This is the published version of a paper published in *Global health action*.

Citation for the original published paper (version of record):

The development and experience of epidemiological transition theory over four decades: a systematic review.
*Global health action*, 7: 23574
http://dx.doi.org/10.3402/gha.v7.23574

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-90531
The development and experience of epidemiological transition theory over four decades: a systematic review

Ailiana Santosa1*, Stig Wall1, Edward Fottrell2, Ulf Högberg3 and Peter Byass1,4

1Division of Epidemiology and Global Health, Department of Public Health and Clinical Medicine, Umeå Centre for Global Health Research, Umeå University, Umeå, Sweden; 2Institute for Global Health, University College London, London, UK; 3Department of Women’s and Children’s Health, Uppsala University, Uppsala, Sweden; 4MRC/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: Barthélemy 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: ailiiana.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).
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
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.
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 elderly 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...
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...
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).

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

<table>
<thead>
<tr>
<th>Year</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Stage 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1870–1874</td>
<td>United Kingdom (153, 157), Sweden (153, 157), Finland (153, 157), Iceland (153, 157), USA (66), Germany (167)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1875–1879</td>
<td>The Netherlands (113)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1880–1884</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1885–1889</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1890–1894</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1895–1899</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1900–1904</td>
<td>Spain (114, 116, 117)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1905–1909</td>
<td>United Kingdom (124, 153, 157), Sweden (153, 157), Finland (153, 157), Iceland (153, 157), The Netherlands (113)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1910–1911</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1915–1919</td>
<td>Mexico (17)</td>
<td>Spain (114, 168)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1920–1924</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1925–1929</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1930–1934</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1935–1939</td>
<td>Japan (91)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1940–1944</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1945–1949</td>
<td>Hong Kong (154), Trinidad and Tobago (64, 65)</td>
<td>Costa Rica (55)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1950–1954</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1955–1959</td>
<td></td>
<td>China (158)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1960–1964</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1965–1969</td>
<td>Hong Kong (154), Singapore (154)</td>
<td></td>
<td>Cuba (140), Puerto Rico (140)</td>
<td></td>
</tr>
<tr>
<td>1970–1974</td>
<td>Peru (77)</td>
<td>Trinidad and Tobago (64, 65)</td>
<td></td>
<td>Portugal (55)</td>
</tr>
<tr>
<td>1975–1979</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1980–1984</td>
<td></td>
<td>Korea (95)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1985–1989</td>
<td>Mexico (59, 140)</td>
<td>Hong Kong (154) Malaysia (154)</td>
<td>Costa Rica (55), Canada (47), Spain (168)</td>
<td></td>
</tr>
<tr>
<td>1990–1994</td>
<td>Peru (77)</td>
<td>Seychelles (37), Ghana (32), Cuba (54)</td>
<td>Japan (158), France (108), Australia (126, 127), Trinidad and Tobago (158)</td>
<td></td>
</tr>
<tr>
<td>1995–1999</td>
<td></td>
<td></td>
<td>China (158)</td>
<td></td>
</tr>
<tr>
<td>2000–2004</td>
<td>Mexico (17), Korea (93)</td>
<td>India (85, 86), Thailand (100, 101)</td>
<td>Mexico (60), Peru (76), Hungary (109), Brazil (69, 73, 118)</td>
<td></td>
</tr>
<tr>
<td>2005–2009</td>
<td></td>
<td></td>
<td>Vietnam (104), Mexico (170)</td>
<td></td>
</tr>
</tbody>
</table>
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, with cardiovascular 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, socioeconomic, 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
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.

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


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, socioeconomics and macro factors such as climate change.
The development and experience of epidemiological transition theory over four decades

51. Chen J, Millar WJ. Are recent cohorts healthier than their predecessors? Health Rep 2000; 11: 9–23 (Eng); 9–6 (Fre).
100. Akgun S, Rao C, Yardim N, Basara BB, Aydin O, Mollahaliloglu
98. Niti M, Ng TP.
97. Pradhan J, Arokiasamy P. Socio-economic inequalities in child
95. Lee ES. Epidemiologic transition in Korea: a new perspective
87. Hoa NP, Rao C, Hoy DG, Hnh ND, Chuc NT, Ngo DA.
54. Wild S, Fischbacher C, Brock A, Griffiths C, Bhopal R. Mortality from all causes and circulatory disease by country of


