vital events update. A second interview for a 'verbal autopsy' is conducted on each death by a trained lay fieldworker to establish the most probable cause of death. The interview is conducted with the closest caregiver to the deceased to establish the most prominent signs and symptoms occurring prior to death. These data are captured onto a database system and transformed into the input format for the InterVA4 analysis software that uses an algorithmic approach, calibrated by knowledge of local disease patterns, to establish the main cause of death, as well as immediate and contributing causes (51–54).

The broad cause of death categories used are CD, NCD, external causes, and undetermined causes. These categories are made up of constituent causes of death and the categorization is shown in a Supplementary file, where constituent causes of death are ordered by rank of frequency. For CDs, the most important three constituent causes were HIV/AIDS, pulmonary TB, and acute respiratory infections; for NCDs: chronic obstructive pulmonary disease, cardiac disease, and stroke; and for external causes of death: assault, motor vehicle accident, and intentional self-harm. The mortality trends in Figures 3 to 5 and the regression results in Tables 2–5 are from the core HDSS residence files, with residence status data merged in and residence criteria of 6 months in the origin

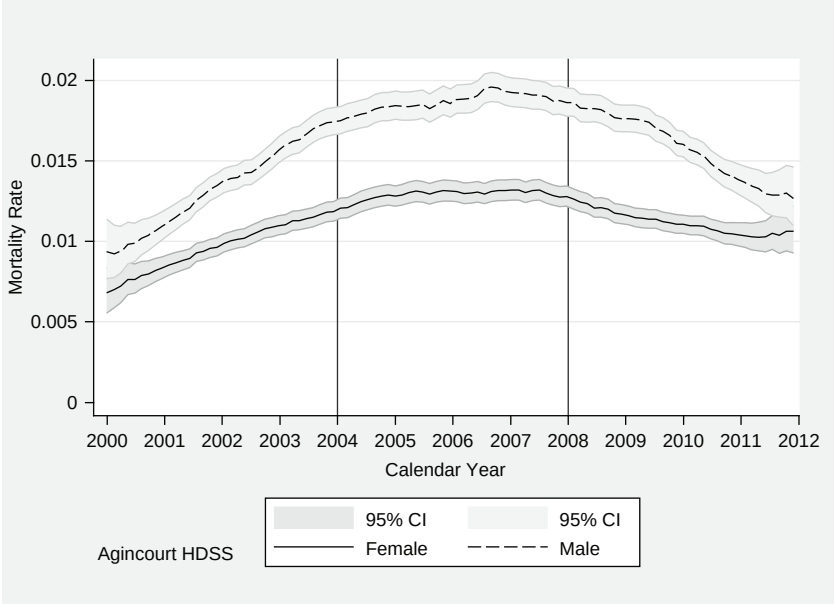

Fig. 3. All-cause mortality rate by calendar year for males and females, Agincourt HDSS.

Citation: Glob Health Action 2014, **7**: 23514 - http://dx.doi.org/10.3402/gha.v7.23514

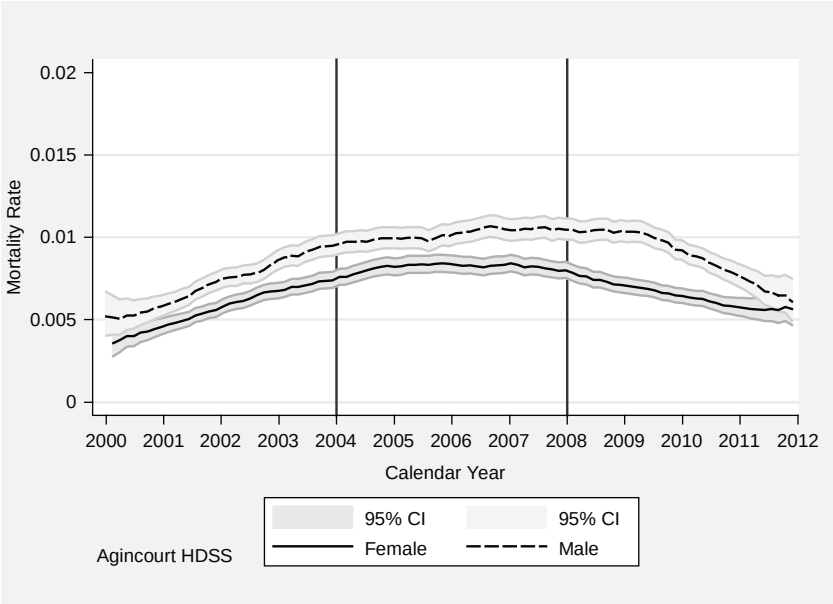

Fig. 4. Communicable disease mortality rate by calendar year for males and females, Agincourt HDSS.

household required to be considered part of the household. Mortality rates by cause of death are provided in a Supplementary file, giving age and sex profiles in the three analytic periods described later.

Migration and mortality analysis

We examine migration and mortality using a continuous, survival time, competing-risk model which estimates the sub-hazard ratios of migration exposures as related to mortality outcomes, comparing migrant with non-migrant categories, after controlling for the age and sex structure of the population. Given that dates are recorded to the day, a continuous-time model is preferred to discrete-time model, which is less precise because of ties, that is, when the order of censoring and event is not known within a time interval. Also, a continuous-time model makes it easier to handle large datasets since dates are recorded only when there is a change in value. To prepare the data, we constructed a residency file, whereby each individual has a sequential record kept of each demographic event that occurred to them. Events include in- or out-migrations, the date of becoming a temporary migrant, and the date of ending a period as a temporary migrant, as well as births or deaths, with deaths classified by probable main cause.

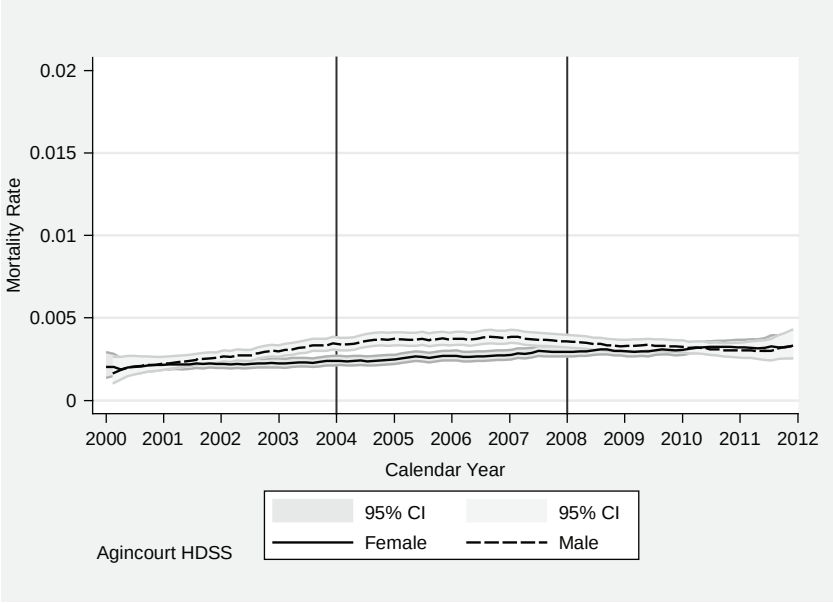

Fig. 5. Non-communicable disease mortality rate by calendar year for males and females, Agincourt HDSS.

Table 2. Migration factors associated with death from communicable disease

	Hazard ratio (95% Confidence intervals); death by communicable disease		
Variables	2000–2003	2004–2007	2008–2011
Sex: male vs. female	1.65*** (1.46–1.87)	1.74*** (1.58–1.93)	2.27*** (2.03–2.54)
Short duration temporary migrant vs. not	2.53*** (1.87–3.41)	1.15 (0.71–1.87)	1.30 (0.73–2.31)
Long duration temporary migrant vs. not	2.70*** (2.33–3.14)	1.65*** (1.48–1.85)	0.78*** (0.68–0.88)
Observations	52,680	52,543	55,847
Wald Chi-square	330.7	245.3	214.7
Log likelihood	−7,719	−11,836	−10,273
Subjects	48,661	49,573	53,271
Failures	1,101	1,660	1,436

***$p < 0.01$.
Comparing temporary migrants with permanent residents, controlling for age and sex.

Individuals were excluded from the regressions if they spent less than 6 months per year in the study population. The biographical files were censored at the end of the observation period on January 1, 2012. In the regression models, age is the temporal dimension and sex is included as a co-variate, so all the models control for age and sex. Migration is differentiated by short- and long-term duration. We did not include other controls because the analyses are already complex.

Competing risk models are used to predict mortality by cause of death categories and explore the extent to which short and long duration temporary migrations are linked to mortality by cause category. The Fine and Gray competing risk model is preferred to the usual Cox model, because it does not make the assumption of independent causes of death (55). Models are run for different cause of death categories in a competing risk model so that each cause of death is treated as sub-category of overall mortality and not as ordinary right-censoring for other causes. Models are also run for each of the three periods used in the mortality models: 2000–2003, 2004–2007, and 2008–2011. The choice of these periods reflects the overall non-linear mortality trends, while preserving enough statistical power for low mortality risk when conducting separate regression analyses by period. In the first period, the mortality trend is increasing; in the middle period the trend flattens; and in the third period, the mortality trend is declining. To account for possibly non-constant trends in the relation between migration status and mortality, we preferred to run separate models for each period rather than to introduce interaction terms between migration status and period. This is because period may have non-constant effect on mortality as well as on the relationship between migration status and mortality.

Findings

Temporary migration

Figures 1 and 2 show temporary migration levels and trends by sex and age group. The migration rates are quite stable over time, considering that they are independently measured each year, but there is a gradually increasing

Table 3. Factors associated with death from non-communicable disease, by migration category and period, controlling for age and sex

	Hazard ratio (95% Confidence intervals); death by non-communicable disease		
Variables	2000–2003	2004–2007	2008–2011
Sex: male vs. female	1.50*** (1.25–1.81)	2.06*** (1.76–2.42)	1.81*** (1.54–2.12)
Short duration temporary migrant vs. not	1.88** (1.05–3.35)	1.70 (0.81–3.59)	1.18 (0.37–3.76)
Long duration temporary migrant vs. not	1.55*** (1.17–2.04)	1.08 (0.87–1.33)	0.60*** (0.48–0.74)
Observations	52,680	52,543	55,847
Wald Chi-square	41.23	90.19	60.68
Log likelihood	−3,195	−4,283	−4,600
Subjects	48,661	49,573	53,271
Failures	493	645	698

***$p < 0.01$, **$p < 0.05$.

Citation: Glob Health Action 2014, **7**: 23514 - http://dx.doi.org/10.3402/gha.v7.23514

Table 4. Hazard ratios showing the association of factors with death from an external cause by period and migration category, controlling for age and sex

	Hazard ratio (95% Confidence intervals); death by external causes		
Variables	2000–2003	2004–2007	2008–2011
Sex: male vs. female	4.53*** (3.04–6.75)	5.46*** (3.78–7.89)	5.18*** (3.56–7.55)
Short duration temporary migrant vs. not	9.50*** (5.47–16.49)	3.30** (1.32–8.28)	2.24 (0.69–7.24)
Long duration temporary migrant vs. not	5.78*** (3.82–8.74)	3.04*** (2.16–4.29)	1.45** (1.00–2.10)
Observations	52,680	52,543	55,847
Wald Chi-square	205.6	135.5	85.82
Log likelihood	−993.0	−1,165	−1,124
Subjects	48,661	49,573	53,271
Failures	147	166	153

***$p < 0.01$, **$p < 0.05$.

trend for some age groups. A striking feature is the high migration rate, especially for males aged 20–59. In the first half of the observation window, men of this age group were 40–60% likely to be a temporary migrant, while in the second half of the observation period they were 50–70% likely to meet the criteria for being a temporary migrant. For women, the ages 20–49 years were the most mobile, with migration levels tending to increase over the observation period, from 20 to 25% being a temporary migrant at the start of the period, to 25–35% of women being a temporary migrant at the end. Thus, for working age adults, both males and females showed an increasing likelihood of being a temporary migrant between 2000 and 2011.

Temporary migration is linked to a change in socio-economic status in rural households in northeast South Africa (7). Table 1 shows two patterns linking migration and economic status for men and women. The upper aspect of the table shows the percent of temporary migrants that were employed. Across the observation period, approximately 75% of male temporary migrants and 50% of female temporary migrants were employed. The gender difference is highly significant with male migrants more likely to be in employment ($p < 0.000$). The lower aspect of the table shows the proportion of employed temporary migrants that remitted cash or another item back to the origin household. In 2002, female employed migrants were slightly more likely than male employed migrants to remit ($p < 0.065$). This difference became clearer over the decade and by 2012 it was significant ($p < 0.000$) that employed female migrants were more likely to remit than employed male migrants.

All-cause mortality

Figures 3 on all-cause mortality shows a clear pattern. The 2000–2003 period displays a constant and high increase in overall mortality; mortality is stable at high level during the 2004–2007 period; and mortality declined in the 2008–2011 period. To note, ART roll-out started systematically in 2008, with almost immediate effect on

Table 5. Factors associated with death from unspecified cause of death, by period and migration status, controlling for age and sex

	Hazard ratio (95% Confidence intervals); death by unspecified causes		
Variables	2000–2003	2004–2007	2008–2011
Sex: male vs. female	1.40* (0.94–2.07)	1.63*** (1.15–2.32)	2.40*** (1.59–3.63)
Short duration temporary migrant vs. not	4.00*** (1.73–9.24)	3.93** (1.23–12.53)	1.98 (0.29–13.78)
Long duration temporary migrant vs. not	3.01*** (1.86–4.88)	2.75*** (1.85–4.07)	1.33 (0.85–2.08)
Observations	52,680	52,543	55,847
Wald Chi-square	38.28	40.22	21.05
Log likelihood	−751.2	−892.8	−649.0
Subjects	48,661	49,573	53,271
Failures	107	126	92

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

the decline of CD cause of death (11). Non-linear trends over the 2000–2011 period are essentially attributable to CDs mortality for both sexes (Fig. 4) and to NCD for males (Fig. 5), with a gradual increase in NCD mortality for females (53).

Mortality from a CD

In Fig. 4, CD mortality trends show a rise and then a fall over the course of the study period. The scourge of AIDS and TB impacted these mortality profiles between 2000 and 2007 (11, 53) and health system made anti-retroviral treatment available in the period 2008–2011 (11).

Table 2 shows three comparable multivariate competing risk regression models, one for each period, enabling us to examine the relationship between temporary migration and CD mortality over time, with temporary migration discriminated by short duration (≤ 3 years) and longer duration (>3 years). There are important associations between migration and mortality from CDs. At the start of the period, temporary migration shows a particularly high risk of mortality with long duration temporary migration having 2.7 times higher risk of death than non-migrants and short duration temporary migration 2.5 times higher risk. Over time the effect lessens although it remains significantly negative for long duration temporary migrants in the middle period. In the last period, shown in the right-hand column, longer duration temporary migration has become a significantly protective factor from CD mortality. The environment for temporary migration has shown an important shift from high risk of exposure to HIV, TB, and acute respiratory infections, possibly due to the associated disruption from normal family supports and the un-healthy environments at destination, to an environment where the benefits of migration have improved, including better access to health care and relative socioeconomic well-being for migrants. By the end of the period, longer duration temporary migration has become positively associated with survival and the picture is mixed for shorter duration temporary migrants.

Mortality from NCD

The age-adjusted level in NCD mortality increased in both sexes over the decade of observation, with male rates increasing in the earlier half of the observation period and then leveling off, and female mortality rates from NCD increasing steadily over the period (53). The Agincourt mortality profiles shown in a Supplementary file indicate that over the 8 years from 2004 to 2011, the percentage of mortality attributable to an NCD increased in all age and sex groups.

Table 3 shows three competing risk regression models, examining the relationship between temporary migration and NCD mortality over time. At the start, temporary migration shows a strong association with NCD mortality with long duration temporary migration having 1.9 times higher risk of death than non-migrants and short-duration temporary migration 1.5 times higher risk. As with CD mortality, the effect lessens over time and becomes non-significant in the middle period, and protective in the last period for long duration temporary migration. This may reflect an overall shift in the dangers associated with long-term temporary migration, from one of negative lifestyle changes, disruption, and stress at the start of the period, associated with high mortality rates from NCD for temporary migrants compared to non-migrants, to a lessening of this relationship over time, associated with an increasingly protective effect of long-term temporary migration at the end of the period.

Migration and mortality from external causes

Temporary migration is significantly associated with mortality from external causes, including assault, motor vehicle accidents, self-harm, and other causes. The gender dimension is striking with men four to five times more likely to die from an external cause compared to women. Both short and long duration temporary migrations stand out as exposures significantly associated with increased risk of mortality from an external cause of death at the start of the decade, but the negative relationships lessen over time.

Migration and mortality from undetermined causes

For undetermined causes of death, the numbers of cases are much fewer than the main causes of death and the general impact of migration follows a similar pattern as seen above with long duration temporary migration showing a significant mortality risk in the first two thirds of the period, which has become insignificant by the end of the period.

Discussion

The availability of long-term demographic and health surveillance data allows us to gain insights into the trends in temporary migration and mortality and their relationship over the period 2000–2011. The results relate to patterns of urban transition in settings with high levels of temporary circular migration. Analytically, temporary migration identifies the individuals that are away for most of the time, but remain linked to the rural households of origin. The data presented show the stability and even growth of temporary migration patterns in rural, north-east South Africa. Remittance patterns are also stable and although men are more likely to migrate for employment reasons, women are increasingly likely to migrate and a half of these female migrants are also employed. In this paper, we do not explore the frequency, quantity, or size of the remittances, but the two variables, 'percent of migrants employed' and 'percent of employed migrants that remit', indicate the intrinsic economic importance of temporary migration. Furthermore, employed female temporary

Citation: Glob Health Action 2014, **7**: 23514 - http://dx.doi.org/10.3402/gha.v7.23514

migrants are more likely than males to remit something back to the original household. Under apartheid, rural South African households were forced to participate in an imposed migrant labor system, but current data show the ongoing relevance of labor migration as a rural livelihood strategy for men and women.

There are features in the mortality trends presented here that are consistent with the conventional epidemiological transition in this rural sub-district. Coupled with severe mortality increases from CDs in the first half of the period, due primarily to AIDS and TB (53), mortality from NCDs has been increasing, albeit more gradually, which has led to the pattern of a dual burden of illness (17). The mortality rates from AIDS/TB started coming down from 2008/9. Although the reasons for this are not directly examined in the paper, there has been a widespread recognition that making anti-retroviral therapy available at public hospitals and health centers has played a major role (11). Thus, we see in this one setting, aspects of the conventional epidemiological transition, but equally important key aspects of the revision to that transition more in step with current thinking regarding the GBD.

Most notably our empirical work suggests that caution should be taken in not making too simplistic an association of age and geography with health outcomes. In many instances, temporary migration is more important than age and sex in explaining variations in mortality. Temporary migration is shown to have a relationship with mortality from CDs, but the relationship is not consistent over time. Notable here is that the Agincourt HDSS data are distinctly privileged in identifying and following temporary migrants; other data sources (often on which discussions of health transition are based) do not have such detail. A single interval of about a decade in length is arguably insufficient to either confirm a new pattern for an epidemiological transition or refute it. Still, the range of changes we see, included dramatic prospective shifts in health prospects due to HIV/AIDS treatment, along with the ongoing dramatic changes in population mobility and the technology of transportation and communication, argue for continued attention to the dynamics of migration and health.

We find a strong association of migration and CD mortality, especially for male temporary migrants, in the early part of the observation period. The relationship is most likely driven by a process whereby temporary labor migrants that become ill as a result of exposure outside home, may eventually die in their rural households (56). In the latter part of the observation period conditions for temporary migrants seem to have improved such that the association of temporary migration and survival becomes a positive relationship. Causes are not directly examined in the study, but three factors can play a role. First, the risk of transmission of infectious disease has possibly declined through temporary migration becoming more selective in nature and the process less stressful for the migrant. Second, temporary migrants and their origin households are more connected through mobile phone technology, which can lessen the impetus for extra-marital sexual partnerships. Third, the improved access to health care may have benefited migrants, especially by ART being made available through government and non-government programs targeting migrant workers.

The association of migration and NCD mortality shows a relationship that transitioned toward a healthier association, with temporary migration starting as a possible risk factor at the beginning of the period, but becoming more positive over time. Although we do not have causal models to directly examine risk factors in the temporary migration and mortality relationships, we have included a small interpretive device by discriminating between shorter and longer duration temporary migrants in the regressions. As explained in the literature review, there is evidence of both positive and negative health implications of migration in different settings, and similarly positive and negative implications for household socioeconomic status reported in different studies. To explore which direction our data point we can assume that longer duration temporary migrants of both sexes are more likely to have succeeded in gaining employment and thus are more likely to be a consistent wage earner. Shorter duration migrants are less likely to be consistently earning a wage. If temporary migrants' relationship to mortality is influenced by exposure to unhealthy environments, diet and or other lifestyle factors, like alcohol and smoking, then shorter duration migrants would be exposed for shorter times and thus show less association with mortality, and, conversely, longer term migrants show more association with mortality. Although far from conclusive, our findings indicate that CD mortality has a stronger association with longer duration migration in the earlier part of the period, consistent with unhealthy exposure, which changes over the observation period into an association with better survival chances, consistent with these migration-related risks lessening over time. NCD mortality shows a somewhat different relationship with migration. A longer duration of migration is consistently less associated with mortality than shorter duration migration throughout the observation period. This may indicate that the economic benefit of longer duration temporary migration is playing a stronger positive role than unhealthy exposures playing a negative role in the relationship between migration and mortality. A similar picture is shown for external causes of death, which is even more sensitive to migration than CD or NCD mortality. In summary, risks of CD mortality may be higher for short-term migrants due to disruption and exposure, whereas for NCD mortality, the relationship to migration is driven more by the positive relation of the earnings of longer

duration migrants making them less likely to succumb to a fatal illness.

While no observational study such as ours can reach the inferential conditions of a true randomized experiment (necessarily so), our design and analysis improve significantly on comparative cross-sectional analysis while also pointing the way to more sophisticated, promising data collection, and analysis for the future. It is probably true that we will not have randomized experiments regarding migrants' status (57). Therefore, carefully crafted longitudinal studies of the redistribution of persons across geography (and associated changes in risk regime and livelihood) will be crucial for understanding the 21st century version of the epidemiological transition in sub-Saharan Africa.

Health and social policy implications arise from the importance and vitality of the rural–urban links shown in the study. When the circular nature of migration from poorer communities is not well understood, the balance between rural and urban development is potentially misconstrued. Migration dynamics that keep households economically vital are conceptually flattened through the temporary migrants being grouped into the urban category while their households of origin, to which they remain economically linked, are classified as rural.

A study limitation arises from the fact that the data are from a small sub-district in remote rural South Africa and we need to consider how representative the data are. We know the levels of circular labor migration are high, and in other areas may not be as high, but it is still valuable to examine the case study to get an empirical understanding of the trends and associations between migration and mortality. Many HDSS data systems reveal active migration patterns and comparative studies are starting to emerge (13).

Another limitation of our study is that the covariates were restricted to age, sex, period, and temporary migration status, while the outcome was restricted to large causes of death. Whereas this was useful to keep enough statistical power for our estimates, further research should aim at examining the migration–health relationship for different socioeconomic categories and also for specific diseases. This requires reconstructing individual, household and community histories consistently over the observation period in relation to migration histories. This is currently being tackled. Barring data availability, this will help target sub-populations with specific health needs.

In this study, we examined the evolution of the relationship between temporary migration status and causes of death. However, the permanent residents who form the baseline category in our study are not necessarily homogenous. A number of selective processes – some of them related to health – may change the composition of the permanent population. Previous research (56, 58) based on data from the 1990s to the mid-2000s has shown how the migrants returning sick after a long period of residence out of the site have highly contributed to reduction in life expectancy. While these studies confirm the 'unhealthy return migrant' hypothesis, studies are lacking to confirm in Agincourt HDSS the more famous 'healthy migrant' hypothesis. Future research will benefit considerably from giving more attention to the influence of individual and household migration histories on the prevalence and incidence of both communicable and NCDs, and longitudinal data-collection platforms such as the one we employ are well-suited to provide such information. Moreover, such extensive longitudinal approaches can be readily supplemented with data collection on the migrants at destination and information about social and monetary connections between origin and destination. As epidemiologists and health policy-makers demand more definitive information about the determinants of the contemporary health transition, such expanded inquiry offers considerable promise for understanding health implications of migration and urbanization.

Main findings

- In rural, northeast South Africa, temporary migration involving migrants that relocate mainly for work purposes and remain linked to the rural household is more important than age and sex in explaining variations in mortality, whatever the cause.
- In this setting, the changing relationship between temporary migration and communicable disease mortality is primarily affected by reduced exposure of the migrant to unhealthy conditions.
- The study suggests that the changing relationship between temporary migration and non-communicable disease mortality is mainly affected by increased livelihood benefits of longer duration migration.

Key messages for action

- Health care access should be expanded for migrants, especially by making patient information available to health practitioners in rural healthcare system to enhance continuity of care when a migrant returns home.
- The rural healthcare system should be improved, as migrants tend to return to the rural households when in need of health care.
- Temporary migration is associated with both non-communicable and communicable diseases; therefore, public health policies should account for population mobility whatever the targeted health risk.

Citation: Glob Health Action 2014, **7**: 23514 - http://dx.doi.org/10.3402/gha.v7.23514

Acknowledgements

We greatly value the involvement of community leaders, study participants, and district and provincial managers in health, education, and other government sectors; the School of Public Health and Faculty of Health Sciences, University of the Witwatersrand, the Medical Research Council and Limpopo and Mpumalanga Provinces, South Africa, the Wellcome Trust, UK (Grants 058893/Z/99/A; 069683/Z/02/Z; 085477/Z/08/Z), The William and Flora Hewlett Foundation, National Institute on Aging (NIA) of the NIH, the Andrew W Mellon Foundation, the (Lifespan/Tufts/Brown) Center for AIDS Research (P30 AI042853), and the Brown University, Population, Studies and Training Center (NICHD R24 527018), USA.

Conflict of interest and funding

The authors have not received any funding or benefits from industry or elsewhere to conduct this study.

References

1. Murray C. Families divided: the impact of migrant labour in Lesotho. Johannesburg: Ravan; 1981.
2. Hosegood V, Benzler J, Solarsh GC. Population mobility and household dynamics in rural South Africa: implications for demographic and health research. South Afr J Demogr 2005; 10: 43–68.
3. Posel D, Casale D. What has been happening to internal labour migration in South Africa, 1993 – 1999? S Afr J Econ 2002; 71: 455.
4. Collinson MA, Tollman SM, Kahn K. Migration, settlement change and health in post apartheid South Africa: triangulating Agincourt demographic surveillance with national census data. Scand J Publ Health 2007; 35(Suppl 69): 77–84.
5. Wilson F. Minerals and migrants: how the mining industry has shaped South Africa. Daedalus 2001; 130: 99–122.
6. Reed HE. Moving across boundaries: migration in South Africa, 1950–2000. Demography 2013; 50: 71–95. DOI: 10.1007/s13524-012-0140-x.
7. Collinson MA. Striving against adversity: the dynamics of migration, health and poverty in rural South Africa. Glob Health Action 2010; 3. DOI: 10.3402/gha.v3i0.5080. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2882287/
8. McCulloch J. Counting the cost: gold mining and occupational disease in contemporary South Africa. Afr Aff 2009; 108: 221–40.
9. Lurie M, Harrison A, Wilkinson D, Abdool Karim S. Circular migration and sexual networking in rural Kwazulu/Natal: implications for the spread of HIV and other sexually transmitted diseases. Health Trans Rev 1997; 7(Suppl 3): 17–27.
10. Abdool Karim SS, Churchyard GJ, Abdool Karim Q, Lawn SD. HIV infection and tuberculosis in South Africa: an urgent need to escalate the public health response. Lancet 2009; 374: 921–33.
11. Gómez-Olivé FX, Angotti N, Houle B, Klipstein-Grobusch K, Kabudula C, Menken J, Williams J, Tollman SM, Clark SJ. Prevalence of HIV among those 15 and older in rural South Africa. AIDS Care 2013; 25: 9: 1122–1128.
12. Packard RM. White plague, black labor: tuberculosis and the political economy of health and disease in South Africa. University of California Press; 1989.
13. Collinson MA, Adazu K, White MJ, Findley SE, eds. The dynamics of migration, health and livelihoods: INDEPTH Network perspectives. Aldershot, UK: Ashgate; 2009.
14. Vallin J. Commentary: 'epidemiologic transition' interrupted or sweep to the second stage of 'health transition'. Int J Epidemiol 2007; 36: 384–6.
15. 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.
16. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2013; 380: 2095–128.
17. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2013; 380: 2224–60.
18. Omran AR. The epidemiologic transition: a theory of the epidemiology of population change. Milbank Mem Fund Q 1971; 49: 509–38.
19. Popkin BM. Nutritional patterns and transitions. Popul Dev Rev 1993; 19: 138–57.
20. Salomon JA, Murray CJ. The epidemiologic transition revisited: compositional models for causes of death by age and sex. Popul Dev Rev 2002; 28: 205–28.
21. Ezzati M, Vander Hoorn S, Lawes CM, Leach R, James WP, Lopez AD, et al. Rethinking the "diseases of affluence" paradigm: global patterns of nutritional risks in relation to economic development. PLoS Med 2005; 2. DOI: 10.1371/journal.pmed.0020133.
22. Lesthaeghe R. The unfolding story of the second demographic transition. Popul Dev Rev 2010; 36: 211–51.
23. Dyson T. Population and development: the demographic transition. London: Zed Books; 2010.
24. Caldwell JC. Toward a restatement of demographic transition theory. Popul Dev Rev 1976; 2: 321–66.
25. Popkin BM. Global nutrition dynamics: the world is shifting rapidly toward a diet linked with noncommunicable diseases. Am J Clin Nutr 2006; 84: 289–98.
26. Prentice AM. The emerging epidemic of obesity in developing countries. Int J Epidemiol 2006; 35: 93–9.
27. World Health Organization (2011). WHO global Infobase. Available from: https://apps.who.int/infobase/Mortality [cited 6 July 2011].
28. Boutayeb A. The double burden of communicable and non-communicable diseases in developing countries. Trans R Soc Trop Med Hyg 2006; 100: 191–9.
29. Doak CM, Adair L, Bentley M, Monteiro C, Popkin BM. The dual burden household and the nutrition transition paradox. Int J Obes 2004; 29: 129–36.
30. Dalal S, Beunza JJ, Volmink J, Adebamowo C, Bajunirwe F, Njelekela M, et al. Non-communicable diseases in sub-Saharan Africa: what we know now. Int J Epidemiol 2011; 40: 885–901.
31. Normile D. A sense of crisis as China confronts ailments of affluence. Science. 2010; 328: 422.
32. World Health Organization (2005). Preventing chronic diseases: a vital investment. Geneva. Available from: http://www.who.int/chp/chronic_disease_report/contents/en/index.html [cited 10 January 2014]
33. Garenne M, Tollman SM, Kahn K, Collinson MA. Fertility trends and net reproduction in Agincourt, rural South Africa: 1992–2004. Scand J Publ Health 2007; 35(Suppl 69): 68–76.
34. World Bank (2009). World Development Report: Reshaping economic geography. Washington, DC: The World Bank.

35. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. Lancet 2006; 367: 1747–57.
36. BeLue R, Okoror TA, Iwelunmor J, Taylor KD, Degboe AN, Agyemang C, et al. An overview of cardiovascular risk factor burden in sub-Saharan African countries: a socio-cultural perspective. Global Health 2009; 5: 1–12.
37. Mbanya JC, Motala AA, Sobngwi E, Assah FK, Enoru ST. Diabetes in sub-Saharan Africa. Lancet 2010; 375: 2254–66.
38. Population Reference Bureau (2011). World Population Data Sheet. Washington, DC: Population Reference Bureau.
39. Harpham T, Tanner M, eds. Urban health in developing countries: progress and prospects. London: Earthscan; 1995.
40. Zulu EM, Beguy D, Ezeh AC, Bocquier P, Madise NJ, Cleland J, et al. Overview of migration, poverty and health dynamics in Nairobi City's slum settlements. J Urban Health 2011; 88: 185–99.
41. Yusuf S, Reddy S, Ôunpuu S, Anand S. Global burden of cardiovascular diseases part I: general considerations, the epidemiologic transition, risk factors, and impact of urbanization. Circulation 2001; 104: 2746–53.
42. Harpham T. Urban health in developing countries: what do we know and where do we go? Health Place 2009; 15: 107–16.
43. Lurie M, Williams B, Zuma K, Mkaya-Mwaburi D, Garnett GP, Sturm AW, et al. The impact of migration on HIV1 transmission in South Africa: A study of migrant and non-migrant men and their partners. Sex Trans Dis 2003; 30: 149–56.
44. White MJ, Lindstrom DP. Internal migration. In: Poston DL, Micklin M, eds. Handbook of Population. New York: Springer; 2005, pp. 307–42.
45. Van Wey LK. Altruisitic and contractual remittances between male and female migrants and households in rural Thailand. Demography 2004; 41: 739–56.
46. White MJ, Muhidin S, Andrzejewski C, Tagoe E, Reed H, Knight R. Urbanization and fertility: an event-history analysis of coastal Ghana. Demography 2008; 45: 803–16.
47. Montgomery MR, Stren R, Cohen B, Reed HE, eds. Cities transformed: demographic change and its implications in the developing world. Washington, DC: National Acadamies Press; 2003.
48. Potts D. The slowing of sub-Saharan Africa's urbanization: evidence and implications for urban livelihoods. Environ Urban 2009; 21: 253–9.
49. Kahn K, Collinson MA, Gómez-Olivé FX, Mokoena O, Twine R, Mee P, et al. Profile: Agincourt health and socio-demographic surveillance system (Agincourt HDSS). Int J Epidemiol 2012; 41: 988–1001. DOI: 10.1093/ije/dys115.
50. Kahn K, Tollman SM, Collinson MA, Clark SJ, Twine R, Clark BD, et al. Research into health, population and social transitions in rural South Africa: data and methods of the Agincourt health and demographic surveillance system. Scand J Publ Health 2007; 35(Suppl 69): 8–20.
51. Kahn K, Tollman SM, Garenne M, Gear JSS. Validation and application of verbal autopsies in a rural area of South Africa. Trop Med Int Health 2000; 5: 824–31.
52. Kahn K, Tollman S, Garenne M, Gear J. Who dies from what? Determining cause of death in South Africa's rural north-east. Trop Med Int Health 1999; 4: 433–41.
53. Tollman SM, Kahn K, Sartorius B, Collinson MA, Clark SJ, Garenne ML. Implications of mortality transition for primary health care in rural South Africa: a population-based surveillance study. Lancet 2008; 372: 893–901.
54. Byass P, Kahn K, Fottrell E, Collinson MA, Tollman SM. Moving from data on deaths to public health policy in Agincourt, South Africa: approaches to analysing and understanding verbal autopsy findings. PLoS Med 2010; 7: e1000325.
55. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999; 94: 496–509.
56. Clark SJ, Collinson MA, Kahn K, Drullinger K, Tollman SM. Returning home to die: circular labour migration and mortality in rural South Africa. Scand J Publ Health 2007; 35(Suppl 69): 35–44.
57. Fu H, Van Landingham MJ. Mental health consequences of international migration for Vietnamese Americans and the mediating effects of physical health and social networks: results from a natural experiment approach. Demography 2012; 49: 393–424.
58. Bocquier P, Collinson MA, Clark SJ, Gerritsen AG, Kahn K, Tollman SM. Ubiquitous burden: the contribution of HIV/TB and migration to mortality in rural South-Africa. Afr Popul Stud 2014. (forthcoming).

Global Health Action

PART IV

Closing the mental health treatment gap in South Africa: a review of costs and cost-effectiveness

Helen Jack[1], Ryan G. Wagner[2,3], Inge Petersen[4], Rita Thom[5], Charles R. Newton[1,2], Alan Stein[1,2], Kathleen Kahn[2,3,6], Stephen Tollman[2,3,6] and Karen J. Hofman[2,5]*

[1]Department of Psychiatry, University of Oxford, Oxford, UK; [2]MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; [3]Umeå Centre for Global Health Research, Umeå University, Umeå, Sweden; [4]School of Applied Human Sciences, University of KwaZulu Natal, Durban, South Africa; [5]PRICELESS SA (Priority Cost Effective Lessons in System Strengthening South Africa), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; [6]INDEPTH Network, Accra, Ghana

Background: Nearly one in three South Africans will suffer from a mental disorder in his or her lifetime, a higher prevalence than many low- and middle-income countries. Understanding the economic costs and consequences of prevention and packages of care is essential, particularly as South Africa considers scaling-up mental health services and works towards universal health coverage. Economic evaluations can inform how priorities are set in system or spending changes.

Objective: To identify and review research from South Africa and sub-Saharan Africa on the direct and indirect costs of mental, neurological, and substance use (MNS) disorders and the cost-effectiveness of treatment interventions.

Design: Narrative overview methodology.

Results and conclusions: Reviewed studies indicate that integrating mental health care into existing health systems may be the most effective and cost-efficient approach to increase access to mental health services in South Africa. Integration would also direct treatment, prevention, and screening to people with HIV and other chronic health conditions who are at high risk for mental disorders. We identify four major knowledge gaps: 1) accurate and thorough assessment of the health burdens of MNS disorders, 2) design and assessment of interventions that integrate mental health screening and treatment into existing health systems, 3) information on the use and costs of traditional medicines, and 4) cost-effectiveness evaluation of a range of specific interventions or packages of interventions that are tailored to the national context.

Keywords: *mental health*; *South Africa*; *economics*; *health planning*; *policy*; *costs and cost analysis*

Responsible Editors: Nawi Ng, Umeå University, Sweden; Barthélémy Kuate Defo, University of Montreal, Canada.

*Correspondence to: Karen J. Hofman, PRICELESS and MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, University of the Witwatersrand, Room 231, 27 St Andrews Road, Parktown, Johannesburg 2193, South Africa, Email: Karen.Hofman@wits.ac.za

This paper is part of the Special Issue: *Epidemiological Transitions – Beyond Omran's Theory*. More papers from this issue can be found at http://www.globalhealthaction.net

Received: 25 November 2013; Revised: 20 December 2013; Accepted: 21 December 2013; Published: 15 May 2014

Mental, neurological, and substance use (MNS) disorders accounted for 10% of the global burden of disease (GBD) in 2010 (1), yet on average, mental health accounts for less than 1% of national health budgets in Africa and South East Asia (2). In South Africa, as in many low- or middle-income countries (LMICs), the burden of mental disorders has grown over the past 20 years (1990–2010) (1). This rise is expected to continue, in part due to the ongoing epidemiological transition from communicable to non-communicable diseases (NCDs) (3) and co-morbidity between MNS disorders, HIV, and other chronic health conditions (4–6).

In 2011, faced with this growing burden, South Africa's Ministry of Health publically committed to increasing by 30% the number of people screened and treated for

Citation: Glob Health Action 2014, **7**: 23431 - http://dx.doi.org/10.3402/gha.v7.23431

mental disorders by 2030, and to reducing by 20% per capita alcohol consumption by 2020 (5, 7, 8). Health budgets and system designs, however, do not currently reflect these new commitments. For example, the District Specialist Teams introduced under the proposed national health insurance provide key specialty services to supplement primary health care at the district-level, but the teams do not include mental health providers (9). Better understanding of cost-effective, context-specific interventions or packages of interventions for treatment and prevention of MNS disorders may contribute to achieving South Africa's ambitious mental health targets.

Economic data is one of several relevant factors, including burden of disease and equity, that policymakers and donors consider as they set priorities and make spending choices in resource-limited settings. Economic data provides an 'external frame' that can be used to make comparisons between competing priorities or interventions and justify investment in programs (10, 11). Cost-effectiveness analyses (CEAs), which compare interventions to determine those likely to yield the most improvements in health per dollar (12), have achieved notable successes as an advocacy tool for system improvements (13, 14).

Despite the importance of economic information in advocating for and designing policy change, there are no reviews of costs related to MNS disorders and cost-effectiveness of their treatment in South Africa, and few reviews that examine costing aspects of MNS disorders in any LMICs (2). Compiling the disparate information on costs and cost-effectiveness at a country-level could facilitate identification of key needs, interventions, and knowledge gaps. A health system approach, which goes beyond single disorders or interventions, is necessary because MNS disorders often occur together and may be co-morbid with other health conditions, and a holistic view can inform the design of treatment packages (15).

Mental health care in South Africa differs from that in other LMICs, making a South Africa-specific review necessary. Although South Africa's gross national income (GNI) bears greater similarity to other middle-income countries than to the rest of sub-Saharan Africa, findings on MNS disorders from middle-income nations, including those in Latin American or South East Asia, cannot be generalized to South Africa. South Africa has a unique post-*apartheid* socioeconomic and cultural context of inequality, with one of the highest Gini coefficients globally (a statistical measure of income inequality in a population) (16), and particular disparities between rural and urban areas. It also has a complex disease burden characterized by high HIV prevalence and a growing burden of non-communicable chronic conditions (17).

Given the paucity of economic assessments of mental health in sub-Saharan Africa and the specific challenges facing South Africa's mental health system, this review summarizes current understanding and highlights key knowledge gaps. Findings may inform future research and the design of mental health policy and interventions in South Africa, other nations in sub-Saharan Africa, and settings with a high prevalence of conditions, including HIV, that may be co-morbid with mental disorders.

Prevalence and epidemiological burden of MNS disorders in South Africa

Globally, the World Health Organization (WHO) estimates that 30.8% of all years lived with disability (YLDs) are due to neuropsychiatric disorders, primarily unipolar depression (11.9%), alcohol use disorder (3.1%), schizophrenia (4.8%), and bipolar mood disorder (4.4%) (18). Three MNS disorders (unipolar depressive disorders, self-inflicted injuries, and alcohol use disorders) are among the top 20 causes of disability-adjusted life years (DALYs) lost globally (18, 19), and MNS disorders account for a larger percentage of lost DALYs than cardiovascular disease or cancer.

The South African Stress and Health (SASH) Study, conducted between 2002 and 2004, provides the only nationally representative data on the prevalence of common mental disorders (20, 21). Other prevalence studies examine specific populations and disorders (22, 23), but do not provide the national representativeness of the SASH (24). Table 1 shows that lifetime prevalence of common mental disorders was 30.3%, and prevalence in the 12 months prior to the survey was 16.5%.

SASH, part of the WHO World Mental Health Survey Initiative, is a cross-national effort to collect country-specific epidemiological data on mental disorders using a single assessment tool and data collection methodology. It includes seven LMICs: China, Columbia, Lebanon, Mexico, Nigeria, South Africa, and Ukraine. Lifetime prevalence rates of select types of MNS disorders are shown in Table 1 (25). Examining all countries surveyed, the highest lifetime prevalence of these disorders is in the US and New Zealand (47.4 and 39.3%, respectively) while the lowest is in Nigeria and China (12.0 and 13.2%, respectively). South Africa has more than twice the lifetime prevalence of mental and substance use disorders than Nigeria (12.0%), the only other African country surveyed (21, 26), and a greater lifetime prevalence than all LMICs except Columbia and Ukraine. However, these comparisons must be interpreted cautiously because researchers acknowledge that prevalence may be underreported in Nigeria and China due to stigma and lack of public familiarity with surveys (27); prevalence of impulse control disorders are measured in all LMICs except South Africa, potentially altering the overall prevalence estimate (25); and the manifestations of mental disorders and their diagnostic criteria may vary between cultural

Citation: Glob Health Action 2014, **7**: 23431 - http://dx.doi.org/10.3402/gha.v7.23431

Table 1. Prevalence of select categories of MNS disorders in South Africa and other LMICs countries included in the WHO World Mental Health Survey Initiative

MNS disorder	Lifetime prevalence (South Africa) (%)	12-month prevalence (South Africa) (%)	Lifetime prevalence (China) (%)	Lifetime prevalence (Columbia) (%)	Lifetime prevalence (Lebanon) (%)	Lifetime prevalence (Mexico) (%)	Lifetime prevalence (Nigeria) (%)	Lifetime prevalence (Ukraine) (%)
Anxiety disorders	15.8	8.1	4.8	25.3	16.7	14.3	6.5	10.9
Substance use disorders	13.3	5.8	4.9	9.6	2.2	7.8	3.7	15.0
Mood disorders	9.8	4.9	3.6	14.6	12.6	9.2	3.3	15.8
Any disorder	30.3	16.5	13.2	39.1	25.8	26.1	12.0	36.1

Sources: Williams et al. [20], Herman et al. [24], Kessler et al. [25].

contexts, making any single diagnostic instrument potentially unfit to capture the range of ways mental disorders may be expressed (28). In this review, we have chosen to focus our epidemiological examination only on data from the WHO World Mental Health Survey to ensure uniform data collection methodology and consistency in definitions of mental disorders. Overall, there are less epidemiological data on mental disorders than on other disorders in LMICs and few panel data available to examine change in their prevalence or burden over time (29).

The SASH data suggest that the high prevalence of common mental disorders may be caused by exposure to stress and trauma during *apartheid* and the ongoing period of racial tension and inequality following *apartheid* (21). The SASH data shows that 74.8% of South Africans have experienced at least one traumatic event, most commonly trauma related to someone close to them (for example, death of a friend or family member), witnessing a traumatic event, or being the victim of criminal or intimate partner violence (20). These traumas and other life stressors, such as economic hardship and relationship problems, were associated with increased 12-month and lifetime prevalence of common mental disorders (30).

In 2000, neuropsychiatric disorders (including mental and nervous system disorders) ranked third in their contribution to South Africa's national burden of disease. Table 2 shows the contributions of individual MNS disorders to that burden (31). YLD data was not directly collected for the country and, consequently, DALY and YLD estimates must be interpreted cautiously. Suicide, the only MNS disorder in the top 20 leading causes of YLLs, ranked thirteenth (1.3%). Suicide causes 5,514–7,582 deaths per year (17), with the number of suicide deaths of people aged under 35 and over 65 increasing between 1968 and 1990 (32).

MNS disorders are commonly co-morbid with HIV, and the conditions are mutually reinforcing (6, 24, 33). Considerable research has focused on the high prevalence of common mental disorders in HIV-positive patients (33–36), with one study reporting that 35% of HIV patients in South Africa (n = 100) meet the criteria for major depressive disorder, 6% for bipolar mood disorder, and 21% for generalized anxiety disorder (33), far higher than prevalence estimates for the general population (24, 33). Among patients with severe mental illness admitted to a psychiatric hospital (n = 206), 29.1% were HIV positive, nearly triple the general population prevalence (37). These co-morbidity data may not reflect the current situation, as they were collected before the rollout of ARVs and at the tertiary health care level where people with the most advanced HIV-related disease present for treatment; those who did not go to the hospital or who visited primary care clinics were not included.

Table 2. Contributions of MNS disorders to South Africa's burden of disease

MNS disorder	Percentage of burden of disease (South Africa)	Ranking in contribution to burden of disease (South Africa)
Unipolar depressive disorders	5.8	2
Alcohol use	2.8	6
Bipolar mood disorder	2.1	9
Schizophrenia	2.1	11
Drug use	1.6	14
Foetal alcohol syndrome	1.1	16
Obsessive compulsive disorder	1.0	18
Panic disorder	1.0	19

Source: Norman et al. [31].

Depression and other mental disorders are of particular concern in patients with HIV because they can lead to suboptimal treatment adherence, and consequently, lower CD4 counts, increased viral load, and a greater chance of developing drug-resistant strains of HIV that require more costly second-line anti-retroviral therapy (38). A diagnosis of HIV also complicates treatment of MNS disorders because of interactions between antiretroviral drugs and other medications. Phenobarbital, for example, a common treatment for epilepsy in sub-Saharan Africa, substantially reduces the half-life of some anti-retrovirals, lowering their therapeutic efficacy (39).

There has been little published research on co-morbidities between MNS disorders and other chronic diseases, yet existing data from South Africa and elsewhere in sub-Saharan Africa suggests association between mental disorders and diabetes, stroke, and epilepsy (40–45).

Service delivery infrastructure

In 2005, South Africa devoted 2.7% of its health budget to mental health care (19), more than twice that of Ghana, Uganda, and many other low- to middle-income countries (46), but less than high-income nations, such as the UK, which uses 10.8% of its health budget on mental health. Brazil and India, South Africa's middle-income peers, spend 2.38 and 0.06% of their health budgets on mental health, respectively (47). Mental health care budget data in the WHO's Mental Health Atlas was updated in 2011, but did not include new South African budget data. Consequently, 2005 data, the most recent available, is given here for South Africa.

As in many LMICs, the mental health system in South Africa is fragmented. Mental health care in South Africa has historically been reliant on psychiatric hospitals, with little attention to mental health in primary care (19, 48, 49). Currently, care for psychiatric disorders, epilepsy and other neurological disorders often occurs in silos, even at the same health facility, and also varies from urban to rural areas. For example, epilepsy may be treated by mental health care providers in rural areas, but by physical health care providers in urban areas. Substance use disorders are also treated in both the health sector and in social development. While the policy is to develop comprehensive care at a primary care level, this is not yet fully realized. Importantly, in South Africa, many people use traditional medicines for MNS disorders, often before or instead of seeking conventional medical treatment (50, 51).

With respect to human resources, there is a substantial mental health workforce shortage, with 1.2 psychiatrists and 7.5 psychiatric nurses per 100,000 people, nearly 10 times less than many high-income countries. South Africa's mental health professionals are concentrated in urban locations, with some rural provinces having one or no psychiatrist, leading to great disparities in care (19, 48).

Review methods

For this narrative overview, we searched Google Scholar and MEDLINE (using PubMed and Ovid) for articles written in English on the economic burden of MNS disorders and for costing data on mental health interventions in South Africa and other sub-Saharan African countries. The narrative overview strategy was selected because it facilitates outlining an area of research that has previously not been widely discussed and highlights key theoretical or empirical gaps in the existing knowledge, yet does not fulfil the methodological criteria of a systematic review (52). To ensure that results provided a sufficient and broad overview of the existing knowledge, we used a deductive approach, generating the paper's headings (direct costs, indirect costs, CEAs) then searching for studies that fit those categories. We broadened the search criteria (from South Africa to sub-Saharan Africa) if there was little literature from South Africa, a strategy that would not be appropriate for a systematic review, but was called for in this case because the amount of economic data available from South Africa and whether data from elsewhere in sub-Saharan Africa is generalizable varies widely between different economic themes (i.e. direct costs, indirect costs, cost-effectiveness). We screened the search results for relevant, methodologically rigorous studies and conducted a forward search of the references of many of the relevant results to identify additional studies. Table 3 displays a summary of all of the studies included in the narrative overview.

Language in South Africa's mental health policy focuses on mental health and substance use disorders

Citation: Glob Health Action 2014, **7**: 23431 - http://dx.doi.org/10.3402/gha.v7.23431

Table 3. Summary of articles included in the narrative overview

Economic information available	Number of studies (total: 18)	Themes (number of study in parentheses)	Study setting (number of study in parentheses)
Direct costs	5	Private sector chronic care (2), public sector workforce costs (2), community interventions (1)	South Africa (5)
Indirect costs	4	Income loss from depression (1), severe mental disorders (1), hospital stay for mental disorder (1), and psychological distress (1)	Ghana (1), Kenya (1), Nigeria (1), and South Africa (1)
Cost-effectiveness	9	Cost-effectiveness of interventions for depression (1), epilepsy (1), bipolar mood disorder (1), heavy alcohol use (2), schizophrenia (1), and many mental disorders (2); cost-effectiveness of group psychotherapy (1)	Low- and middle- income regions, including sub-Saharan Africa (7), Nigeria (1), Uganda (1)

(53), but we broadened the definition to MNS disorders to include patients with neurological disorders, such as epilepsy, who are often treated in mental health care facilities alongside those with mood disorders, anxiety, schizophrenia, and substance use disorders. Neurological and substance use disorders are also often co-morbid with mental disorders and share many of the same co-morbidities as mental disorders, such as HIV and other NCDs. Because the mental health system addresses MNS disorders and these disorders are interlinked, an economic analysis would not be complete without attention to the full burden on the system.

Results

Economic burden of MNS disorders

Economic costs due to mental illness are typically divided into direct and indirect costs (12, 54). Individuals, governments, health insurers or other institutions pay direct costs, usually the costs of medical care and services. Indirect costs include funds spent or lost as a result of the condition, including lost productivity for patients and caregivers, unemployment and disability benefits, and legal, penal, or other costs related to a crime. While distinguishing direct from indirect costs is useful, there is no defined reference case for measuring costs and classifying them as direct or indirect. As a result, there is considerable methodological variation between studies.

Direct costs

Two studies have explored the direct costs of private, outpatient chronic disease care in South Africa. One examined the costs of caring for patients who have private health insurance with chronic disease benefits and have been diagnosed with schizophrenia, epilepsy, or bipolar mood disorder (n = 210,664 health insurance beneficiaries receiving treatment for at least one chronic condition—no specific data on number receiving treatment for other chronic mental disorders). In 2001, outpatient medical management (primarily tests, scans, and doctor visits) for 1 year ranged from R875 (USD$88) for an individual with bipolar mood disorder to R1200 (USD$120) for an outpatient with schizophrenia or epilepsy. Medication costs, on average, are much higher, ranging from R4362 (USD$436) for patients with epilepsy to R7287 (USD$729) for those with schizophrenia and R7512 (USD$751) for those with bipolar mood disorder (55). Additionally, based on data from a private sector pharmaceutical group, in 2008, prescriptions for Alzheimer's disease cost an average of R2659 (USD$266) per patient per year (n = 588 patients) (56).

Public sector workforce expenditures in LMICs account for a substantial portion of health care costs (57), and likely a larger portion of mental health care costs because mental health services, particularly those with adequate capacity for psychosocial care, rely less on laboratory tests or tools and more on trained workers than other forms of healthcare (2). The estimated workforce cost of providing integrated adult mental health services for a limited number of priority mental disorders using a task-shifting approach (dedicating and supporting counsellors and community health workers to work in mental health rather than hiring more expensive specialist health mental workers) in primary health care in South Africa was £28,457 per 100,000 population (approximately USD$44,200 or USD$0.44 per person in the population served by the primary health care facility). The staffing costs of scaling-up integrated primary mental health care, and employing a task-shifting approach was cheaper than alternative staffing models to provide comparable care coverage, although the exact cost difference was not specified (58). The staffing costs associated with implementing inpatient and outpatient child and adolescent mental health care ranged from $5.99 per individual in the population to provide care for 15–30% of children and adolescents with mental disorders to $21.50 for care for 100% of children and adolescents with mental disorders (59). A randomized control trial

in South Africa examined the effects of home-visits for recently discharged psychiatric patients (n = 51) that aimed to prepare the family for home care and assessed health status and care over a 1-year period. This intervention reduced re-admission by 31.5% and the number of days spent in hospital in a year by 55.6%, producing a cost saving of R786 (approximately USD$79) per patient (60).

Indirect costs

Indirect costs, i.e. costs to families and households, may contribute to the total economic burden of mental illness more than direct costs (54). Only one study has assessed indirect costs in South Africa, but several have examined these costs elsewhere in sub-Saharan Africa.

In terms of productivity, the presence of severe depression or anxiety was associated with a reduction in personal income of USD$4798 per adult per year in South Africa, resulting in a national loss of USD$3.6 billion annually (61). In contrast, a Nigerian study found the annual impact of a severe mental disorder on productivity was USD$463 per patient, totalling USD$166.2 million annually (62). The disparity in these costs may be due to differences in purchasing power parity (PPP) between countries and different measurements of productivity. Examining the indirect costs of the institutional care system, a Kenyan study estimated lost productivity over the course of a hospital stay for patients and their families, showing that one psychiatric hospital admission resulted in a USD$453 productivity loss (63). Conversely, household survey data from Ghana examined the general population, many of whom lack access to mental health treatment, and showed that psychological distress, measured using the Kessler 10 Psychological Distress Scale (64), was associated with unemployment and lost work time in both formal and informal sectors. Individuals with moderate or severe psychological distress had reductions in productivity of 11.1 and 24.4%, respectively. From these estimates, the researchers calculated that psychological distress in Ghana is associated with an approximately 6.8% GDP loss, or USD$2.7 million annually (65).

Despite research indicating that common mental disorders co-occur with HIV in South Africa (35–37), no studies specify the direct or indirect costs that stem from the impact that MNS disorders have on other chronic conditions, including diabetes, and stroke.

CEAs of mental health interventions

Existing research suggests that standard treatments, including psychotropic medications and various forms of psychotherapy, are effective in LMICs (66), and preliminary cost-effectiveness data suggests that they may also provide good value for money in terms of DALYs averted.

Several studies examine the cost-effectiveness of interventions for depression (67), epilepsy (68), bipolar mood disorder (69), heavy alcohol use (70, 71), and schizophrenia (72) in low-income regions, including sub-Saharan Africa. Using the same methodology, a single study compared the cost-effectiveness of interventions to address all of these disorders in sub-Saharan Africa. By estimating staffing, drug, and patient care costs (inpatient stay, laboratory tests, outpatient visits, medications), the researchers found that national or regional alcohol control policies (USD$117 per DALY averted by increasing taxation by 50%) and treatment for epilepsy or depression in primary care (epilepsy: USD$265 per DALY averted; depression: USD$858 per DALY averted using newer anti-depressants) were most cost-effective, while inpatient care for schizophrenia using newer psychotropic drugs (USD$11,072 per DALY averted) was least cost-effective. Treating schizophrenia (USD$2,748 per DALY averted) and bipolar affective disorder (USD$2,551 per DALY averted) in the community using older psychotropic drugs paired with psychosocial care was more cost-effective than inpatient treatment (schizophrenia: USD$6,816 and bipolar affective disorder: USD$4,874 per DALY averted). In general, treatments administered in community and primary care settings were more cost-effective than those in hospitals. There were substantial differences in cost-effectiveness between sub-Saharan African and South East Asian regions, underscoring the importance of context-specific cost-effectiveness data (73).

Within sub-Saharan Africa, the only country-specific cost-effective analysis for a range of interventions was conducted in Nigeria and found approximate correspondence with the regional data in the rank order of cost-effectiveness of interventions, but differences in the cost-effectiveness ratios for each intervention (26, 73). A CEA of group psychotherapy for individuals with depression in Uganda found that the therapeutic intervention cost $1,150 per quality-adjusted life year added. The authors concluded that this intervention was cost-effective because it cost less than Uganda's per capita GDP (74), the level that the WHO Commission on Macroeconomics and Health suggests as the upper limit for 'highly cost-effective' interventions (75). However, the study's authors acknowledged that that there is no universally recognized definition of 'cost-effectiveness' or criteria for what makes an intervention cost-effective in any given context.

Discussion

This narrative overview examines available costing data on MNS disorders and the cost-effectiveness of treatments, with a focus on South Africa and data relevant

Citation: Glob Health Action 2014, **7**: 23431 - http://dx.doi.org/10.3402/gha.v7.23431

to South Africa. In the public sector, there is data on workforce expenses at a population level for delivery of specific packages of care (55). In the private sector, there is information on the costs of medication and medical management for three severe and chronic mental disorders (58–60). Only one South African study investigated income reduction associated with depression and anxiety (61), and few other studies estimate indirect costs elsewhere in sub-Saharan Africa (62, 63, 65). There are some region-level data on the cost-effectiveness of interventions for the treatment and prevention of MNS disorders in sub-Saharan Africa, but no data specific to South Africa, even though existing analyses suggest variation in costs and cost-effectiveness between countries and regions.

What we can conclude from what we know

The data in this review suggests that indirect costs from foregone income due to MNS disorders are substantial (61–63, 65), and there are cost-effective interventions for addressing them (73). The most cost-effective interventions incorporate mental health care into primary care or community services without the use of specialized workers (58–60, 73). Such integration may be particularly apt in South Africa because of the high and growing prevalence of MNS disorders co-morbid with HIV and likely, with other chronic conditions (3, 17, 35–37). Integrated interventions could improve coverage of a population that is at high risk for mental disorders and already presenting for care, which would maximize the impact and cost-effectiveness of interventions and improve overall health outcomes by increasing adherence to chronic disease treatment regimes.

Prevention interventions that address alcohol consumption by raising taxes, limiting advertising, or reducing alcohol availability by restricting hours of sale or increasing the drinking age have been shown to be highly cost-effective interventions for reducing DALYs lost due to MNS disorders in LMICs (71, 73). The low cost of these interventions suggests that other prevention programs, such as campaigns to reduce prevalence of risk factors for mental disorders (such as child abuse and sexual violence), may also prove cost-effective.

Knowledge gaps

This review reveals four linked knowledge gaps and associated methodological challenges.

First, there is inadequate research on the health and financial burdens of MNS disorders in South Africa and other LMICs. In terms of health burden, much of the existing research examines prevalence; this likely underrepresents the burden of MNS disorders because of underreporting due to stigma and because much of the burden is due to disability and premature mortality from co-morbid conditions or poor lifestyle and self-care (6). For instance, individuals receiving public mental health services in the US died 13–30 years sooner than people in the general population, although the causes of death were similar to those of the general population (76). Premature mortality in people with mental disorders is not well understood in LMICs.

While DALY measurements, in theory, illustrate disability YLD and premature mortality (YLL) more effectively, consideration must be given to how the DALY is constructed and measured. DALY estimations in South Africa use disability weights that are not context-specific. Disability weighting for YLD calculations varies widely by context depending on the impact of a given condition on a person's lifestyle. In order to develop more accurate burden measures, disability weights should be empirically assessed in South Africa and other LMICs, rather than based on regional data or data from different countries.

Even DALYs cannot fully capture the societal costs of mental disorders because they do not take into account the indirect costs, such as lost productivity of patients and carers, household resources spent caring for a sick family member, travel costs for hospital visits, or the negative impact on patients' children who may not receive adequate attention and care. Indirect costs are particularly high for MNS disorders and must be examined alongside disease burden to fully illustrate their effects on patients, their families, and the broader society (77). Furthermore, most direct and indirect cost studies conducted in sub-Saharan Africa examine mental disorders broadly, with few studies differentiating the costs due to particular conditions and none specifically examining costs of neurological and substance use disorders. More disorder-specific cost data are needed to inform decisions about priority setting and investment.

Second, although use of traditional medicine for mental disorders is common in South Africa (50, 51), there are no data on the prevalence of use, motivations for use, costs, or effectiveness of these treatments. A better understanding of why people use traditional medicines and the health effects of a range of traditional practices would provide insight into beliefs about mental health, how traditional healing could complement or be integrated with conventional medicine, and whether any of these practices put patients at risk. Additionally, existing research on traditional treatments for other health conditions suggests that the costs for traditional medicine are high in South Africa, often as high as those for conventional medical care, and are typically borne by the poorest segment of the population (78), yet there is no cost data available on traditional interventions for MNS disorders. Complete, accurate data on traditional medicines and their costs would provide a

more thorough picture of mental health care and could provide a basis for cost-effective, integrated interventions that operate in tandem with existing traditional practices.

Third, effective strategies for integrating mental health services into other parts of South Africa's health system must be designed and tested. Interventions to incorporate mental health into primary care and into care for people living with HIV have shown promise for use in LMICs (79–81), but must be tested in South Africa. Few interventions that integrate attention to MNS disorders into treatment programs for other chronic diseases (82) or blend conventional treatment of MNS disorders with traditional medicines have been implemented and tested in a LMIC. Before interventions can be implemented on a national scale, they must be tested in South Africa and the effectiveness data used to guide scale-up.

Fourth, in concert with evaluation of the effectiveness of interventions, there is a need for more data on cost-effectiveness and the economic impact of a range of interventions and intervention packages. At present, cost-effectiveness research primarily examines specific treatments, rather than care packages, such as coordinated treatment for patients with co-morbid conditions, prevention efforts integrated with primary health care, or cooperation with traditional healers. Future cost-effectiveness studies will need to examine a broader selection of integrated interventions.

National or provincial cost-effectiveness data may differ substantially from global or regional findings and could be important for bringing about changes in funding priorities (26, 73). For instance, regional cost-effectiveness data do not fully account for inefficiencies in South Africa's fragile health system, such as high absenteeism and unfilled posts. Furthermore, there is a need to examine the broader societal benefits, such as gains in productivity and employment and reduction in costs to other parts of the economy, for instance, policing and crime, child protection, or social work services. Cost-effectiveness data can inform choices on resource allocation, and information on economic gains will help with advocacy for mental health services. These types of economic data are particularly important given the context of South Africa's planned implementation of a national health insurance.

Limitations

The analysis and findings in this review must be acknowledged in light of several limitations. A narrative overview was selected rather than a systematic review because there is little economic research related to MNS disorders in sub-Saharan Africa. As a result, there is great need for an introduction to the topic that challenges current thinking, and defines the future research agenda—appropriate goals for a narrative overview (52). Although the authors defined the section headings prior to conducting the search, they took precautions to ensure the presentation of findings was as unbiased as possible, confining their commentary and interpretation to the discussion section and did not generate their arguments until results were drafted. The grey literature was not searched, and the authors did not approach the Ministry of Health to get additional unpublished data. Finally, costing data was converted to a single currency, but was not adjusted to account for inflation to preserve the integrity of the original data.

Conclusion

This narrative overview examines the epidemiological context of MNS disorders in South Africa and reviews what is known about their costs and the cost-effectiveness of their treatments. Existing data suggests that providing mental health services in the context of other health interventions and prevention efforts aimed at limiting alcohol consumption may be most cost-effective. Further research on the costs related to MNS disorders is greatly needed to develop an evidence base to support effective and efficient implementation and advocacy.

Building political will is critical for the implementation of more integrated models of mental health care. Economic data will be one key factor in making a persuasive case and assisting policymakers to make more informed choices about the importance of investment in mental health care and inclusion of mental health in the basket of options for the proposed national health insurance. While this review has put forward a set of potential priorities for researchers to address, further analysis must be conducted in tandem with conversations with policymakers able to introduce changes based on the findings.

Main findings

- South Africa faces a growing burden of mental, neurological, and substance use (MNS) disorders, which are often co-morbid with HIV and other chronic diseases. A considerable mental health treatment gap exists, with significant care shortages in rural areas.
- Indirect costs, primarily from foregone income due to MNS disorders, are substantial in sub-Saharan Africa.
- The most cost-effective treatment interventions in sub-Saharan Africa incorporate mental health care into community-based services. Taxation of alcohol is a "best buy" for prevention.

Citation: Glob Health Action 2014, **7**: 23431 - http://dx.doi.org/10.3402/gha.v7.23431

Key messages for action

- Four policy-relevant knowledge gaps are identified in South Africa:
  - Epidemiological and economic burdens of MNS disorders must be fully understood to inform spending decisions
  - More data on the use, costs, and effectiveness of traditional therapies for MNS disorders are necessary to develop interventions that combine traditional and biomedical care
  - Effective strategies for integrating mental health services into primary care must be designed and tested
  - Context specific data on the cost-effectiveness of integrated intervention models of cares is essential for advocacy and spending choices
- Economic data is critical for advocacy, to develop integrated models of mental health care and will inform choices between competing spending priorities.

Authors' contributions

KH, AS, RW, and HJ developed the concept for the paper. HJ and RW conducted the literature review. HJ drafted the manuscript with assistance from RW and KH. RT, IP, CN, AS, KH, RW, ST and KK reviewed and provided comments on the manuscript.

Ethical issues

There are no ethical concerns with this paper and ethical review board approval was not required as no human subjects were involved.

Acknowledgements

Thanks to Patrizia Favini and Alex K Smith for their assistance in the preparation of this manuscript.

Conflict of interests and funding

None of the authors declare any conflict of interest with the material in this paper.

References

1. 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 2013; 380: 2197–223.
2. Saxena S, Thornicroft G, Knapp M, Whiteford H. Resources for mental health: scarcity, inequity, and inefficiency. Lancet 2007; 370: 878–89.
3. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 2006; 3: e442.
4. Beyenburg S, Mitchell AJ, Schmidt D, Elger CE, Reuber M. Anxiety in patients with epilepsy: systematic review and suggestions for clinical management. Epilepsy Behav 2005; 7: 161–71.
5. Mayosi BM, Lawn JE, van Niekerk A, Bradshaw D, Abdool Karim SS, Coovadia HM. Health in South Africa: changes and challenges since 2009. Lancet 2012; 380: 2029–43.
6. Prince M, Patel V, Saxena S, Maj M, Maselko J, Phillips MR, et al. No health without mental health. Lancet 2007; 370: 859–77.
7. Ramokgopa G. A milestone for mental health in South Africa. Afr J Psychiatry 2012; 15: 379.
8. Summit PitNMH. The Ekurhuleni declaration on mental health—2012. Afr J Psychiatry 2012; 15: 381–3.
9. National health insurance in South Africa. In: DoHSA, ed. 2011, p. 1–59. Available from: http://us-cdn.creamermedia.co.za/assets/articles/attachments/34471_nhi.pdf [cited 18 July 2013].
10. Tomlinson M, Lund C. Why does mental health not get the attention it deserves. PLoS Med 2012; 9: e1001178.
11. Shiffman J, Smith S. Generation of political priority for global health initiatives: a framework and case study of maternal mortality. Lancet 2007; 370: 1370–9.
12. Drummond MF, Sculpher MJ, Torrence GW, O'Brien BJ, Stoddort GL. Methods for the economic evaluation of health care programmes. Oxford, UK: Oxford University Press; 2005.
13. Disease control priorities related to mental, neurological, developmental and substance abuse disorders. Geneva: World Health Organization; 2006. Available from: http://whqlibdoc.who.int/publications/2006/924156332x_eng.pdf [cited 18 July 2013].
14. Clark DM. Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: the IAPT experience. Int Rev Psychiatry 2011; 23: 318–27.
15. Collins PY, Insel TR, Chockalingam A, Daar A, Maddox YT. Grand challenges in global mental health: integration in research, policy, and practice. PLoS Med 2013; 10: e1001434.
16. World Bank (2013). GINI index. Available from: http://data.worldbank.org/ondicators/si.pov.gini [cited 15 October 2013].
17. Mayosi BM, Flisher AJ, Lalloo UG, Sitas F, Tollman SM, Bradshaw D. The burden of non-communicable diseases in South Africa. Lancet 2009; 374: 934–47.
18. World health report 2001: mental health: new understanding, new hope. Geneva, Switzerland: World Health Organization; 2001.
19. World Health Organization. Mental health atlas: 2005. World Health Organization. Available from: http://www.who.int/mental_health/evidence/mhatlas05/en/ [cited 16 July 2013].
20. Williams SL, Williams DR, Stein DJ, Seedat S, Jackson PB, Moomal H. Multiple traumatic events and psychological distress: the South Africa stress and health study. J Trauma Stress 2007; 20: 845–55.
21. Stein DJ, Seedat S, Herman A, Moomal H, Heeringa SG, Kessler RC, et al. Lifetime prevalence of psychiatric disorders in South Africa. Br J Psychiatr 2008; 192: 112–17.
22. Kleintjes S, Flisher A, Fick M, Railoun A, Lund C, Molteno C, et al. The prevalence of mental disorders among children, adolescents and adults in the Western Cape, South Africa. Afr J Psychiatry 2006; 9: 157–60.
23. Havenaar JM, Geerlings MI, Vivian L, Collinson M, Robertson B. Common mental health problems in historically disadvantaged urban and rural communities in South Africa: prevalence and risk factors. Soc Psychiatr Psychiatr Epidemiol 2008; 43: 209–15.
24. Herman AA, Stein DJ, Seedat S, Heeringa SG, Moomal H, Williams DR. The South African Stress and Health (SASH)

study: 12-month and lifetime prevalence of common mental disorders. S Afr Med J 2009; 99: 339–44.
25. Kessler R, Aguilar-Gaxiola S, Alonso J, Chatterji S, Lee S, Ormel J, et al. Special articles. The global burden of mental disorders: an update from the WHO World Mental Health (WMH) surveys. Epidemiol Psychiatr Sci 2009; 18: 23.
26. Gureje O, Chisholm D, Kola L, Lasebikan V, Saxena S. Cost-effectiveness of an essential mental health intervention package in Nigeria. World Psychiatr 2007; 6: 42–8.
27. 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 Psychiatr 2007; 6: 168.
28. Lynskey MT, Strang J. The global burden of drug use and mental disorders. Lancet 2013; 382: 1540–2.
29. Baxter AJ, Patton G, Scott KM, Degenhardt L, Whiteford HA. Global epidemiology of mental disorders: what are we missing? PLoS One 2013; 8: e65514.
30. Seedat S, Stein DJ, Jackson PB, Heeringa SG, Williams DR, Myer L. Life stress and mental disorders in the South African stress and health study. S Afr Med J 2009; 99: 375–82.
31. Norman R, Bradshaw D, Schneider M, Pieterse D, Groenewald P. Revised burden of disease estimates for the comparative risk factor assessment, South Africa 2000. Cape Town: Medical Research Council; 2006.
32. Flisher AJ, Liang H, Laubscher R, Lombard CF. Suicide trends in South Africa, 1968–90. Scand J Publ Health 2004; 32: 411–18.
33. Els C, Boshoff W, Scott C, Strydom W, Joubert G, Van der Ryst E. Psychiatric co-morbidity in South African HIV/AIDS patients. S Afr Med J – Cape Town Medical Association of South Africa 1999; 89: 992–4.
34. Olley BO, Seedat S, Stein DJ. Persistence of psychiatric disorders in a cohort of HIV/AIDS patients in South Africa: a 6-month follow-up study. J Psychosom Res 2006; 61: 479–84.
35. Myer L, Smit J, Roux LL, Parker S, Stein DJ, Seedat S. Common mental disorders among HIV-infected individuals in South Africa: prevalence, predictors, and validation of brief psychiatric rating scales. AIDS Patient Care STDs 2008; 22: 147–58.
36. Freeman M, Nkomo N, Kafaar Z, Kelly K. Factors associated with prevalence of mental disorder in people living with HIV/AIDS in South Africa. AIDS Care 2007; 19: 1201–9.
37. Singh D, Berkman A, Bresnahan M. Seroprevalence and HIV-associated factors among adults with severe mental illness: a vulnerable population. S Afr Med J 2009; 99: 523–7.
38. Cook JA, Cohen MH, Burke J, Grey D, Anastos K, Kirstein L, et al. Effects of depressive symptoms and mental health quality of life on use of highly active antiretroviral therapy among HIV-seropositive women. J Acquir Immune Defic Syndr 2002; 30: 401–9.
39. Epilepsy and HIV—a dangerous combination. Lancet Neurol 2007; 6: 747.
40. Lin EH, Korff MV. Mental disorders among persons with diabetes—results from the World Mental Health Surveys. J Psychosom Res 2008; 65: 571–80.
41. James BO, Omoaregba JO, Eze G, Morakinyo O. Depression among patients with diabetes mellitus in a Nigerian teaching hospital. S Afr J Psychiatr 2010; 16: 61–4.
42. Issa BA, Yussuf AD, Baiyewu O. The association between psychiatric disorders and quality of life of patient with diabetes mellitus. Iranian J Psychiatr 2007; 2: 30–4.
43. Oladiji J, Akinbo S, Aina O, Aiyejusunle C. Risk factors of post-stroke depression among stroke survivors in Lagos, Nigeria. Afr J Psychiatr 2009; 12: 47–51.
44. Nubukpo P, Clément J, Houinato D, Radji A, Grunitzky E, Avodé G, et al. Psychosocial issues in people with epilepsy in Togo and Benin (West Africa) II: quality of life measured using the QOLIE-31 scale. Epilepsy Behav 2004; 5: 728–34.
45. Mbewe EK, Uys LR, Nkwanyana NM, Birbeck GL. A primary healthcare screening tool to identify depression and anxiety disorders among people with epilepsy in Zambia. Epilepsy Behav 2013; 27: 296–300.
46. Flisher AJ, Lund C, Funk M, Banda M, Bhana A, Doku V, et al. Mental health policy development and implementation in four African countries. J Health Psychol 2007; 12: 505–16.
47. World Health Organisation (2011). Mental health atlas 2011. Available from: http://www.who.int/mental_health/publications/mental_health_atlas_2011/en/ [cited 16 July 2013].
48. Lund C, Kleintjes S, Kakuma R, Flisher AJ. Public sector mental health systems in South Africa: inter-provincial comparisons and policy implications. Soc Psychiatr Psychiatr Epidemiol 2010; 45: 393–404.
49. Petersen I, Lund C. Mental health service delivery in South Africa from 2000 to 2010: one step forward, one step back. S Afr Med J 2011; 101: 751–7.
50. Sorsdahl K, Flisher A, Wilson Z, Stein D. Explanatory models of mental disorders and treatment practices among traditional healers in Mpumulanga, South Africa. Afr J Psychiatry 2010; 13: 284–90.
51. Sorsdahl K, Stein DJ, Flisher AJ. Traditional healer attitudes and beliefs regarding referral of the mentally ill to Western doctors in South Africa. Transcult Psychiatr 2010; 47: 591–609.
52. Green BN, Johnson CD, Adams A. Writing narrative literature reviews for peer-reviewed journals: secrets of the trade. Journal of Chiropractic Medicine 2006; 5: 101–17.
53. South African Department of Health (2012). Mental health policy framework for South Africa and strategic plan 2014–2020 (final draft) Pretoria. In: DOHS SA, ed. Government Printer; 2012, pp. 1–48.
54. Hu TW. Perspectives: an international review of the national cost estimates of mental illness, 1990–2003. J Ment Health Pol Econ 2006; 9: 3.
55. McLeod H, Rothberg A, Pels L, Eekhout S, Mubangizi DB, Fish T. The costing of the proposed chronic disease list benefits in South African Medical Schemes in 2001. Centre for Actuarial Research, University of Cape Town; 2002. Available from: http://www.commerce.uct.ac.za/Research_Units/CARE/RESEARCH/Papers/Chronic Disease List Report.pdf [cited 18 July 2013].
56. Truter I. Prescribing of drugs for Alzheimer's disease: a South African database analysis. Int Psychogeriatr 2010; 22: 264.
57. Provinces spend 46% of combined capital budgets. Available from: http://www.sanews.gov.za/south-africa/provinces-spend-46-combined-capital-budgets [cited 19 July 2013].
58. Petersen I, Lund C, Bhana A, Flisher AJ. A task shifting approach to primary mental health care for adults in South Africa: human resource requirements and costs for rural settings. Health Pol Plann 2012; 27: 42–51.
59. Lund C, Boyce G, Flisher AJ, Kafaar Z, Dawes A. Scaling up child and adolescent mental health services in South Africa: human resource requirements and costs. J Child Psychol Psychiatry 2009; 50: 1121–30.
60. Gillis L, Koch A, Joyi M. The value and cost-effectiveness of a home-visiting programme for psychiatric patients. S Afr Med J 1990; 77: 309.
61. Lund C, Myer L, Stein DJ, Williams DR, Flisher AJ. Mental illness and lost income among adult South Africans. Soc Psychiatr Psychiatr Epidemiol 2013; 48: 845–51.
62. Esan OB, Kola L, Gureje O. Mental disorders and earnings: results from the Nigerian National Survey of Mental Health and Well-being (NSMHW). J Ment Health Pol Econ 2012; 15: 77.

63. Kirigia J, Sambo L. Cost of mental and behavioural disorders in Kenya. Ann Gen Psychiatr 2003; 2: 7.
64. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand S-LT, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 2002; 32: 959–76.
65. Canavan ME, Sipsma HL, Adhvaryu A, Ofori-Atta A, Jack H, Udry C, et al. Psychological distress in Ghana: associations with employment and lost productivity. Int J Ment Health Syst 2013; 7: 9.
66. Patel V, Araya R, Chatterjee S, Chisholm D, Cohen A, De Silva M, et al. Treatment and prevention of mental disorders in low-income and middle-income countries. Lancet 2007; 370: 991–1005.
67. Chisholm D, Sanderson K, Ayuso-Mateos JL, Saxena S. Reducing the global burden of depression Population-level analysis of intervention cost-effectiveness in 14 world regions. The British Journal of Psychiatry 2004; 184: 393–403.
68. Chisholm D. Cost-effectiveness of first-line antiepileptic drug treatments in the developing world: a population-level analysis. Epilepsia 2005; 46: 751–9.
69. Chisholm D, van Ommeren M, Ayuso-Mateos JL, Saxena S. Cost-effectiveness of clinical interventions for reducing the global burden of bipolar disorder. The British Journal of Psychiatry 2005; 187: 559–67.
70. Chisholm D, Rehm J, Van Ommeren M, Monteiro M. Reducing the global burden of hazardous alcohol use: a comparative costeffectiveness analysis. J Stud Alcohol Drugs 2004; 65: 782.
71. Anderson P, Chisholm D, Fuhr DC. Effectiveness and costeffectiveness of policies and programmes to reduce the harm caused by alcohol. Lancet 2009; 373: 2234–46.
72. Chisholm D, Gureje O, Saldivia S, Villalón Calderón M, Wickremasinghe R, Mendis N, et al. Schizophrenia treatment in the developing world: an interregional and multinational cost-effectiveness analysis. Bull World Health Organ 2008; 86: 542–51.
73. Chisholm D, Saxena S. Cost effectiveness of strategies to combat neuropsychiatric conditions in sub-Saharan Africa and South East Asia: mathematical modelling study. BMJ 2012; 344: e609.
74. Siskind D, Baingana F, Kim J. Cost-effectiveness of group psychotherapy for depression in Uganda. J Ment Health Policy Econ 2008; 11: 127.
75. Choosing interventions that are cost effective. Available from: http://www.who.int/choice/costs/CER_thresholds/en/ [cited 1 October 2013].
76. Colton CW, Manderscheid RW. Congruencies in increased mortality rates, years of potential life lost, and causes of death among public mental health clients in eight states. Prev Chronic Dis 2006; 3: A42.
77. Hyman SE. The diagnosis of mental disorders: the problem of reification. Annu Rev Clin Psychol 2010; 6: 155–79.
78. Nxumalo N, Alaba O, Harris B, Chersich M, Goudge J. Utilization of traditional healers in South Africa and costs to patients: findings from a national household survey. J Publ Health Pol 2011; 32(Suppl 1): S124–36.
79. Blank MB, Hanrahan NP, Fishbein M, Wu ES, Tennille JA, Ten Have TR, et al. A randomized trial of a nursing intervention for HIV disease management among persons with serious mental illness. Psychiatr Serv 2011; 62: 1318–24.
80. Crepaz N, Passin WF, Herbst JH, Rama SM, Malow RM, Purcell DW, et al. Meta-analysis of cognitive-behavioral interventions on HIV-positive persons' mental health and immune functioning. Health Psychol 2008; 27: 4–14.
81. Kaaya S, Eustache E, Lapidos-Salaiz I, Musisi S, Psaros C, Wissow L. Grand challenges: improving HIV treatment outcomes by integrating interventions for co-morbid mental illness. PLoS Med 2013; 10: e1001447.
82. Ngo VK, Rubinstein A, Ganju V, Kanellis P, Loza N, Rabadan-Diehl C, et al. Grand challenges: integrating mental health care into the non-communicable disease agenda. PLoS Med 2013; 10: e1001443.

Global Health Action

PART IV

Public policy, health system, and community actions against illness as platforms for response to NCDs in Tanzania: a narrative review

Emmy Metta[1]*, Beverly Msambichaka[1], Mary Mwangome[1], Daniel J. Nyato[1], Marjolein Dieleman[2], Hinke Haisma[3], Paul Klatser[2] and Eveline Geubbels[1]

[1]Ifakara Health Institute, Dar es Salaam, Tanzania; [2]Royal Tropical Institute, Amsterdam, The Netherlands; [3]Faculty of Spatial Sciences, Population Research Centre, University of Groningen, Groningen, The Netherlands

Background: Most low- and middle- income countries are facing a rise of the burden of non-communicable diseases (NCDs) alongside the persistent burden of infectious diseases. This narrative review aims to provide an inventory of how the existing policy environment, health system, and communities are addressing the NCDs situation in Tanzania and identify gaps for advancing the NCD research and policy agenda.
Methodology: A literature search was performed on PubMed and Google scholar with full text retrieval from HINARI of English language articles published between 2000 and 2012. Documents were read to extract information on what Tanzanian actors were doing that contributed to NCDs prevention, treatment, and control, and a narration was written out of these. Reference lists of all retrieved articles were searched for additional relevant articles. Websites of organizations active in the field of NCDs including the Government of Tanzania and WHO were searched for reports and grey literature.
Results: Lack of a specific and overarching NCD policy has slowed and fragmented the implementation of existing strategies to prevent and control NCDs and their determinants. The health system is not prepared to deal with the rising NCD burden although there are random initiatives to improve this situation. How the community is responding to these emerging conditions is still unknown, and the current health-seeking behavior and perceptions on the risk factors may not favor control of NCDs and their risk factors.
Conclusion and recommendation: There is limited information on the burden and determinants of NCDs to inform the design of an integrative and multisectorial policy. Evidence on effective interventions for NCD services in primary care levels and on community perceptions on NCDs and their care seeking is virtually absent. Research and public health interventions must be anchored in the policy, health system, and community platforms for a holistic response.

Keywords: *NCD*; *policy*; *health system*; *community response*

Responsible Editors: Nawi Ng, Umeå University, Sweden; Barthélémy Kuate Defo, University of Montreal, Canada.

*Correspondence to: Emmy Metta, Ifakara Health Institute, P.O. Box 78373, Dar es Salaam, Tanzania, Email: emetta@ihi.or.tz; emetta2000@yahoo.com

This paper is part of the Special Issue: *Epidemiological Transitions – Beyond Omran's Theory*. More papers from this issue can be found at http://www.globalhealthaction.net

Received: 25 November 2013; Revised: 10 March 2014; Accepted: 13 March 2014; Published: 15 May 2014

Tanzania's population of 44.9 million in 2012 is more than three times the 12.3 million estimated in 1967 (1). The population growth rate is estimated at 2.7% annually with a crude birth rate of 12.9 per 1,000 inhabitants. The fertility rate at 5.4 children per woman in 2012 is a drop from the 6.5 and 6.3 children per woman estimated in 1988 and 2002 respectively, but is still very high and contributes to the rapid increase in population (1). The Tanzanian demographic profile, like that of other low- and middle-income countries, is largely young, with 42.5% of the population being younger than 15 years, 51.9% between 15 and 59 years, and 5.6% aged 60 years or older (1). The increasing life expectancy at birth from 50 years in 1988 to 55 years in 2010 (2) is, however, changing this picture by slowly increasing the middle-aged and older populations. This improved survival has been attributed to improved health standards and hence reduced mortality (3). Eighty percent of Tanzania's population lives in rural areas (1), and the adult literacy level was 71% in 2010 (4). Tanzania's gross national

EM, BM, MM, and DJN claim equal authorship.

Citation: Glob Health Action 2014, **7**: 23439 - http://dx.doi.org/10.3402/gha.v7.23439

income per capita was USD 551 in 2010 (4), making it one of the poorest countries worldwide. Health sector spending amounted to 7% of the national gross domestic product in 2011 (5), with the government contributing to 53% of all health expenditure in 2010/11 fiscal year (6).

Tanzania has for decades struggled with the burden of infectious and deficiency disorders along with poor maternal and child health indicators. Successes have been documented in the form of a steady decline of the maternal mortality ratio from 529 deaths in 1996 to 452 deaths per 100,000 live births in 2010 (4); and the under-5 mortality ratio from 137 in 1978 to 81 deaths per 1,000 live births in 2010 (4). Malaria attributable child mortality decreased from 53 in 1999 to 36 in 2007–2008 (7, 8), a reduction in HIV prevalence among persons aged 15 to 49 years was seen from 7% in 2003–2004 to 5.1% in 2012 (9, 10), and evidence of the impact of antiretroviral treatment on population mortality has been documented (11). Although these efforts are ongoing amidst resource constraints (2), the rising prevalence of degenerative non-communicable diseases (NCDs) is exerting further strains on the meager resources.

In 2010, NCDs were responsible for 27% of all deaths in Tanzania (12), a figure that is comparable to that of neighboring countries like Mozambique (13). As in the rest of the world, cardiovascular diseases, cancers, diabetes, and chronic obstructive respiratory diseases (CORDs) have been highlighted as main contributors to premature mortality (14). In 2012, the prevalence of hypertension in Tanzanian adults of 25–64 years of age was 26% and that of raised fasting blood glucose in the same population was 9.1% (15). Findings of a study conducted in selected rural and urban communities in Tanzania in 2003 found that crude yearly stroke incidences were 95 per 100,000 and 107.9 per 100,000, respectively (16). Despite this shrinking rural–urban difference in burden, it has been found that the availability of NCD diagnostic and management services is twice as much in urban compared to rural areas (17). National representative data on heart diseases, cancers, and CORDs could not be obtained. These NCDs are reported to share four main risk factors: poor dietary habits, excessive alcohol use, tobacco use, and lack of physical exercise (18).

The health system is anchored on 5,416 health facilities, at an average ratio of 1.5 facilities per 10,000 persons, with dispensaries being the point of first contact for the majority of the population. Two-thirds of facilities are government-owned, and 236 of the facilities are of hospital level and above (19). The provider to population ratio in 2012 was 7.1 per 10,000 population including only professional health workers. Efforts to tackle infectious diseases have been targeting the health system, policies, and the community, which are important constituents of any health sector response. This paper is an inventory of how the policy environment, health system, and communities are addressing the NCDs situation. It highlights the ways to adapt these platforms for responses to the increasing burden of NCDs and identifies the knowledge gaps for advancing the research and policy agendas for NCDs in Tanzania.

Methodology

A literature search was performed using a set of comprehensive topic-related search terms. Inclusion criteria were English-written articles on original work conducted in Tanzania. We excluded systematic or narrative reviews, opinion papers, documents including expectant women as participants, and articles on drug evaluations and diagnostics. The search was restricted to publications between 1 January 2000 and 31 December 2012 to ensure that the retrieved articles reflected the current situation and most recent responses. The search was performed on PubMed and Google scholar, with full text articles retrieved from HINARI. We also undertook targeted grey literature search, focusing on large institutions including WHO, NCD interest groups in Tanzania, and documents and reports of the government of Tanzania. Articles selected included primary research articles, evaluation, and situation analyses reports. The references of retrieved articles were manually searched for additional material. The search terms were based on key terms aligned to policy, health care system, and community, the important constituents of a health sector response (20). The terms were combined by Boolean operators 'AND' to narrow the search appropriately and 'OR' to expand it with similar terms. The search strategy included the following strings: 'non-communicable diseases' OR 'non communicable diseases' OR 'NCDs' AND Tanzania; 'non-communicable diseases' OR 'non communicable diseases' OR 'NCDs' AND health services OR health care AND Tanzania; 'non-communicable diseases' OR 'non communicable diseases' OR 'NCDs' AND health seeking OR health-seeking AND Tanzania; 'non-communicable diseases' OR 'non communicable diseases' OR 'NCDs' AND policy AND Tanzania. The search resulted in 10,594 articles as depicted in Fig. 1. Ten thousand, four hundred and seventy-three of these were excluded for reasons including title not focused on NCDs, language other than English, or publication date was out of target. One hundred and twenty-one titles were found to be relevant to our subject area and their abstracts were retrieved and screened to determine if they matched our criteria. Forty-one abstracts were excluded based on content relevance to our topic and study design. From this step, 80 abstracts were identified to be relevant and their full text versions were retrieved from HINARI. Bibliographies of retrieved documents were also searched for relevant papers.

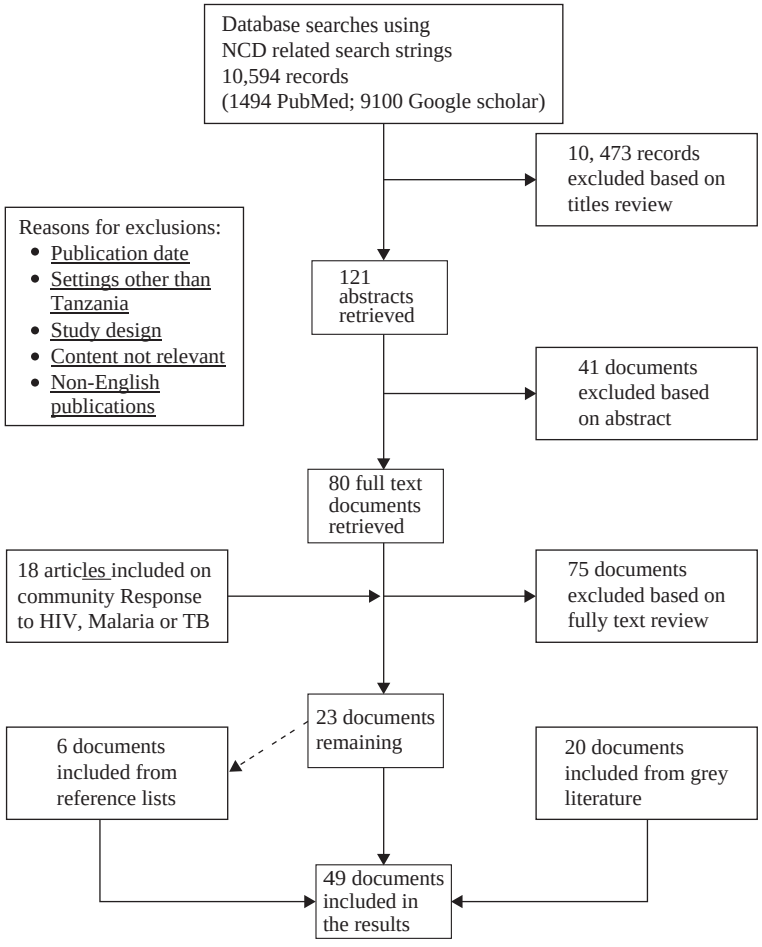

Fig. 1. The process of selecting articles.

Six documents were identified this way, three of which were peer-reviewed articles, and the other three were reports. When our initial search yielded only two papers on community responses to NCD, we expanded our inclusion criteria to include articles on health seeking for HIV, TB, or malaria. We obtained 18 relevant publications based on this expansion.

Analysis was done based on pre-determined themes on NCD policy prevention and control, health system financing, NCD health services, and human recourses of health regarding NCD services and community health-seeking behavior and practices. These themes were predetermined taking into account the review objectives and focused on the main health sector response constituents. Documents were handled manually in that no data management software was used during the analysis. They were read to extract information on what Tanzanian actors were doing that contributed to NCDs prevention, treatment, and control, and a narration was written out of these responses. There were discussions among the first four authors on content of the results and in case of disagreements resolution was by consensus.

Results

This review included 49 documents that were comprised of 26 peer-reviewed articles which were mainly cross-sectional studies published between 2006 and 2012. Three of the 26 articles are included in the policy section, one in the health systems section, and 22 in the community section. The other 23 documents were grey literature mainly from the Tanzanian government and WHO sources, published between 2001 and 2012. Six of these documents were reviewed for the policy section, 17 for the health system, and none was reviewed for the community section. Three of the six documents obtained through searching the reference lists of retrieved articles were peer-reviewed articles, the remaining were grey literature. In these results the NCD-related policy environment is described first and then the preparedness of the health system for the rising NCD burden. This will be followed by an inventory of community responses to NCDs.

Policies and strategies on NCDs and their determinants in Tanzania

The government has established a unit at the Ministry of Health and Social Welfare (MoHSW) to steer formulation

Citation: Glob Health Action 2014, **7**: 23439 - http://dx.doi.org/10.3402/gha.v7.23439

of NCD policies and guidelines (21). It has also included prevention and control of NCDs as strategic objectives of the health sector (7). This National Strategy for NCDs 2008–2013 identifies primary, secondary, and tertiary prevention as important components in addressing NCDs (7). The strategy is integrated and generic and addresses NCDs only from the health sector perspective (7, 12). There was neither an overarching national NCD policy nor evidence of harnessing of non-health sectorial policies in the current response to NCDs despite the multifaceted etiologies of these diseases. Regarding addressing of the four main risk factors of NCDs, by 2011, Tanzania had integrated strategies for alcohol; smoking; physical inactivity; and unhealthy diets, overweight, and obesity (12).

The government signed (in 2004) and approved for implementation (in 2007) the WHO's Framework Convention for Tobacco Control (FCTC) (22). This convention clearly presents both demand and supply reduction provisions for controlling exposure to tobacco, among other provisions (23). Fortunately, some of its demand–reduction provisions such as the introduction of excise taxation on tobacco were already in place (24), but still more actions are needed considering that smoking tobacco costs the nation more than 30 million dollars annually in treatment of tobacco-related cancers (25). In 2011, it was reported by WHO that Tanzania had implemented none of the five tobacco (m)POWER measures to the highest level of achievement (12). Concerning control of the use of alcohol, Tanzania's alcohol policy includes excise taxes on alcohol, a minimum drinking age of 18, a zero tolerance policy for drinking and driving, regulations on alcohol advertising and sponsorships, and restrictions for on- and off-premise sales of alcoholic beverages (26).

The scope of the current Tanzania National Nutrition Strategy states: *the strategy seeks to ensure the nutritional status of all citizens of Tanzania throughout their life cycle* (27). However, the strategy focuses more on women of reproductive age and children under 5 years of age with special emphasize on children less than 2 years of age because malnutrition's most serious and lasting damage occurs during pregnancy and the first two years of life. None of the targets set are related to NCDs or their dietary risk factors.

There was no policy on physical exercise that was found other than a reference to a policy on physical education in schools (28).

Health systems and the response to NCDs in Tanzania

In Tanzania, health services for NCDs are mostly provided from district hospitals to higher level health facilities (29). In 2010, the public health system had only two out of eight NCD-related screening and diagnostic tests nationwide (cervical cancer screening and breast cancer screening) at primary level facilities (21).

In regards to medications, by 2009, the problem of stock-outs for all drugs including for NCDs was persistent and was associated with poor health outcomes, especially for those who could not access services in the private facilities (30). In 2010, 15 out of 17 NCD-related drugs were available in public health system (21).

The Human Resource for Health (HRH) deficit in Tanzania stood at 65% in 2008 (7, 19). With the projected increase of the NCD burden, the government is rethinking its HRH management strategies (7, 19). The immediate strategy involves capacity building of HRH for NCD care in the form of in-service trainings offered by various disease-interest groups (29, 31) such as the Tanzania Diabetes Association which has trained staff from even lower level facilities in the Lake Region (Zachariah Ngoma, personal communication, 26 March 2013). The introduction of NCD prevention and control modules into courses in local training institutions (32) is availing the training to a wider pre-service trainee population, which may help to build the HRH pool of different cadres for the future. The Association of Private Health Facilities in Tanzania (APHFTA) has conducted in-service trainings for diabetes and hypertension care in 18 out of 27 regions of Tanzania (33). In the public health system, comprehensive clinical guidelines have been developed only for diabetes and hypertension (21) but evidence for their distribution, training, and utilization is limited.

Currently, both health facility and population-based information systems reflect limited measurement of NCD-related variables (34, 35). The standards of Health Management Information System (HMIS) in Tanzania are slowly improving with efforts underway to minimize the existing challenges in quality and utilization of health information and to expand information sources such as the cancer registry (36). For population-based NCD information to inform policy and other interventions, the MoHSW resorted to search for evidence on NCDs and their risk factors through projects. The National Institute for Medical Research conducted WHO STEP wise approach to chronic disease risk factor surveillance (WHO STEPs survey), to quantify the risk factors and burden for NCDs (15) nationally. Variables for NCD risk factors have been included in the Tanzania demographic health surveys (DHS) tool (34) and into the MZIMA adult health community cohort (37). There is lack of literature on utilization of even wider existing platforms such as the sentinel Panel of Districts (SPD) (38).

Underfunding of health services is a persistent challenge contributing to a myriad of other malfunctions in the system (39). As a signatory of the Abuja declaration 2001 which required Governments to allocate at least 15%

of its annual budget to health sector, Tanzania allocated 8.9% for the 2011–2012 fiscal year (40). Financing of health care is through, among others, health insurance mechanisms including community health funds that are being rolled out, though uptake is reported to be low (41). Integration of services is one way to deal with NCDs proposed in the Health Sector Strategic Plan III and provides a cost-efficient means of quality service delivery (7). The government has collaborated with other stakeholders in a private hospital in Dar es Salaam whereby NCD services were introduced in the HIV clinic and staff trained in management of co-morbidities. In 2011, this initiative was evaluated and results showed that 15% of its 3,400 patients on antiretroviral therapies had co-morbidities including hypertension, hyperlipidemia, diabetes, and other metabolic disorders (42). Other NCD-related programs that attest to advantages of private–public partnership include the Ocean Road Cancer Institute, and the APHFTA (33, 43).

Communities in preventive health and health care seeking

In public health, community involvement is key to achieving the goals of preventive and promotional health programs (44). In Tanzania the importance of involving communities in effective disease control interventions for various infectious diseases has been well documented (45). This practice was associated with increased access and acceptance of palliative services (46), effectiveness and efficiency of disease control interventions (47), as well as equity, sustainability, and communities' self-reliance (48). Private facility initiatives such as those under APHFTA through their NCD program, involved primary schools and communities around health facilities reached by their program in healthy lifestyle teachings including healthy eating and physical activities (33). The impact of these interventions however, is not yet known.

Factors contributing to development of NCDs include social determinants which are highly linked to complex sociocultural practices and beliefs, making it challenging to effect lifestyle changes (49). For example, fatness was associated with beauty and economic prosperity of the household (50); and excessive alcohol intake was fuelled by cheaply available local brew, and the need for entertainment and relaxation on the part of users (51).

Literature suggests that where illness symptoms were thought to have spiritual etiology, traditional healers became the preferred source of treatment (52–54), because they were perceived to have 'appropriate skills' for managing diseases considered as *out of order* (55).

Socioeconomic determinants have been shown to influence the timeliness of peoples' responses to the utilization of emergency and in-patient services (56) and the initiation and continuation of treatments for chronic illnesses (55), all of which influence patient outcomes (57). Studies revealed that services at health care facilities are only sought when chronic illness symptoms persist or worsen, after using over-the-counter medicines (58, 59). In addition, accessibility to health facilities has been shown to facilitate prompt care seeking (60) and continuity of care for chronic illnesses (61).

Several factors influence adherence to treatments to infectious disease in Tanzania including; people's perceptions of the illness and quality of medicines (62), fear of side effects (63), patient–provider relationships, cultural pressure, cost and availability of medicines (61), and the patient's understanding of the medication schedule (64). Others include poor access of re-filling prescriptions, inadequate nutrition especially for medicines perceived to increase hunger (63), and long hours of absence from home (65). Adherence to medications was reported to also be influenced by gender (65).

Discussion

Our findings have shown that there is no existing NCD policy in Tanzania. For clarity of vision and purpose and coherence of interventions, a national policy is important (66, 67). Reasons for lack of a policy include the lack of evidence to inform such a policy (67). There are several opportunities in available guidelines and strategies for controlling exposure to the major NCD risk factors, implementation of which is partly hampered by lack of an NCD policy. Being multifaceted in their etiology, NCD policies being designed must as well be multisectorial. In Tanzania, there was no evidence of including other sectors such as education and agriculture in policy development for NCD prevention and control. The country could learn from experiences shaping the HIV response where multisectoral approaches have been adopted in prevention and communication activities (68). Regarding prevention of exposure to NCD risk factors, the adoption and implementation of FCTC is a positive step towards tobacco control. However, more effort is needed towards fully implementing the convention through the mPOWER measures.

Although some aspects of alcohol policy are in place, implementation of these, especially in rural areas where local brew is mostly produced and consumed, will be a challenge (68). Addressing diet-related NCDs in the national nutrition strategy without setting targets related to NCDs or their dietary risk factors indicates that overnutrition is not yet prioritized. Whereas there are still no policies focusing on physical activity, it has been proposed that the policies will have to address activity at work places, leisure activities, and means of transport (69). Physical activity policies should also be sensitive to sociocultural differences across Tanzania.

As policies are being developed to tackle the rising NCD burden, the health system in Tanzania must gear up for roles in primary, secondary, and tertiary prevention

Citation: Glob Health Action 2014, **7**: 23439 - http://dx.doi.org/10.3402/gha.v7.23439

of NCDs. Concentration of the NCD services at higher level health facilities leads to late diagnosis and delayed access to appropriate treatments, especially in rural areas (70). Accessibility of NCD services also entails availability of diagnostic and treatment services at the primary level facilities. Lack of screening and diagnostic tests at primary level facilities creates a challenge for NCD care, especially the long-term patient monitoring and management. Whereas the WHO observatory indicates that NCD medication availability was impressive, it is the consistent medical supply of affordable medicines that contributes to favorable health outcomes. A recent national survey in Tanzania showed that, only 9% of rural and 20% of the urban health facilities provided diabetes diagnosis and management services and only half of these had the staff, drugs, diagnostics, and guidelines to actually provide the service on the day of the survey (17). Generally the health care system is not prepared to tackle the rising burden of NCDs (17, 71). The HRH crisis hinders quality preventive, diagnostic, and treatment service provision for NCDs close to the community and requires creative thinking on how to more efficiently use the existing HRH. The pre-service and in-service training efforts, though prudent, do not completely address the shortage of competent cadres in lower level facilities in the country. Creativity including task shifting where lower cadres of HRH are prepared for NCD prevention and management could ease this crisis as has been done elsewhere (72). Availing tools such as clinical guidelines has facilitated task shifting and could be adopted in Tanzania's primary care facilities (72).

The rising burden of NCDs will increase the financial strain on the health system because of their chronic nature; hence, new approaches to minimize costs of interventions must be embraced. Integration ensures that, 'clients receive a continuum of preventive and curative services, according to their needs over time and across different levels of the health system' (73).Various integration models are possible, but little has been documented on pilot projects. The benefits of integration of services that have been observed in the Tanzanian example have been demonstrated in other settings where chronic diseases clinics have been piloted (74). This idea of chronic disease clinics could be replicated in primary care settings where HIV prevention, care, and treatment services are offered.

Improvement of HMIS would positively influence quality of care for NCD patients and overall health system functioning (75). Systematic information concerning public awareness and practices regarding NCD risky behaviors as well as factors that shape risky lifestyle is still limited. Existing population survey platforms such as the demographic surveillance system and SPD could be adapted to fit chronic disease surveillance by introducing outcome measures that realistically measure morbidity burden of NCDs. Mortality outcomes may not capture true burden of NCDs whose onset and course is insidious.

The financial strains caused by NCDs are also experienced by patients and their families. Ensuring effective implementation of the policy on health care fee waivers which exempts patients with chronic diseases from user fee charges will improve accessibility and equity in utilization of NCD services. However, this has implications to the government resources and may necessitate government's re-allocation of its financial resources to the health sector.

Considering that NCDs are insidious in their onset, slow in their progression, and long-term, communities have to be involved in their prevention and control (76). Community involvement has been shown to be effective in other chronic diseases prevention and control efforts (77, 78). This involvement has also been recommended in the Tanzania HSSP III as an objective in the NCD control strategy (7). Since community responses, including lifestyle changes, are shaped by sociocultural aspects and health system factors, effective mitigation of NCDs will therefore have to integrate community-based and individually targeted interventions (79) that are sensitive to variations in gender and cultural norms, for them to be acceptable (80).

Knowing that health services are commonly sought when illnesses persist or worsen after failure of self-medication, interventions to promote prompt health care seeking must be designed. Self-medication may present a challenge when it comes to NCDs, most of which are initially asymptomatic and symptoms indicate worsening of disease intensity. Information is lacking on NCD self-treatment patterns and their motivations, which may hinder the possibility of addressing any undesirable patterns in the community.

As NCD medication use is long-term, it is important to understand how sociocultural aspects shape people's sustained compliance to medicines. Compliance to long-term medication has been studied for infectious chronic diseases (81, 82); however, community perceptions towards infectious disease and NCDs may not be the same. Therefore application of the findings from these studies to NCDs medication may not be valid. Exploration of the aspects that might shape sustained compliance to long-term NCDs medication would inform designing of strategies that foster compliance to long-term treatments.

The knowledge gaps for both research and policy agendas regarding responses to the rising NCDs burden in Tanzania include inadequate information on the epidemiological patterns of NCDs and their determinants across Tanzania and the limited awareness of the local context to inform the design of an integrative and multisectorial policy. Apart from better understanding the epidemiological pattern of NCDs, it will be important to project the implications of the current dual burden on population life expectancy, taking into account not just the

health systems and policy environment, but also the wider demographic and development indicators. There is limited evidence on effective interventions for NCD services in primary care levels and information on community perceptions on NCDs and their care seeking is virtually absent.

Limitations and strength of the review

This review included mainly published documents and thus *omitted* insights from non-published developments. Through the inclusion of only English language documents we might have missed relevant results published in Kiswahili. However, a Kiswahili language research journal does not exist in Tanzania and English is one of the country's official languages. We therefore think that this has not resulted in non-inclusion of formal government-produced documents nor of research publications. The search terms did not include 'chronic disease' because not all chronic diseases are NCDs. We thus may have missed articles that used only 'chronic disease' as a term to describe NCDs. Nevertheless, the holistic nature of the information reviewed provides a broader view of the NCDs situation and the responses in Tanzania. The narrative review approach was chosen rather than a systematic review because it is well suited to present a broad perspective on a newly emerging problem (83, 84).

Conclusions and recommendations

NCDs and their risk factors are largely lifestyle related, making multisectorial responses unavoidable. These conditions are posing a critical challenge to the government, health system, and communities that have to face both communicable and NCDs. There is limited context-relevant information on the burden and determinants of NCDs which may hamper the design of effective interventions, especially for prevention. The existing research platforms such as the health and demographic surveillance system (HDSS), DHS, and SPD can be leveraged to address these NCD knowledge gaps.

Current efforts to address NCDs in Tanzania are fragmented due to lack of a NCD policy. The success of NCD prevention and control requires such a clearly defined NCD policy to provide a roadmap for implementation of multisectoral strategies and plans. For effective interventions targeting NCDs risk factors at local level, empowerment of local government authorities tasked with implementation of existing policies in the country is needed.

The country's health system is not adequately prepared to accommodate the requirements of NCDs. There is an urgent need to design and evaluate low-tech, low-cost interventions for prevention, diagnosis, treatment, and continuity of care that can be scaled up at primary care levels. This may necessitate the adoption of an integrated health care model to address both NCDs and other chronic communicable diseases as a strategy to address the HRH and financing challenges. Also, supporting informed decision making for NCDs at clinical and policy levels requires accurate clinical record keeping and diseases registry maintenance for NCDs and their risk factors.

Information on community awareness and practices in prevention and health care seeking regarding NCDs is limited. The role of the community in NCD development, prevention, and management must be explored by understanding the motivations for health-seeking, self-medication practices, and aspects shaping continuity of care in their particular contexts.

Main findings

- There is limited context-relevant information on the epidemiological patterns of non-communicable diseases (NCDs) and their determinants across Tanzania; this may hamper the design of effective interventions, especially prevention.
- Current efforts to address NCDs in Tanzania are fragmented due to a lack of a NCD policy.
- The county's health system is not adequately prepared to accommodate the requirements of NCDs and the information on community awareness, practices in prevention and health care seeking regarding NCDs is limited.

Key messages for action

- The existing research platforms such as the health and demographic surveillance system (HDSS), demographic health surveys (DHS) and the sentinel panel of districts (SPD) should be leveraged to address the knowledge gaps in epidemiological patterns, in best practices in health care and in community roles in prevention and management of NCDs in Tanzania.
- Considering the multifaceted nature of NCD risk factors and causation, there is a need to formulate a clear, multi-dimensional policy on NCDs prevention and management.
- There is an urgent need to design and evaluate low-tech, low-cost interventions for prevention, diagnosis, treatment and continuity of care that can be scaled up at primary care levels. This may include the adoption of an integrated health care model to address both NCDs and other chronic communicable diseases as a strategy to address the HRH and financing challenges.

Acknowledgements

This review was supported by grants from the Netherlands Organization for International Co-operation in Higher Education and the Ifakara Health Institute, Tanzania.

Conflict of interest and funding

The authors have not received any funding or benefits from industry to conduct this study.

References

1. NBS, OCGS (2013). Population and Housing Census Report. Population Distribution by Administrative Units; Key Findings. Dar es Salaam: NBS and OCGS.
2. URT (2011). Tanzania Country Report on Millennium Development Goals 2010. Dar es Salaam: Tanzania Printers.
3. REPOA. Delivery of Social Services on Mainland Tanzania. Are people satisfied? AFROBAROMETER Briefing Paper 34. Available from: http://www.repoa.or.tz/documents/Afrobrief No34.pdf [cited 5 March 2014].
4. NBS (2011). National Bureau of statistics, Ministry of Finance, June 2011. "Tanzania in figures 2010." Dar es Salaam: NBS.
5. The World Bank. Working for a world free poverty. Available from: http://data.worldbank.org/indicator/SH.XPD.TOTL.ZS [cited 19 November 2013].
6. HSER. Directorate of policy and planning, Ministry of Health and Social Welfare, July 2012. "Health Sector Expenditure Review, 2010/11." Dar es Salaam: Tanzania and health systems 20/20 project, Abt Associates Inc; 2010/11.
7. MoHSW (2008). Health sector strategic plan III for 2009–2015: partnerships for delivering MDGs. Dar es Salaam, Tanzania: M. o. H. S. Welfare.
8. Masanja H, de Savigny D, Smithson P, Schellenberg J, John T, Mbuya C, et al. Child survival gains in Tanzania: analysis of data from demographic and health surveys. Lancet 2008; 371: 1276–83.
9. THMIS (2012). Tanzania HIV/AIDs and Malaria Indicator Survey 2011–12. Dar es Salaam, Tanzania: NBS.
10. Hallett TB, Aberle-Grasse J, Bello G, Boulos L, Cayemittes M, Cheluget B, et al. Declines in HIV prevalence can be associated with changing sexual behaviour in Uganda, urban Kenya, Zimbabwe, and urban Haiti. Sex Trans Infect 2006; 82: i1–i8.
11. Marston M, Michael D, Wringe A, Isingo R, Clark BD, Jonas A, et al. The impact of antiretroviral therapy on adult mortality in rural Tanzania. Trop Med Int Health 2012; 17: e58–e65.
12. WHO. United Republic of Tanzania: NCD country profile 2011. Available from: http://www.who.int/nmh/countries/tza_en.pdf [cited 7 July 2012].
13. Silva-Matos C, Beran D. Non-communicable diseases in Mozambique: risk factors, burden, response and outcomes to date. Glob Health 2012; 8: 37.
14. Atun R, Jaffar S, Nishtar S, Knaul F, Lima Barreto M, Nyirenda M. Improving responsiveness of health systems to non-communicable diseases. Lancet 2013; 20132: 64–71.
15. WHO. WHO Tanzania STEPS Survey-2012 Fact sheet. Available from: http://www.who.int/chp/steps/UR_Tanzania_FactSheet_2012.pdf [cited 12 August 2012].
16. Walker R, Whiting D, Unwin N, Mugusi F, Swai M, Aris E, et al. Stroke incidence in rural and urban Tanzania: a prospective, community-based study. Lancet Neurol 2010; 9: 786–92.
17. MoHSW (2013). Tanzania service availability and readiness assessment (SARA) 2012. Dar es Salaam. Dar es Salaam, Tanzania: Ifakara Health Institute.
18. Nigatu T. Integration of HIV and noncommunicable diseases in health care delivery in low-and middle-income countries. Prev Chronic Dis 2012; 9: 1–3.
19. MoHSW (2008). Human resource for health strategic plan 2008–2013. Dar es Salaam: MoHSW.
20. Pruitt S, Annandale S, Epping-Jordan J, Femandez Diaz JM, Khan M, Kisa A, et al. Innovative care for chronic conditions. Building blocks for action. Geneva, Switzerland: WHO; 2002.
21. Health systems response and capacity. Global Health Observatory. Available from: http://apps.who.int/ghodata/ [cited 19 August 2012].
22. WHO (2010). Report card on the WHO framework convention on tobacco control. Geneva: WHO.
23. WHO (2003). Framework convention on tobacco control. Geneva: World Health Organization, Tobacco Free Initiative.
24. Osoro N, Mpango P, Mwinyimvua H. An analysis of excise taxation in Tanzania. Alexandria, Viginia: EAGER Publications/BHM; 2001.
25. Kagaruki LK. Community-based advocacy opportunities for tobacco control: experience from Tanzania. Glob Health Promot 2010; 17: 41–4.
26. United Republic of Tanzania (the) Socio economic context. Available from: http://www.who.int/substance_abuse/publications/global_alcohol_report/profiles/tza.pdf?ua=1 [cited 22 June 2012].
27. National Nutrition Working Group. National Nutrition Strategy: July 2011/12–June 2015/16. Available from: http://scalingupnutrition.org/wp-content/Tanzania_National-Nutrition-Strategy.pdf [cited 13 November 2012].
28. Mafuniko FM, Pangani IN. Physical education in Tanzanian secondary schools: perceptions towards physical education as an academic discipline. NUE J Int Educ Coop 2008; 3: 51–61.
29. Ramaiya K. Setting up diabetes clinics in Tanzania. Pract Diabetes Int 2006; 23: 339–40.
30. MoHSW: Ministry of Health & Social welfare TZ (2009). In-depth assessment of the medicines supply system in Tanzania. Dar es Salaam: MoHSW.
31. WHO (2009). WHO country cooperation strategy 2010–2015, Tanzania. Brazzaville, Republic of Congo: WHO Regional Office for Africa.
32. Diabetes Management Course. Available from: http://www.muhas.ac.tz/Advertisements/short courses/Diabetes course advertisment AUGUST 2012.pdf [cited 15 November 2012].
33. Association of Private Health Facilities in Tanzania (2011). Non communicable diseases programme. Available from: http://www.aphfta.org/index.php?option=com_content&view=article&id=115&Itemid=154 [cited 25 September 2012].
34. TDHS (2010). National Bureau of Statistics (NBS) [Tanzania] and ICF Macro. 2011.Tanzania Demographic and Health Survey 2010. Dar es Salaam, Tanzania: NBS and ICF Macro.
35. TFNC (2010). National nutrition strategy JULY 2011/12 – JUNE 2015/16. ed. M. o. H. a. S. Welfare. Dar es Salaam, Tanzania.
36. Mwakigonja AR, Loon KV. Proposal to strengthen health information system. ed. M. o. H. a. S. Welfare. Dar es Salaam: Ifakara Health Research & Development Center, UDSM, MoH&SW, UiO; 2008.
37. Eveline Geubbels. The Ifakara MZIMA cohort. An open adult health community cohort on chronic diseases. Available from: http://www.ihi.or.tz/a/ihi.or.tz/ihi-main-site/projects/mzima [cited 22 November 2012].
38. Sentinel Panel of Districts. A new platform for health monitoring and evaluation in Tanzania Spotlight. Issue 8, October 2011. Available from: http://ihi.eprints.org/1831/1/IHI_Spotlight_-_SPD_Final_Vol_8.pdf [cited 5 March 2012].
39. HERA (2006). District health services delivery in Tanzania: where are we in terms of quantity and quality of health care provision? Belgium: HERA.

40. MOF. Government Budget for Financial year 2011/2012 Citizens Budget Edition. Available from: http://www.opengov.go.tz/files/publications/attachments/CITIZEN_ENGLISH_2011_12_FINAL_en_sw.pdf [cited 24 September 2012].
41. Musau S, Chee G, Patsika R, Malangalila E, Chitama D, Praag EV, et al. Tanzania health system assessment 2010. In: Bethesda M, ed. Bethesda, Maryland: Health systems 20/20 project; 2011. p. 1–110.
42. MSH. Non-communicable and chronic ciseases: a health systems approach. Available from: http://projects.msh.org/resource-center/fact-sheets/upload/NCD-factsheet.pdf [cited 2 June 2012].
43. Ocean Road Cancer Institute. Available from: http://www.uicc.org/membership/ocean-road-cancer-institute [cited 6 October 2012].
44. Mlozi M, Shayo E, Senkoro K, Mayala B, Rumisha S, Mutayoba B, et al. Participatory involvement of farming communities and public sectors in determining malaria control strategies in Mvomero District, Tanzania. Tanzan Health Res Bull 2006; 8: 134–40.
45. Egwaga S, Mkopi A, Range N, Haag-Arbenz V, Baraka A, Grewal P, et al. Patient-centred tuberculosis treatment delivery under programmatic conditions in Tanzania: a cohort study. BMC Med 2009; 7: 80.
46. Nanney E, Smith S, Hartwig K, Mmbando P. Scaling up palliative care services in rural Tanzania. J Pain Symptom Manage 2010; 40: 15–18.
47. Kisinza W, Kisoka W, Mutalemwa P, Njau J, Tenu F, Nkya T, et al. Community directed interventions for malaria, tuberculosis and vitamin A in onchocerciasis endemic districts of Tanzania. Tanzan J Health Res 2008; 10: 232–9.
48. Mkumbo K, Schaalma H, Kaaya S, Leerlooijer J, Mbwambo J, Kilonzo G. The application of Intervention Mapping in developing and implementing school-based sexuality and HIV/AIDS education in a developing country context: The case of Tanzania. Scand J Public Health 2009; 37: 28–36.
49. Shayo GA, Mugusi FM. Prevalence of obesity and associated risk factors among adults in Kinondoni municipal district, Dar es Salaam Tanzania. BMC Public Health 2011; 11: 365.
50. Maletnlema T. A Tanzanian perspective on the nutrition transition and its implications for health. Public Health Nutr 2002; 5: 163–8.
51. Mbatia J, Jenkins R, Singleton N, White B. Prevalence of alcohol consumption and hazardous drinking, tobacco and drug use in urban Tanzania, and their associated risk factors. Int J Environ Res Public Health 2009; 6: 1991–2006.
52. Kamat VR. Dying under the bird's shadow: narrative representations of Degedege and child survival among the Zaramo of Tanzania. Med Anthropol Q 2008; 22: 67–93.
53. Foster D, Vilendrer S. Two treatments, one disease: childhood malaria management in Tanga, Tanzania. Malar J 2009; 8: 240.
54. Comoro C, Nsimba S, Warsame M, Tomson G. Local understanding, perceptions and reported practices of mothers/guardians and health workers on childhood malaria in a Tanzanian district—implications for malaria control. Acta Tropica 2003; 87: 305–13.
55. Simpson K. Diabetes in Tanzania: insulin supply and availability. J R Coll Physicians Edinb 2003; 33: 181–201.
56. Ferry GA, Dickson SR, Mbaruku G, Freedman LP, Kruk ME. Equity of inpatient health care in rural Tanzania: a population- and facility-based survey. Int J Equity Health 2012; 11: 7.
57. Simba DO, Kakoko DC, Warsame M, Premji Z, Gomes MF, Tomson G, et al. Research understanding caretakers' dilemma in deciding whether or not to adhere with referral advice after pre-referral treatment with rectal artesunate. Malaria Journal 2010; 9: 1–9.
58. Mangesho P, Shayo E, Makunde W, Keto G, Mandara C, Kamugisha M, et al. Community knowledge, attitudes and practices towards tuberculosis and its treatment in Mpwapwa District, central Tanzania. Tanzan J Health Res 2007; 9: 38–43.
59. Haram L. Assessment of health-care seeking behaviour: the case of co-infection of TB and HIV/AIDS in Temeke, Tanzania. Trondheim, Norway: Nowegiean University of Science and Technology (NTNU); 2008.
60. Obrist B, Iteba N, Lengeler C, Makemba A, Mshana C, Nathan R, et al. Access to health care in contexts of livelihood insecurity: a framework for analysis and action. PLoS Med 2007; 4: e308.
61. Kolling M, Winkley K, von Deden M. Research "For someone who's rich, it's not a problem." Insights from Tanzania on diabetes health-seeking and medical pluralism among Dar es Salaam's urban poor. Globalization and Health 2010; 6: 8. DOI: 10.1186/1744-8603-6-8.
62. Dillip A, Hetzel M, Gosoniu D, Kessy F, Lengeler C, Mayumana I, et al. Socio-cultural factors explaining timely and appropriate use of health facilities for degedege in south-eastern Tanzania. Malar J 2009; 8: 144.
63. Roura M, Busza J, Wringe A, Mbata D, Urassa M, Zaba B. Barriers to sustaining antiretroviral treatment in Kisesa, Tanzania: a follow-up study to understand attrition from the antiretroviral program. AIDS Patient Care STDs 2009; 23: 203–10.
64. Kabanywanyi AM, Lengeler C, Kasim P, King'eng'ena S, Schlienger R, Mulure N, et al. Adherence to and acceptability of artemether-lumefantrine as first-line anti-malarial treatment: evidence from a rural community in Tanzania. Malar J 2010; 9: 48.
65. Watt MH, Maman S, Earp JA, Eng E, Setel PW, Golin CE, et al. "It's all the time in my mind": facilitators of adherence to antiretroviral therapy in a Tanzanian setting. Soc Sci Med 2009; 68: 1793–800.
66. Mendis S, Fuster V. National policies and strategies for noncommunicable diseases. Nat Rev Cardiol 2009; 6: 723–7.
67. Mfinanga SG, Kivuyo SL, Ezekiel L, Ngadaya E, Mghamba J, Ramaiya K. Public health concern along side with global initiative on the priority action for "silent uprising epidemic" on non-communicable diseases in Tanzania. Tanzan J Health Res 2012; 13: 1–6.
68. National Multi-sectoral HIV Prevention Strategy-NACP. Available from: http://www.nacp.go.tz/documents/Prevention Strategy.pdf [cited 24 January 2014].
69. Armstrong T, Bull F. Development of the world health organization global physical activity questionnaire (GPAQ). J Public Health 2006; 14: 66–70.
70. Ramaiya K. Personal view: Tanzania and diabetes—a model for developing countries? Br Med J 2005; 330: 679.
71. Mbanya JCN, Motala AA, Sobngwi E, Assah FK, Enoru ST. Diabetes in sub-Saharan Africa. Lancet 2010; 375: 2254–66.
72. Lekoubou A, Awah P, Fezeu L, Sobngwi E, Kengne AP. Hypertension, diabetes mellitus and task shifting in their management in sub-Saharan Africa. Int J Environ Res Public Health 2010; 7: 353–63.
73. WHO: World Health Organization (2008). Integrated health services – what and why. Technical Brief. Geneva: WHO.
74. Janssens B, Van Damme W, Raleigh B, Gupta J, Khem S, Soy Ty K, et al. Offering integrated care for HIV/AIDS, diabetes and hypertension within chronic disease clinics in Cambodia. Bull World Health Organ 2007; 85: 880–5.
75. Epping-Jordan J, Pruitt S, Bengoa R, Wagner E. Improving the quality of health care for chronic conditions. Qual Saf Health Care 2004; 13: 299–305.
76. Nissinen A, Berrios X, Puska P. Community-based noncommunicable disease interventions: lessons from developed coun-

tries for developing ones. Bull World Health Organ 2001; 79: 963–70.
77. Dawad S, Jobson G. Community-based rehabilitation programme as a model for task-shifting. Disabil Rehabil 2011; 33: 1997–2005.
78. Hatcher A, de Wet J, Bonell CP, Strange V, Phetla G, Proynk PM, et al. Promoting critical consciousness and social mobilization in HIV/AIDS programmes: lessons and curricular tools from a South African intervention. Health Educ Res 2011; 26: 542–55.
79. Krishnan A, Ekowati R, Baridalyne N, Kusumawardani N, Kapoor S, Leowski J. Evaluation of community-based interventions for non-communicable diseases: experiences from India and Indonesia. Health Promot Int 2011; 26: 276–89.
80. Jagoe K, Edwards R, Mugusi F, Whiting D, Unwin N. Tobacco smoking in Tanzania, East Africa: population based smoking prevalence using expired alveolar carbon monoxide as a validation tool. Tob Control 2002; 11: 210–14.
81. Gonzalez JS, Penedo FJ, Antoni MH, Durán RE, McPherson-Baker S, Ironson G, et al. Social support, positive states of mind, and HIV treatment adherence in men and women living with HIV/AIDS. Health Psychol 2004; 23: 413.
82. Munro SA, Lewin SA, Smith HJ, Engel ME, Fretheim A, Volmink J. Patient adherence to tuberculosis treatment: a systematic review of qualitative research. PLoS Med 2007; 4: e238.
83. Murphy CM. Writing an effective review article. J Med Toxicol 2012; 8: 89–90.
84. Green BN, Johnson CD, Adams A. Writing narrative literature reviews for peer-reviewed journals: secrets of the trade. J Chiropr Med 2006; 5: 101–17.

Global Health Action

PART IV

Essential evidence for guiding health system priorities and policies: anticipating epidemiological transition in Africa

Peter Byass[1,2*#], Don de Savigny[3,4] and Alan D. Lopez[5]

[1]Umeå Centre for Global Health Research, Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden; [2]MRC/Wits Rural Public Health and Health Transitions Research Unit, School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; [3]Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Basel, Switzerland; [4]University of Basel, Basel, Switzerland; [5]Melbourne School of Population and Global Health, University of Melbourne, Carlton, Australia

Background: Despite indications that infection-related mortality in sub-Saharan Africa may be decreasing and the burden of non-communicable diseases increasing, the overwhelming reality is that health information systems across most of sub-Saharan Africa remain too weak to track epidemiological transition in a meaningful and effective way.

Proposals: We propose a minimum dataset as the basis of a functional health information system in countries where health information is lacking. This would involve continuous monitoring of cause-specific mortality through routine civil registration, regular documentation of exposure to leading risk factors, and monitoring effective coverage of key preventive and curative interventions in the health sector. Consideration must be given as to how these minimum data requirements can be effectively integrated within national health information systems, what methods and tools are needed, and ensuring that ethical and political issues are addressed. A more strategic approach to health information systems in sub-Saharan African countries, along these lines, is essential if epidemiological changes are to be tracked effectively for the benefit of local health planners and policy makers.

Conclusion: African countries have a unique opportunity to capitalize on modern information and communications technology in order to achieve this. Methodological standards need to be established and political momentum fostered so that the African continent's health status can be reliably tracked. This will greatly strengthen the evidence base for health policies and facilitate the effective delivery of services.

Keywords: *health services*; *epidemiological transition*; *health information*; *sub-Saharan Africa*; *health policy*

Responsible Editors: Nawi Ng, Umeå University, Sweden; Barthélémy Kuate Defo, University of Montreal, Canada.

*Correspondence to: Peter Byass, Umeå Centre for Global Health Research, Department of Public Health and Clinical Medicine, Umeå University, SE-90187, Umeå, Sweden, Email: peter.byass@epiph.umu.se

This paper is part of the Special Issue: *Epidemiological Transitions – Beyond Omran's Theory*. More papers from this issue can be found at http://www.globalhealthaction.net

Received: 14 November 2013; Revised: 24 February 2014; Accepted: 27 February 2014; Published: 15 May 2014

There is increasing evidence that, in many sub-Saharan African populations, death rates from major communicable diseases are declining, especially in childhood. As a result, countries are likely to advance through a process of epidemiological transition toward a greater burden of non-communicable diseases (NCDs), while still bearing a heavy communicable disease burden (1–5). It is much less evident that sub-Saharan African health systems or their health information subsystems have adequate processes in place to adapt their systems and policies accordingly (6). Epidemiological transitions involve changes in patterns of births and deaths, and particularly in causes of death, and are inevitably accompanied by health transitions, in terms of the risks and diseases experienced, and changing patterns of health care needs. Measured trends in all-cause mortality suggest that patterns of morbidity and mortality are shifting, both in terms of cause and age distribution, with ensuing changes in therapeutic needs and demands (the most obvious example in sub-Saharan Africa [SSA] being anti-retroviral therapy against HIV/AIDS). Changing disease patterns are tending to increase

[#]PB is Deputy Editor of Global Health Action, but played no part in the editorial process for this paper.

Citation: Glob Health Action 2014, **7**: 23359 - http://dx.doi.org/10.3402/gha.v7.23359

prevalence compared with incidence for some key diseases (7, 8). Changing patterns of risk factors – whether in terms of vector exposure risks for infectious diseases or factors such as tobacco and alcohol consumption for NCDs – constitute a further critical component (9, 10). Long neglected health issues such as mental health are now increasingly seen as needing a health systems response (11). In low-income countries under-five mortality is decreasing at an impressive rate, albeit slower in SSA (12). These changes are occurring against a background of extremely scanty and often dubious data about what is actually happening (13). There is now, more than ever, a need to proactively update strategies for essential health data in SSA in order to increase the visibility of the continent's current and future population health trends and priorities.

It is unrealistic to expect that all SSA countries will achieve adequately high performance of national health information systems to global standards of timeliness, completeness, quality, and data use over the next 10–20 years. Thus, a transitional prioritised approach to improving the availability of critical health information in the short-term needs to be considered that is relevant to the essential policy actions that SSA countries must take now as the epidemiological transition unfolds. Currently, estimates of Africa's population health parameters tend to be made at the global level (14, 15), using such data as may be available as inputs to increasingly sophisticated models. The consequence of this is that national estimates for SSA countries tend to be a by-product of global estimates rather than contributing to them. This is completely counter to the principle of helping countries to strengthen their health systems through better health intelligence resulting from better information systems. Ideally, we need to move toward a model where a within-country data cycle becomes the normative source of information used locally and then fed into global estimates (Fig. 1), rather than *vice versa.* A typology of within-country data sources that might contribute to this is shown in Table 1.

The concept of health transition necessarily implies that not only health status but also rates of change need to be monitored and hence within-country data cycles need to be established on an on-going basis. This is a prerequisite for ensuring a continuous supply of timely health information, rather than relying on a series of random, cross-sectional snapshots.

In this paper, we set out the principles of a minimum essential dataset that we believe all African countries should and could establish in order to guide health systems and policies through this epidemiological and health transition that is looming, and for which most sub-Saharan African countries are wholly unprepared. This paper is not intended to be an implementation guide but could inform the design of in-country data systems.

We propose that the essential components of such a minimum dataset are as follows:

1. continuous, reliable, unbiased *documentation of age- and sex-specific mortality, including the major causes of deaths* in the population (through routine civil registration with vital statistics, supplemented with sentinel or sample mortality surveillance systems with verbal autopsy [VA], where necessary);
2. biennial *documentation of exposure to the top 10 major risk factors for the leading causes of mortality* by age and sex (through population-based national household surveys); and

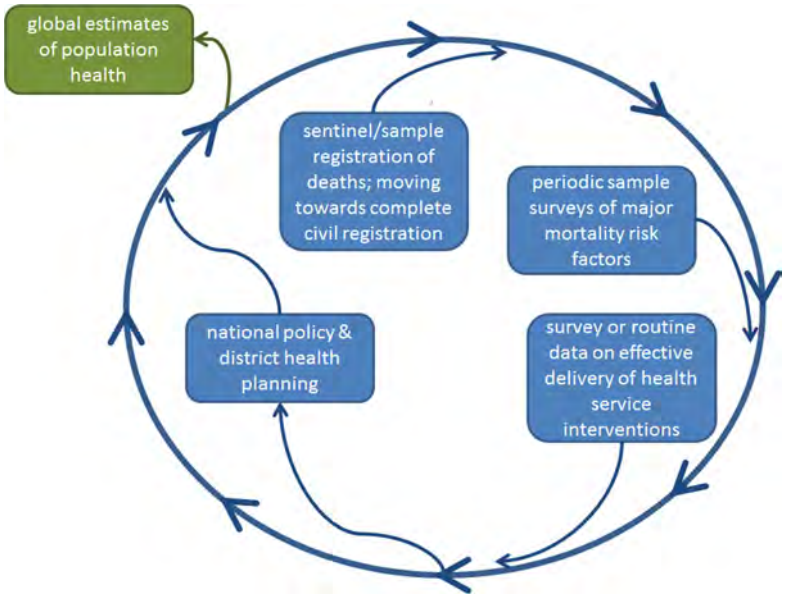

Fig. 1. The concept of an in-country data cycle, also able to feed into global data.

Table 1. Typology of selected population and health facility level data sources potentially contributing to the specific national health information needs as described in this paper

Level	Model	Sample	Approach	Examples
National	National census	All	Complete cross-section	Most countries
	Civil registration with vital statistics	All	Complete longitudinal	Industrialized countries
	Sample registration or sentinel districts	1–2% of population	Longitudinal sample	China, India, Tanzania
	Cluster surveys	Cluster sample size	Repeatable cross-section	DHS surveys, WHO-SAGE
	Fixed panel surveys	Cohort sample size	Longitudinal cohort	Millennium Cohort Study
	Health facility surveys	All or sample of facilities	Self-selected group	Service availability and quality
Regional/Provincial	Complete population	All	Complete longitudinal	Universal registration
	Cluster surveys	Cluster sample size	Cross-section	Intervention coverage
	Individual surveillance	Defined area population	Complete in defined area	INDEPTH centres
District/local area	One-off or annual surveys	Survey sample size	Cross-sectional	Ad-hoc enquiries and district situation analyses
	Health facility surveys	All or sample of facilities	Self-selected group	Service availability, quality and use (for coverage numerators)
	Specific research	Context dependent	Specific issues of interest	Academic studies

3. annual *documentation of the district-level coverage of key preventive and curative health interventions* for these major causes and risk factors for district planning purposes supplemented through a mix of periodic health facility assessments and surveys to determine full effective coverage of interventions for national policy purposes.

Some critically important strategic questions which need to be addressed in this context are:

1. *how* can these data sources be cost-effectively integrated within national health information systems to reliably describe the epidemiological transition dynamics for national populations?
2. *what* are the methodological implications for upgrading national health information systems to reliably measure epidemiological transition?
3. *which* ethical and political issues might drive long-term improvements in national health information in this direction?

We discuss each of these issues below.

1. **Documentation of age- and sex-specific mortality, including the major causes of deaths.**
Complete and timely registration of all births and deaths at national level, including medical certification of cause of death, and collection of other key information about each birth and death, such as age of mother or age at death - for an entire population - is the optimal solution for monitoring epidemiological transition, as pioneered in Scandinavian countries and progressively implemented in other nations along with socioeconomic development (16). A circular issue arises, however, in that epidemiological transition normally accompanies other processes of population development, including the improvement of health information systems, and it is important not to misinterpret apparent longitudinal changes in population health that may actually reflect developments in its monitoring. It is equally clear that near-complete individual registration of births and deaths is not going to be widely implemented in SSA countries in the near future, for logistic and economic reasons, although this must be the primary goal, as elsewhere, of national health information development strategies (17). Therefore, it is important to consider what the 'best-buy' strategies for health data in SSA might be for the immediate decades, particularly emphasizing the need to reliably measure changes in health patterns over time. Representativeness of sub-national data is a crucial but difficult concept in this context. For reasons that are not always obvious, the frequent default assumption is that sub-national data are unrepresentative. This partly arises because it is very hard to demonstrate that any restricted set of data accurately represents a wider but unknown context. However, empirical evidence suggests that data may often actually be more generalizable than is thought

Citation: Glob Health Action 2014, **7**: 23359 - http://dx.doi.org/10.3402/gha.v7.23359

to be the case. For example, in 1925, when Sweden was in many ways similar to a modern-day low- or middle-income country (LMIC), around 80% of counties had health indices closely comparable to national levels, meaning that most single counties chosen at random could each have adequately represented this relatively small country (18). This is an important consideration in the medium-term, when universal registration is not likely to be implemented in SSA countries. National estimates of proportional mortality show major similarities among neighbouring countries. In Tanzania, district burden of disease data have been used to inform district health priorities and resource allocation for other districts in nearby administrative regions (19).

The overall choice of data sources therefore needs to combine a variety of sources, each with different strengths, which are complementary, and also each differently viable in a particular national context, but particularly considering the need to capture change over time. Factors such as the size and diversity of a country, the nature and coverage of its health system, local costs of relevant items such as wages, travel, communications, etc., and local history of more and less successful data collection strategies will also be important (20).

A relevant resource in terms of practical steps toward tracking epidemiological transition has been provided by AusAID's Health Information Systems Knowledge Hub (http://www.uq.edu.au/hishub/) at the University of Queensland (17). Although this was designed primarily for Asian and Pacific countries, the principles translate well to SSA.

Data sources for the first dataset on causes of death are several. Ideally, physician-certified causes of death incorporated into a civil registration system is the standard source of vital statistics to which all countries should aspire to develop and maintain. However, attaining adequate coverage of all deaths at national level (at least 90%) with sufficient quality of cause of death coding has proved elusive for low- and many middle-income countries and could potentially require decades to achieve, without concerted effort and resources. In the meantime, as civil registration and vital statistics (CRVS) systems slowly develop, the WHO and the former Health Metrics Network have recommended interim data sources (21). These are sentinel (urban & rural) demographic surveillance sites as a minimum, or where possible, more statistically representative sample registration sites, both with VA on all deaths in the sentinel or sample populations (17). These can be designed, funded, and implemented within 1–2 years and will produce useful longitudinal data thereafter on trends and dynamics in mortality by cause. Moreover, they will strengthen capacity both to produce and use cause of death data at country and sub-national levels. There is now extensive experience in implementing such sentinel mortality surveillance systems - Health and Demographic Surveillance Sites (HDSS) with VA - and growing experience in implementing sample registration with VA (SAVVY) (17, 22). VA methods for low- and middle-income countries are becoming increasingly standardized, adapted and simplified through machine coding of causes of death (23, 24).

2. **Documentation of exposure to the top 10 major risk factors of mortality.**
   Data sources for the second essential dataset on risk factor exposure can use standard adapted survey instruments for each risk factor (smoking, nutrition, high blood pressure, obesity, HIV sero-status, solid fuel smoke exposure, etc.) (25). However, these are rarely assembled into an omnibus national sample survey along the lines of the WHO STEPS (26). The Health Metrics Network and the Household Survey Network have been promoting greater integration and more strategic scheduling of national household surveys. Where this is done, the strategies should be reviewed to ensure that minimum indicators to inform epidemiological and health transition are included.
3. **Documentation of the effective coverage of key preventive and curative health interventions for these causes and risk factors.**
   Data sources for the third dataset on effective coverage of key health interventions targeting the top causes of death and risk factors are the least developed. Indicators of health service and intervention coverage and quality have the most sub-national (District) heterogeneity and can differ widely between neighbouring districts (27–29). District Health Information Systems have a long tradition of counting cases served (numerators). But there is a dearth of simple methodology that District Health Managers can apply to understand the actual annual reach and coverage of their own services, mainly due to missing denominators for many conditions. There is much room and need to innovate here by combining epidemiological and demographic information to provide districts with annual estimates of denominators (that is, likely cases of a particular disease to treat, size of target population for preventive services such as immunisation, etc.) to allow them to report and manage local coverage. In the meantime, coverage is usually only available from periodic cross-sectional national household or health facility surveys. This is one of the largest health information system weaknesses today (30). It is impossible strategically to manage

district health systems without knowing effective coverage of key services (31). The key importance of this data source is it allows national health policy stakeholders to determine the alignment of their services and policies with population health needs.

How can different sources of data be effectively integrated to yield a national picture of health transition?

Making connections between different data sources is always a major problem, particularly in countries where there is no universal unique personal identifier system. In countries with unique personal identification, it is possible in principle (subject to suitable ethical and confidentiality safeguards) to cross-link data of different types that relate to the same individuals (e.g. population-based data and health facility data) (32). However, that is not normally the case for SSA populations, and consequently there are very real difficulties in relating different data sources. SSA health facilities also tend to operate in very flexible catchment areas, with people often opting to consult out-of-area facilities because of perceived differences in issues such as quality of care and stigma, making links between population and facility data difficult. Consequently, connections between different data sources are very unlikely to be possible at the individual level, and even on a geographical basis (e.g. within districts) may be problematic.

The temporal component of any data intended to be used for understanding transition further complicates the issue. The example of the widely implemented Demographic and Health Survey (DHS) (33) and other national cluster sample surveys such as MICS (34), which typically draw a fresh cluster sample (up to around 10,000 households) at each 5-year interval, is important here. Although these surveys are designed to yield both time-of-survey cross-sectional data and retrospective data on factors such as mortality, the interpretation in terms of transition is complex because of the 5-year survey intervals. Repeated DHS surveys in a particular country permit pseudo-longitudinal approaches to analysis, but it has to be remembered that there is no intention in the DHS design to re-interview any individual or family longitudinally, meaning that issues of recall bias and inter-sample variation in these repeated cross-sections have to be considered carefully. Figure 2 shows under-five cumulative mortality for Nigeria, as estimated by four separate DHS survey rounds (35–38), which together make the interpretation of changes in this important index of epidemiological transition very difficult. Within the results from each survey, one might expect that increasing recall bias associated with deaths at longer periods before interviews (increasing recall effects at the left end of each survey line) could compete with temporal trends in decreasing mortality; from these four surveys it is difficult to interpret inter- or intra-survey trends over the 30-year period covered, even given the carefully controlled DHS methodologies used.

HDSS data from the INDEPTH Network (39) are an important source of detailed, longitudinal population data that are otherwise largely unavailable in SSA countries. INDEPTH centres typically cover geographically

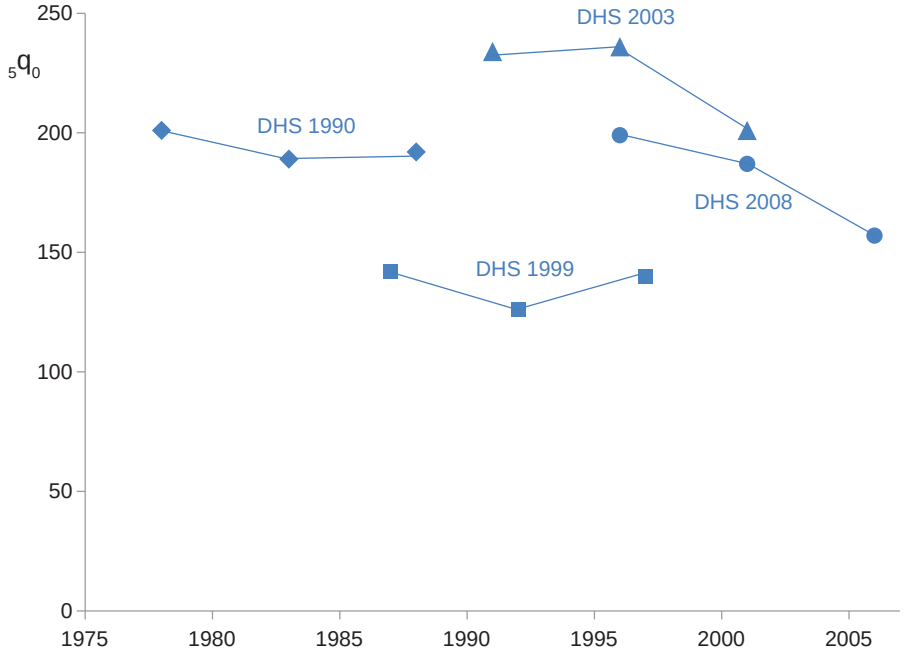

Fig. 2. Under-5 child mortality ($_5q_0$) for Nigeria over three decades, as measured in four Demographic and Health Surveys (DHS).

Citation: Glob Health Action 2014, **7**: 23359 - http://dx.doi.org/10.3402/gha.v7.23359

circumscribed populations (commonly between 50,000 and 150,000 people) longitudinally with regular household visits to record vital events and other data. It has been argued that these local, detailed data may constitute a more appropriate way of monitoring changes over time, such as is needed to monitor progress toward MDGs (40), although counter-arguments that such data may be unrepresentative are frequently made (4).

One possible way to use different sources of country data is to attempt cross-triangulations that compare outcomes across sources, although this can be difficult to accomplish in practice if comparable indices are not covered in different sources. One study from Mozambique compared census, HDSS and DHS data (41), suggesting that it should be possible to use all three sources synergistically in a national system. Several studies in SSA countries have made direct comparisons between national DHS surveys and HDSS data from specific areas within countries, though within-district numbers in DHS data are insufficient for local comparisons (42–44).

Household health survey programmes such as the DHS can add considerable value to national health information systems by providing timely and disaggregated data on mortality levels and trends. Recent advances in methodology to analyse survey responses to measure mortality levels and trends (45, 46) have greatly increased the utility of these methods by controlling for biases, particularly recall error and timeliness of measurements close to the date of the survey. Previously, the most recent estimates could only be made at least 2–3 years preceding the survey. Similarly, improvements to methods for measuring sibling survival have greatly increased the utility of survey-based responses for measuring levels and trends in adult survival.

What would be the methodological implications for measuring epidemiological and health transitions in Africa in support of policy direction?

As health transitions eventually progress and possibly accelerate in SSA countries, it will become increasingly urgent to resolve the major challenges facing a reorientation of African health policies and systems from acute curative services to long-term chronic care. The fragile health systems in SSA, in many instances, have weak dynamic efficiency in terms of being able to respond to new developments (47). These challenges are mirrored for health information systems, which need to move beyond mainly documenting (in practice counting) acute episodes of illness, expanding in scope to include long-term chronic illnesses and repeated patient encounters and continuity of care. Clearly this is a more fundamental challenge than simply considering different diseases, since it implies a greatly increased need for reliable longitudinal follow-up, and in turn this indicates a more important role for longitudinal rather than cross-sectional data.

Accepting the premise that longitudinal data is an important component of information systems capable of tracking epidemiological transition, the additional resources required to implement such systems in most SSA countries would be considerable. Age and cause-specific death registration, at least within some defined population samples, is a pre-requisite for assessing epidemiological transition, and, in the absence of routine medically certified cause-specific death registration, this implies that introducing VA methods in a systematic manner is essential (24). Sampling considerations for population cause-of-death data are not trivial. As well as the conventional considerations of covering age and sex groups, assumptions need to be made about the smallest detectable cause-specific mortality fraction (CSMF). Since the size of ranked CSMFs in most populations follows an approximately exponential distribution, in effect this means deciding how far down the ranking it is worthwhile to go, and then sampling to make that *n*th rank measurable (48). Proof-of-principle for implementing VA methods, including probabilistic modeling of cause of death, with mobile technology (49) indicates that this is an approach ready to move out of research contexts and into routine usage for monitoring cause-specific mortality. Automated VA methods not involving physicians should be actively promoted as a major strategic component of any health information system for routinely tracking epidemiological transition in SSA since the performance of these methods has been found to be generally superior to physician interpretation of VA data, and are rapid, cheap, and consistent over time and place (50–52). One of the principal limitations to the widespread use of VA for measuring population cause of death patterns has been the non-standardized manner in which physicians interpret VA data, and their tardiness in doing so (53).

Less challenging are the national risk factor surveys from national household surveys. SSA countries routinely host repeated and uncoordinated, internationally funded, national household sample surveys such as the living standards surveys, household budget surveys, DHS and MICS surveys, malaria and HIV indicator surveys, and so on. The Health Metrics Network and the International Household Survey Network have encouraged greater integration, harmonization and synergistic scheduling of these surveys. Still there needs to be consensus methodology on the risk factor indicators and their methods of measurement to take advantage of these opportunities.

Where more innovation is needed is measuring both coverage (access) and full effective coverage (outcome) of essential health interventions aimed at the key causes of disease and injury burdens, including risk factors. There is a lack of practical methodology to do so, with

the exception of a few interventions such as immunisation. For routine coverage estimates for district level planning, current Health Management Information Systems at local level do a good job of assessing numerators of the demand, but do not know the respective denominators for each service in the population at risk and in need of that service. Without this critical information, districts cannot estimate their local coverage, and therefore cannot plan and allocate local resources rationally. However without much effort, national programmes could provide expected prevalence or incidence figures from epidemiological surveys and other sources that could be coupled with district demographic data extrapolated from census data to provide the annual service related denominators. This type of epidemiological data sharing between national programmes and district implementers is feasible but practically never done. This is an easy step to determine a crude measure of service coverage (essentially access of those in need), which will already be a useful advance for district planning. Even this crude coverage can be useful for district planning because the rates monitored will often be well below what is expected. But this still stops short of determining effective coverage, which will be even lower than crude coverage. Innovation is needed for methodologies to determine the actual effective coverage of these essential interventions, which goes far beyond just access, to determining the quality and actual health outcomes of the interventions. Effective coverage is a measure of systems effectiveness and needs to be known periodically for policy purposes at national level across a sample of the health system. This will provide a barometer of the performance of the intervention in the real world health system concerned. Determining effective coverage requires a combination of survey and research effort across many disciplines in implementation science. Research and development is needed on these approaches and is already underway (54–56).

Are there particular ethical and political considerations for measuring epidemiological and health transitions in Africa?

In countries with highly functional individual registration as the basis for all national information, including health, there generally exists a widespread (though perhaps not explicit) public confidence that the data system will operate in a confidential and ethical manner, and be under effective governmental regulation. However, it cannot be assumed that this is a readily transferable concept to other settings such as SSA countries. This may be a major barrier to expanding health information systems in SSA in acceptable ways.

While all individual health data should be handled within a robust framework of confidentiality, some data are understandably regarded as more sensitive than others. In SSA, the needs which have emerged for handling large numbers of HIV-related personal records and test results have to some extent brought requirements for patient confidentiality into focus, in contrast to the past when local health centres tended to be very casual about such matters. This demonstrates that paradigm shifts in practice are possible.

There are also important ethical and political considerations surrounding the use of outputs from an effective health information system. In one extreme example, a maternity facility in Burkina Faso was burnt down by protesters angry about the alleged neglect of a woman who died in labour, illustrating that information about health outcomes is by no means a neutral commodity (57), and carries implications for accountability.

National governance issues also impinge on the organization of health services and hence health information systems. In larger countries, operating on a federal or quasi-federal basis with considerable regional autonomy, it may be more appropriate to focus efforts on running good information systems at the regional level. Individual regions in larger countries may be larger than small countries. If health information systems operate well at regional level, then aggregating to national level should be simple. Nigeria and Ethiopia, the two most populous countries in SSA, would be key examples of countries where regional approaches would be valuable.

Moving forward to improve understanding of epidemiological transition in SSA

The current scenario where the health information map of Africa is largely characterized by empty spaces is manifestly unacceptable in today's information age and given the critical need for essential information support to health policy development in Africa. This deplorable situation requires urgent action. Every SSA country needs encouragement and technical resources to improve health information systems to the point where at least reliable in-country national estimates of key parameters become available. Because health in SSA countries is increasingly dominated by longer-term diseases and conditions, good national estimates must have a longitudinal basis and thus be able to reflect change reliably. Sub-national data are also critically important for proper national planning, especially in larger countries, wherein wider variations in health and population exposure to risk factors might be expected. Timeliness of data, particularly on leading causes of death and how they are changing is also critically important if the data are to be useful to inform policy.

For locations where, in the medium-term, universal individual registration of births and deaths and medical certification of cause of death is not feasible, it is strategically important to integrate established population surveillance sites into national data systems, replicating

Citation: Glob Health Action 2014, **7**: 23359 - http://dx.doi.org/10.3402/gha.v7.23359

such sites where appropriate. This implies moving from an isolated field site model of surveillance toward sentinel site networks, such as have been implemented successfully in India and China (58, 59). A balance needs to be struck among emphases on measuring mortality, morbidity and risk-factor outcomes.

A strategic focus is needed such that resources and effort are not dispersed or duplicated over all elements of a health information system, but targeted on the critically important elements of health system management. As argued in this paper, these are: 1) mortality (by age, sex and cause), but not morbidity; 2) periodic data on population-level exposures to major (selected) risk factors, particularly those important for NCDs and injuries; and 3) some measure of the health system response to these health threats, especially the effective coverage (as opposed to measured or estimated coverage) of essential health interventions against the major causes of disease burden.

Clearly these are not strategies that can be implemented without effort or cost. Human resources, particularly in relation to some of the technical issues of data management and quality control, are very scarce in some SSA countries, especially in the government sector. Therefore, careful national implementation plans need to be worked out that are contextually appropriate. There will undoubtedly be implementational challenges; however, the question must be asked whether SSA, and indeed the world, can continue to drift without effective and actionable information on the health of 856 million of the 6,895 million (12%) of the world's population (60). Since this proportion is projected to rise to 1,960 of 9,306 million (21%) by 2050, any delay in tackling the problem will only increase the magnitude of the difficulties.

Conclusions

A more pro-active and strategic approach to health information development in SSA is urgently required by the international community. Local policy is likely to be more effective when based on local evidence. This implies a shift away from relying on global comparative exercises that generate sets of national estimates toward facilitating the development of disaggregated within-country health information systems that are targeted toward monitoring both the leading causes of ill-health in populations, and the response of the health system to controlling them. Lessons must be learnt from effective systems that have been implemented in other regions, but adapted to the African context so that effective data linkages can be made across various levels of health systems and corresponding populations. SSA countries have a chance to take full advantage of the ICT revolution and leap ahead in the design of efficient and interoperable data collection, integration and dissemination technologies in e-Health and m-Health (61). A culture of using health information critically and openly for steering new health policies and system strengthening needs to be encouraged and fostered, and this demands the availability of timely and reliable health statistics.

Health status in some low-income SSA countries is changing more rapidly now than at any prior time in history. To track and steer these population health dynamics and understand what such transitions mean for health systems and policies requires radical change and strengthening of national health information systems to provide essential information. What we propose in this paper is concerted investment on three fronts:

1. interim, strategic investments in sentinel or sample registration systems that provide timely, quality, longitudinal data on deaths and causes of deaths while developing effective civil registration systems for vital statistics;
2. periodic national cross-sectional omnibus sample surveys of the top 10 major risk factors for the leading causes of death; and
3. development of new approaches to estimating district level intervention coverage (access) for annual district planning purposes, and for periodically estimating full effective coverage for national policy purposes through better combinations of routine health service statistics, demographic and epidemiological data for the essential health interventions relevant to these causes.

Countries in SSA and their international funding partners and the global health research community need to prioritize these dimensions in their approach to national strategies for health information systems strengthening. However, information without an appropriate health policy and health system response is not enough. The strengthening of the health information system in this direction needs to be coupled with health system strengthening in order to respond effectively to the dynamics imposed by the health transition (6, 62–65).

The main impediment to implementing these three re-directions of the Health Information System is presently methodological. We need innovations in monitoring burden of cause-specific mortality; in monitoring its attendant risk factors; and in monitoring and verifying the actual coverage of required health system responses. These innovations should focus on improving: sentinel and full CRVS approaches; national household risk factor survey methodologies; and practical coverage assessment methods that can be implemented at, and by the district level.

There is growing momentum toward Universal Health Coverage in the Post-2015 Sustainable Development Goals. These goals will take greater cognizance of demographic change and the social determinants of health (66).

The data collection platforms and priorities proposed in this paper will be essential for understanding universal coverage and how it can be attained. The time is right to consider a proactively deliberate updating of strategies for essential health data in SSA in order to more reliably understand the continent's current needs to improve population health and to better prepare for future trends.

Main findings

- Although epidemiological transition is generally considered to be underway in sub-Saharan Africa, health information systems in most places are insufficient for adequately tracking developments in population health, leading to major gaps in the knowledge needed to plan effective health services.
- Many countries in sub-Saharan Africa get key national health indicators from global estimation processes, rather than gathering adequate national data to understand their own situation and feed into the construction of global estimates; this is insufficient for health planning.
- Very few sub-Saharan African countries currently have or are close to having effective systems for national universal civil registration and vital statistics (CRVS; registering all citizens, including births, deaths and cause of death, and using those data effectively in national statistics).

Key messages for action

- Most sub-Saharan countries are unlikely to implement complete death registration with physician-certified cause any time soon. Although this should be the long-term aim, interim solutions must use standardised verbal autopsy procedures with automated cause of death assignment.
- A consensus is needed on a minimum essential dataset to underpin effective national health information systems in sub-Saharan Africa. Key components would include complete civil registration including verbal autopsy, population-based documentation of major risk factors, and documentation of health service coverage for key preventive and curative measures.
- A key priority across sub-Saharan Africa is the investment in national-level systems that can gather and handle health information effectively, including procedures, equipment and human resources. Countries and their international partners need to understand health status on a population-wide basis.

Conflict of interest and funding

No specific funding was received from any source for writing this paper. Some of the ideas were developed from a US National Academy of Sciences workshop in Johannesburg during October 2011. All of the authors are members of the INDEPTH Network's Scientific Advisory Committee.

References

1. Maher D, Sekajugo J. Research on health transition in Africa: time for action. Health Res Policy Syst 2011; 9: 5.
2. Gill GV, Mbanya JC, Ramaiya K, Tesfaye S. A Sub-Saharan African perspective of diabetes. Diabetologia 2009; 52: 8–16.
3. Mensah G. Epidemiology of stroke and high blood pressure in Africa. Heart 2008; 94: 697–705.
4. Jamison DJ, Feachem RG, Makgoba MW, Bos E, Baingana FK, Hofman KJ, et al. Disease and mortality in Sub-Saharan Africa. Washington, DC: World Bank; 2006, pp. 43–58.
5. Boutayeb A. The double burden of communicable and non-communicable diseases in developing countries. Trans R Soc Trop Med Hyg 2006; 100: 191–9.
6. Atun R, Jaffar S, Nishtar S, Knaul FM, Barreto ML, Nyirenda M, et al. Improving responsiveness of health systems to non-communicable diseases. Lancet 2013; 381: 690–7.
7. UNAIDS (2011). World AIDS Day Report 2011. Geneva: UNAIDS.
8. Murray CJL, Lopez AD. Measuring the global burden of disease. N Engl J Med 2013; 369: 448–57.
9. Dalal S, Beunza J, Volmink J, Adebamowo C, Bajunirwe F, Njelekela M, et al. Non-communicable diseases in Sub-Saharan Africa: what we know now. Int J Epidemiol 2011; 40: 901.
10. Danaei G, Finucane M, Lu Y, Singh G, Cowan M, Paciorek C, et al. Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Blood Glucose). National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants. Lancet 2011; 378: 31–40.
11. Jenkins R, Baingana F, Ahmad R, McDaid D, Atun R. How can mental health be integrated into health system strengthening? Ment Health Fam Med 2011; 8: 115–17.
12. Bryce J, Black R, Victora C. Millennium development goals 4 and 5: progress and challenges. BMC Med 2013; 11: 225.
13. Byass P. The unequal world of health data. PLoS Med 2009; 6: e1000155.
14. Byass P. The imperfect world of global health estimates. PLoS Med 2010; 7: e10001006.
15. Murray CJ, Lopez AD. Production and analysis of health indicators: the role of academia. PLoS Med 2010; 7: e1001004.
16. Sundin J, Willner S. Social Change and Health in Sweden: 25 Years of politics and practice. Stockholm: Swedish National Institute of Public Health; 2007.
17. University of Queensland, Health Information Systems Knowledge Hub. Strengthening practice and systems in civil registration and vital statistics: a resource kit. Working Paper no. 19. Brisbane: University of Queensland; 2012.
18. Byass P, Sankoh O, Tollman S, Högberg U, Wall S. Lessons from history for designing and validating epidemiological surveillance in uncounted populations. PLoS One 2011; 6: e22897.
19. de Savigny D, Kasale H, Mbuya C, Reid G. Fixing health systems. 2nd ed. Ottawa, Canada: International Development Research Centre; 2008.

Citation: Glob Health Action 2014, **7**: 23359 - http://dx.doi.org/10.3402/gha.v7.23359

20. Health Metrics Network. Framework and standards for country health information systems. 2nd ed. Geneva: World Health Organization; 2008.
21. WHO, HMN, University of Queensland (2013). Strengthening civil registration and vital statistics for births, deaths and causes of death: resource kit. Geneva: WHO; 1–238.
22. INDEPTH Network (2013). INDEPTH resource kit for demographic surveillance systems. Accra: INDEPTH Network.
23. Fottrell E, Byass P. Verbal autopsy – methods in transition. Epidemiol Rev 2010; 32: 38–55.
24. Murray CJL, Lopez AD, Shibuya K, Lozano R. Verbal autopsy: advancing science, facilitating application. Popul Health Metr 2011; 9: 18.
25. Lopez AD. Reducing risks to health: what can we learn from the Global Burden of Disease 2010 Study? Int J Public Health 2013; 58: 645–6.
26. World Health Organization (2013). WHO STEPS manual. Geneva: WHO.
27. Boerma T. Public health information needs in districts. BMC Health Serv Res 2013; 13: S12.
28. O'Neill K, Takane M, Sheffel A, AbouZahr C, Boerma T. Monitoring service delivery for universal health coverage: the Service Availability and Readiness Assessment. Bull World Health Organ 2013; 91: 923–31.
29. Massya N, Day C, Dambo M, Barron P, English R, Padarath A. District health barometer 2012/13. Durban: Health Systems Trust; 2013, pp. 1–420.
30. Enkhtuya B, Badamusuren T, Dondog N, Khandsuren L, Elbegtuya N, Jargal G, et al. Reaching every district – development and testing of a health micro-planning strategy for reaching difficult to reach populations in Mongolia. Rural Remote Health 2009; 9: 1045.
31. Lozano R, Soliz P, Gakidou E, Abbott-Klafter J, Feehan DM, Vidal C, et al. Benchmarking of performance of Mexican states with effective coverage. Lancet 2006; 368: 1729–41.
32. Serwaa-Bonsu A, Herbst A, Reniers G, Ijaa W, Clark B, Kabudula C, et al. First experiences in the implementation of biometric technology to link data from Health and Demographic Surveillance. Glob Health Action 2010; 3: 2120.
33. Rutstein S, Rojas G. Guide to DHS statistics. Maryland: ORC Macro; 2006.
34. Bornstein M, Britto P, Nonoyama-Tarumi Y, Ota Y, Petrovic O, Putnick D. Child development in developing countries: introduction and methods. Child Dev 2012; 83: 16–31.
35. Federal Office of Statistics of Nigeria (1992). Nigeria Demographic and Health Survey 1990. Lagos: Federal Office of Statistics.
36. National Population Commission (2000). Nigeria Demographic and Health Survey 1999. Abuja: National Population Commission.
37. National Population Commission (2004). Nigeria Demographic and Health Survey 2003. Abuja: National Population Commission.
38. National Population Commission (2009). Nigeria Demographic and Health Survey 2008. Abuja: National Population Commission.
39. Evans T, AbouZahr C. INDEPTH @ 10: celebrate the past and illuminate the future. Glob Health Action 2008; 1: 1899.
40. Bangha M, Diagne A, Bawah A, Sankoh O. Monitoring the millennium development goals: the potential role of the INDEPTH Network. Glob Health Action 2010; 3: 5517.
41. Nhacolo A, Nhalungo D, Sacoor C, Aponte J, Alonso P. Levels and trends of demographic indices in southern rural Mozambique: evidence from demographic surveillance in Manhiça district. BMC Public Health 2006; 6: 291.
42. Fottrell E, Enquselassie F, Byass P. The distribution and effects of child mortality risk factors in Ethiopia: a comparison of estimates from DSS and DHS. Ethiop J Health Dev 2009; 23: 163–8.
43. Hammer G, Kouyaté B, Ramroth H, Becher H. Risk factors for childhood mortality in sub-Saharan Africa: a comparison of data from a Demographic and Health Survey and from a Demographic Surveillance System. Acta Tropica 2006; 98: 212–18.
44. Byass P, Worku A, Emmelin A, Berhane Y. DSS and DHS. Longitudinal and cross-sectional viewpoints on child and adolescent mortality in Ethiopia. Popul Health Metr 2007; 5: 12.
45. Obermeyer Z, Rajaratnam JK, Park CH, Gakidou E, Hogan MC, Lopez AD, et al. Measuring adult mortality using sibling survival: a new analytical method and new results for 44 countries, 1974–2006. PLoS Med 2010; 7: e1000260.
46. Rajaratnam JK, Tran LN, Lopez AD, Murray CJ. Measuring under-five mortality: validation of new low-cost methods. PLoS Med 2010; 7: e1000253.
47. Brooks A, Smith TA, de Savigny D, Lengeler C. Implementing new health interventions in developing countries: why do we lose a decade or more? BMC Public Health 2012; 12: 683.
48. Rao C, Porapakkham Y, Pattaraarchachai J, Polprasert W, Swampunyalert N, Lopez AD. Verifying causes of death in Thailand: rationale and methods for empirical investigation. Popul Health Metr 2010; 8: 11.
49. Hounton S, Sombié I, Meda N, Bassane B, Byass P, Stanton C, et al. Methods for evaluating effectiveness and cost-effectiveness of a Skilled Care Initiative in rural Burkina Faso. Trop Med Internat Health 2008; 13: 14–24.
50. Lozano R, Lopez AD, Atkinson C, Naghavi M, Flaxman AD, Murray CJ. Performance of physician-certified verbal autopsies: multisite validation study using clinical diagnostic gold standards. Popul Health Metr 2011; 9: 32.
51. Byass P, Kahn K, Fottrell E, Mee P, Collnson M, Tollman S. Using verbal autopsy to track epidemic dynamics: the case of HIV-related mortality in South Africa. Popul Health Metr 2011; 9: 46.
52. Leitao J, Chandramohan D, Byass P, Jakob R, Bundhamcharoen K, Choprapawon C, et al. Revising the WHO verbal autopsy instrument to facilitate routine cause-of-death monitoring. Glob Health Action 2013; 6: 21518.
53. Gakidou E, Lopez AD. What do children die from in India today? Lancet 2010; 376: 1810–11.
54. Alonso PL, Bell D, Hanson K, Mendis K, Newman RD, de Savigny D, et al. A research agenda for malaria eradication: health systems and operational research. PLoS Med 2011; 8: e1000397.
55. Littrell M, Miller JM, Ndhlovu M, Hamainza B, Hawela M, Kamuliwo M, et al. Documenting malaria case management coverage in Zambia: a systems effectiveness approach. Malar J 2013; 12: 371.
56. Rao VB, Schellenberg D, Ghani AC. Overcoming health systems barriers to successful malaria treatment. Trends Parasitol 2013; 29: 164–80.
57. Anon. Safe World for Women. Mother's death sparks riot in Burkina Faso; 2011. [cited 2011 Oct 20]. Available from: http://www.asafeworldforwomen.org/womens-rights/in-the-news/1180-mothers-death-sparks-riot-in-burkina-faso.html
58. Saikia N, Jasilionis D, Ram F, Shkolnikov V. Trends and geographic differentials in mortality under age 60 in India. Popul Stud J Demogr 2011; 65: 73–89.
59. Yang G, Hu JT, Rao K, Ma J, Rao C, Lopez AD. Mortality registration and surveillance in China: history, current situation and challenges. Popul Health Metr 2005; 3: 3.

60. United Nations Population Division (2010). World population prospects, 2010 revision. New York: United Nations Population Division; 2010.
61. Kabadi G, Mwanyika H, de Savigny D. Innovations in monitoring vital events: mobile phone SMS support to improve coverage of birth and death registration: a scalable solution. Brisbane: Health Information Systems Knowledge Hub, School of Population Health, The University of Queensland; 2013, pp. 1–25.
62. Swanson RC, Cattaneo A, Bradley E, Chunharas S, Atun R, Abbas KM, et al. Rethinking health systems strengthening: key systems thinking tools and strategies for transformational change. Health Policy Plan 2012; 27(Suppl 4): iv54–iv61.
63. Atun R. Health systems, systems thinking and innovation. Health Policy Plan 2012; 27(Suppl 4): iv4–iv8.
64. de Savigny D, Adam T. Systems thinking for health systems strengthening. Geneva: Alliance for Health Policy and Systems Research; 2009.
65. Adam T, de Savigny D. Systems thinking for strengthening health systems in LMICs: need for a paradigm shift. Health Policy Plan 2012; 27: iv1–iv3.
66. Kickbusch I, Brindley C. Health in the post 2015 development agenda. Geneva: WHO, pp. 1–71.

COACTION PUBLISHING

MEDICINE & HEALTH
SERIES