



Scan to access more
free content

Using 'amenable mortality' as indicator of healthcare effectiveness in international comparisons: results of a validation study

Johan P Mackenbach,¹ Rasmus Hoffmann,¹ Bernadette Khoshaba,² Iris Plug,¹ Grégoire Rey,³ Ragnar Westerling,⁴ Kersti Pärna,⁵ Eric Jougl,³ José Alfonso,⁶ Caspar Looman,¹ Martin McKee²

► An additional appendix is published online only. To view this file please visit the journal online (<http://dx.doi.org/10.1136/jech-2012-201471>).

¹Department of Public Health, Erasmus MC, Rotterdam, the Netherlands

²London School of Hygiene and Tropical Medicine, London, UK

³CépiDc, Institut National de la Santé et de la Recherche Médicale (INSERM), Le Vésinet, France

⁴Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden

⁵Department of Public Health, University of Tartu, Tartu, Estonia

⁶Fundación de Investigación, Hospital General Universitario de Valencia, Valencia, Spain

Correspondence to

Professor Dr Johan P Mackenbach, Department of Public Health, Erasmus MC, P.O. Box 2040, Rotterdam 3000 CA, the Netherlands; j.mackenbach@erasmusmc.nl

Accepted 19 July 2012

Published Online First
25 September 2012

ABSTRACT

Background and study aims There is widespread consensus on the need for better indicators of the effectiveness of healthcare. We carried out an analysis of the validity of amenable mortality as an indicator of the effectiveness of healthcare, focusing on the potential use in routine surveillance systems of between-country variations in rates of mortality. We assessed whether the introduction of specific healthcare innovations coincided with declines in mortality from potentially amenable causes in seven European countries. In this paper, we summarise the main results of this study and illustrate them for four conditions.

Data and methods We identified 14 conditions for which considerable declines in mortality have been observed and for which there is reasonable evidence in the literature of the effectiveness of healthcare interventions to lower mortality. We determined the time at which these interventions were introduced and assessed whether the innovations coincided with favourable changes in the mortality trends from these conditions, measured using Poisson linear spline regression. All the evidence was then presented to a Delphi panel.

Main results The timing of innovation and favourable change in mortality trends coincided for only a few conditions. Other reasons for mortality decline are likely to include diffusion and improved quality of interventions and in incidence of diseases and their risk factors, but there is insufficient evidence to differentiate these at present. For most conditions, a Delphi panel could not reach consensus on the role of current mortality levels as measures of effectiveness of healthcare.

Discussion and conclusions Improvements in healthcare probably lowered mortality from many of the conditions that we studied but occurred in a much more diffuse way than we assumed in the study design. Quantification of the contribution of healthcare to mortality requires adequate data on timing of innovation and trends in diffusion and quality and in incidence of disease, none of which are currently available. Given these gaps in knowledge, between-country differences in levels of mortality from amenable conditions should not be used for routine surveillance of healthcare performance. The timing and pace of mortality decline from amenable conditions may provide better indicators of healthcare performance.

INTRODUCTION

There is widespread consensus on the need for better indicators of the effectiveness of healthcare, both for inclusion in routine surveillance systems and as a tool for research.¹ Several decades ago, it was proposed that rates of mortality from certain causes that are amenable to medical care could be used as indicators of the effectiveness of healthcare. This gave rise to the concept of 'avoidable' or 'amenable' mortality, originally developed by Rutstein *et al*² for application in a clinical context, and operationalised for application at the population level by Holland *et al*.^{3–6}

The concept attracted considerable interest in the 1980s,⁷ gaining momentum through a European Commission-funded Concerted Action Project on 'Health Services and Avoidable Deaths'. This culminated in publication of the European Community Atlas of 'Avoidable Death' in 1988,⁸ a major contribution that was subsequently updated⁹ and later replicated in Central and Eastern Europe.¹⁰ Eurostat is currently considering whether to include mortality from conditions amenable to healthcare in its standard set of mortality indicators, the English Department of Health has selected it as one of its high-level performance indicators¹¹ and the Commonwealth Fund in the USA uses it to compare performance of the 50 states.¹² The concept has also been widely used in a research context in Europe and in other parts of the world, for example, to measure differences in performance of national health systems,^{13–20} the relationship between expenditure and health outcomes^{21–22} or variations in effectiveness of health services between socioeconomic or ethnic groups.^{23–27}

The selection of causes of death that were considered by Rutstein, Holland and others to be amenable to medical intervention was, however, based on an informal procedure, drawing heavily on expert opinion. Although subsequent studies have taken a more systematic approach, including much more extensive reviews of evidence,²⁸ the validity of these causes of death as indicators of the effectiveness of healthcare has never been formally and systematically demonstrated. Furthermore, since the first publications on the concept of 'amenable mortality', scientific methods for the measurement and analysis of population health have improved, making it all the more necessary to apply rigorous standards to the selection of indicator conditions.

We have therefore carried out a partial analysis of the validity of amenable mortality as an indicator of the effectiveness of healthcare. We have focused on the potential use of these indicators in routine surveillance systems in which variations between countries in rates of mortality at one point in time would be interpreted as indicating variations in effectiveness of healthcare. In our validation study, we sought to exploit differences between countries in the time of introduction of healthcare innovations and assessed whether the introduction of innovations in healthcare coincided with declines in mortality from potentially amenable causes in seven European countries. Here, we summarise the main conclusions of this analysis, illustrated with key examples, and discuss possible ways forward. A full report is available online.²⁹

SUMMARY OF STUDY DESIGN AND MAIN RESULTS

We defined 'amenable mortality' as causes of death for which mortality rates are likely to reflect variations in the effectiveness of healthcare, with healthcare being limited to primary care, hospital care and personalised public health services (eg, immunisation and screening).

We first selected 54 causes of death for which a considerable decline in mortality has occurred in recent years and for which there was still a sufficient number of deaths in 2000 to allow a meaningful analysis of between-country variations in mortality. These 54 causes were selected by screening all three-digit cause-of-death codes from England and Wales, as an example of one of the larger European Union countries and one where coding specificities were well understood. We took the period 1979–2000, when ICD-9 was in use, to avoid problems arising from a change of ICD version. We selected causes of death for which mortality declined more than 30% between 1979 and 2000 and for which the number of deaths in 2000 exceeded 100.

For these conditions, we conducted systematic reviews of the literature in order to identify healthcare interventions, which were introduced after 1970 and which, according to evidence from patient- or population-level studies, have effectively reduced mortality from these conditions. The strength of the evidence was highly variable, and our literature reviews revealed a number of important weaknesses in the evidence base. Nevertheless, for 14 conditions, we were able to identify interventions that more or less fulfilled this criterion (HIV/AIDS, colon cancer, cervical cancer, testicular cancer, Hodgkin's disease, rheumatic heart disease, hypertension, ischaemic heart disease, heart failure, cerebrovascular disease, peptic ulcer, renal failure, congenital heart disease and perinatal conditions).

For these interventions, we then identified the timing of introduction in seven European countries (Estonia, France, (West) Germany, the Netherlands, Spain, Sweden, United Kingdom). The selection of countries was based on several considerations: sufficient size to generate robust findings (mortality by cause is very unstable in small countries such as Iceland, Luxembourg or Malta), access to data (there are significant delays in production of mortality data from some countries such as Belgium and Denmark), geographical distribution (and thus economic conditions, with countries from North, South, East and West Europe) and finally, the existence of willing partners with sufficient expertise. The research was undertaken by partners in these countries, for example, by hand searching formularies for the first date of introduction of a drug, tracking sales data, identifying early trials, reviewing documents of professional societies and interviewing knowledgeable informants. Published sources of information are given in the online

appendix. There were wide variations between countries in time of introduction of new interventions: for half of the innovations, the difference between the earliest and the latest adopter was more than 10 years.

For each of the 14 conditions, we also obtained mortality data for the seven countries in the study, covering the period 1970–2005. For this purpose, Estonian mortality data from the years before 1990 had to be reclassified from the abridged Soviet version of ICD-9 to the international version of ICD-9. Possible coding changes that might influence mortality trends were identified by an automated jump detection method.³⁰ Correction factors were derived, and these were applied to create a consolidated database.

For each of the 14 conditions, and for men and women separately, mortality trends were analysed using Poisson regression analysis. Age (in 5-year age groups) was controlled for, and in the analyses reported here, no age limits were imposed. In a second series of mortality analyses, we applied an upper age limit of 75 years, and this analysis produced essentially the same results. We applied linear spline regression to identify points in time ('knots') at which the mortality trend changed significantly (eg, a mortality decline started or accelerated). We found that although mortality decline was the general pattern, this was far from uniform, both across time and between countries. A model with three knots described all mortality patterns sufficiently well, and we then related these knots to the approximate periods in which a decline in mortality from the introduction of innovations could be expected. We found slightly more matches between the two than could be expected to occur by chance alone. Introduction of innovation and favourable change in mortality trend coincided for only four conditions out of the 14 (ie, HIV/AIDS, colorectal cancer, ischaemic heart disease and cerebrovascular disease).

The results of each of the preceding steps were summarised for each condition in a one-page vignette, and experts were asked to assess the likelihood that variations in current mortality levels from these conditions reflect variations in the effectiveness of healthcare. We involved 23 experts from 16 countries who were either producers or users of evidence on health system performance, seeking participation by both genders and health professionals as well as non-health professionals. We held a Delphi procedure with two rounds³¹ in which the experts completed an online form for each vignette and were asked to score each cause of death on a scale of 1–9, where 1 was an entirely inappropriate indicator of health system performance and 9 was an entirely appropriate indicator. At the end of the first round, the distribution of scores for each cause was recorded and each participant was fed back both the distribution and their personal score, along with a summary of the comments that had been made. They were then asked to rescore them. In keeping with conventional usage, consensus was deemed to exist when, after excluding the highest and lowest individual scores, the remainder lay within a three-point range. After the second round, the experts had reached consensus on only three out of these 14 conditions (colorectal cancer, cervical cancer and cerebrovascular disease). Their comments suggested that they differed in their expectations and interpretations of indicators and, in particular, the ability to partition the impact of healthcare from other factors.

FOUR ILLUSTRATIONS

We illustrate these results and their interpretation for four candidate conditions, representing both 'positive' and 'negative' findings: HIV/AIDS, Hodgkin's disease, breast cancer and

Table 1 Selected numerical data for four candidate conditions

Condition	Innovation	Timing of introduction of innovation (range of variation between earliest and latest adopter)	Number of matches between innovation and mortality trend change (out of total number of possible matches)	Timing of corresponding change in mortality trend (range of variation between earliest and latest change, by gender)	Delphi results: median score† and whether consensus was achieved	Correlation between recent mortality level and mortality level before favourable change in trend‡	Correlation between recent mortality level and mortality decline since favourable change in trend§
HIV/AIDS	Antiretroviral treatment	1987 the Netherlands—1995 Estonia	11/14	M: 1989 the Netherlands—1997 Estonia F: 1994 France—1995 Germany, Spain, Sweden	7; no	M: 0.53 F: 0.91**	M: 0.50 F: -0.27
Hodgkin's disease¶	High-dose therapy and peripheral blood stem cell transplantation	1985 Sweden—1999 UK	3/12	NA	7; no	M: -0.40 F: 0.00	M: 0.76* F: 0.73
Breast cancer	1) Screening with mammography 2) Treatment with Tamoxifen	1) 1975 the Netherlands—2005 Estonia 2) 1976 UK and Sweden—1992 Estonia	5/14	NA	8; no	F: 0.84**	F: -0.19
Cerebrovascular disease	1) Detection and treatment of hypertension 2) Intensive management of stroke††	1) 1970 Sweden—1997 Estonia 2) 1992 France—1998 UK and Estonia	17/24	1) M: 1974 Sweden—2003 Estonia F: 1975 Sweden—2003 Estonia 2) M: 1996 Germany—2003 Estonia, UK F: 2002 the Netherlands—2003 Estonia, UK	7; yes	M: 0.56 F: 0.80**	M: 0.79** F: 0.66*

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

†On a scale from 1 to 9.

‡Recent=2005; mortality level before favourable change=1995 for HIV, 1984 for Hodgkin's disease, 1991 for breast cancer and 1970 for cerebrovascular disease.

§Positive correlations mean: lower mortality levels in 2005 are associated with steeper decreases in mortality. See†† for years.

¶No mortality data on Hodgkin's disease were available for Estonia.

††No timing of introduction of intensive management of stroke was available for Spain and Sweden.

F, female, M, male.

cerebrovascular disease. Table 1 gives numerical data on these conditions, and figure 1 illustrates mortality trends (for reasons of space, this was restricted to one gender only for each condition). Our reasoning was that for variations in current national mortality levels to be valid indicators of healthcare effectiveness, current mortality levels should primarily reflect the degree of mortality decline induced by the introduction of healthcare interventions. They should not reflect pre-existing mortality levels or the extent of autonomous trends in mortality.

In the case of HIV/AIDS, the main innovation occurring in the 1970–2005 period was treatment with antiretroviral drugs, which was introduced in the late 1980s and early 1990s, with some variations between countries³² (table 1). The few randomised clinical trials with mortality as an outcome^{33–35} cannot easily be generalised to the population now receiving combination therapy. Trials with mortality as an endpoint rapidly gave way to those with surrogate markers, as it was felt that, given the pace of therapeutic developments, results would otherwise be obsolete by the time they were published.³⁶ Despite these problems, however, there is widespread consensus, justified by the combination of evidence from trials and observational studies that these drugs effectively lower mortality among patients,³⁷ even if it is difficult to quantify the impact.³⁸ In most of the countries participating in this study mortality from HIV/AIDS shows steep increases in the late 1980s and early 1990s, related to the spread of the epidemic among homosexual men, injecting drug users and other risk groups,³⁹ followed by a sudden decline after the mid-1990s (figure 1A). The introduction of antiretroviral drugs coincided with a favourable change in mortality trend in most countries (sometimes preceding the peak in the mid-1990s). Countries in which antiretroviral drugs were introduced earlier tended to have earlier changes in mortality trend, as shown by the large number of 'matches' between innovation introduction and mortality trend change. The Delphi panel, however, could not reach consensus on the validity of this indicator, mainly because it felt that mortality from HIV/AIDS is also strongly determined by incidence. The final two columns of table 1 show the correlations of the recent mortality level from HIV/AIDS with, respectively, the mortality level before the favourable change in trend and the mortality decline since the favourable change. The first are much more positive than the second, suggesting that current national levels of mortality from HIV/AIDS primarily reflect the peak levels of mortality from this condition and not the extent of mortality decline since these peaks were reached. This casts doubt on the validity of current national mortality levels from HIV/AIDS as indicators of the effectiveness of healthcare for this specific condition.

In the case of Hodgkin's disease, considerable progress had already been made with chemotherapy,⁴⁰ with a substantial improvement in long-term survival by 1970.⁴¹ The main innovation between 1970 and 2005 was high-dose chemotherapy combined with peripheral blood stem cell transplantation for those whose disease was not responding to conventional chemotherapy.⁴² This was introduced in the late 1980s and 1990s, with considerable variation between countries (table 1). Mortality from Hodgkin's disease, however, followed a more or less uniform decline during this period (figure 1B), and there were almost no 'matches' between innovation introduction and mortality trend change. The Delphi panel could not reach consensus on the validity of this indicator.

In the case of breast cancer, we identified two important innovations, screening with mammography and treatment with Tamoxifen. There is considerable evidence that the former is

Avoidable mortality

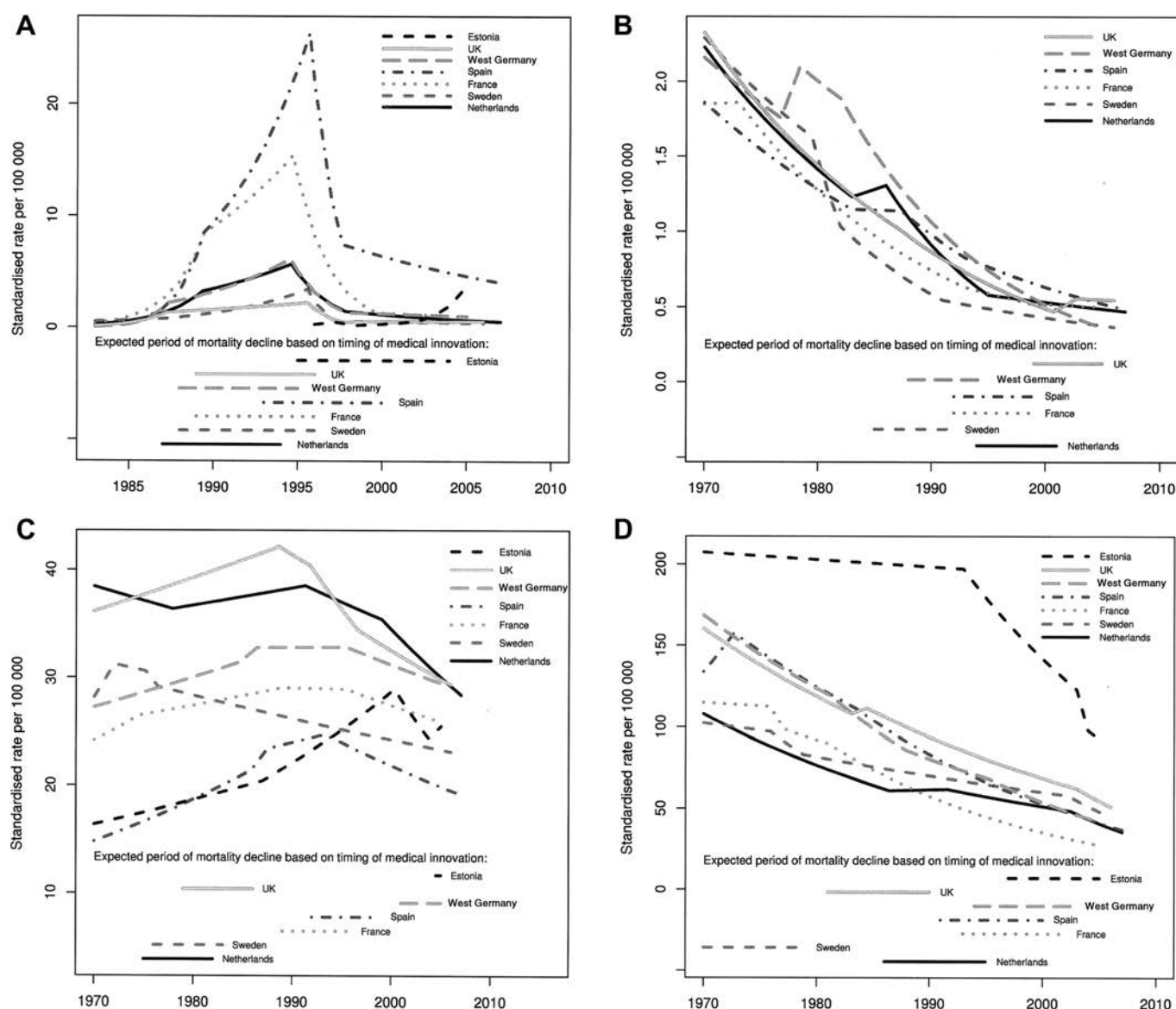


Figure 1 Mortality from four conditions and expected periods of mortality decline based on timing of innovation. (A) Estimated mortality trends from HIV/AIDS (men only) and expected periods of mortality decline based on the introduction of antiretroviral drugs. (B) Estimated mortality trends from Hodgkin's disease (men only) and expected periods of mortality decline based on the introduction of high-dose therapy and peripheral blood stem cell transplantation. (C) Estimated mortality trends from malignant neoplasm of breast (women only) and expected periods of mortality decline based on the introduction of screening by mammography. (D) Estimated mortality trends from cerebrovascular disease (women only) and expected periods of mortality decline based on the introduction of prevention by detection and treatment of hypertension.

effective in saving lives,⁴³ even if this is disputed,^{44 45} while the latter has been shown to be effective in prolonging survival, especially in women with oestrogen receptor-positive disease.⁴⁶ Although the benefits of screening are still controversial, it is likely that the observed reductions in population-level mortality reflected the combined effect of both innovations. They were introduced over a protracted period, with considerable variation between countries and with substantial overlap between the periods when both innovations were introduced in each of the countries we studied (table 1). In many countries, mortality from breast cancer increased during the 1970s and 1980s, peaked in the early 1990s and then declined (figure 1C). There were very few 'matches' between the introduction of innovations and changes in the mortality trend. The Delphi panel could not reach consensus on the validity of this indicator. As the final columns of table 1 show, the between-country variations in

current breast cancer mortality level reflect the mortality level before mortality started to decline, which is likely to reflect disease incidence.

In the case of cerebrovascular disease, we identified two important innovations, detection and treatment of hypertension and intensive management of stroke. Both have been shown to lower mortality.^{47 48} The first was introduced progressively over a protracted period from the 1950s onwards, with enormous variations in the coverage and intensity of treatment between countries. The second was introduced only recently, with less variation (table 1). Mortality from cerebrovascular disease has declined substantially in all countries in this study but with some accelerations and decelerations (figure 1D). We found a large number of 'matches' between innovation introduction and mortality trend change, and in this case, the Delphi panel also reached consensus on the validity of this indicator.

Nevertheless, as the final columns of table 1 show, the current levels of cerebrovascular disease mortality not only reflect the degree of decline since 1970, but also the mortality level before it started to decline, suggesting that between-country variations in current mortality levels may be driven to a considerable extent by the prevalence of hypertension and associated risk factors and, potentially, patterns of death certification.

THE WAY FORWARD

The 14 conditions that we studied have been selected on the basis of evidence that there has been a substantial reduction in mortality and that effective interventions exist, but despite accumulating much more information than has previously been brought together, in most cases, we have not found clear evidence for a link between the reductions in mortality and the introduction of specific interventions. What does this mean?

It could of course mean that contrary to what we assumed, the interventions have not been effective in lowering mortality. Indeed, there is no doubt that the evidence on the effectiveness of interventions that we found in the literature is in many ways unsatisfactory. Few clinical trials assess mortality, instead focusing on intermediate or surrogate outcomes, even though these can be quite misleading.^{49–51} Most are undertaken with highly selected subjects, often in potentially atypical centres of excellence, so it is difficult to extrapolate results to the general population.⁵² Few clinical trials compare a new intervention against no treatment, instead showing incremental improvements on existing treatments. Consequently, while the cumulative effect of progressively better treatments may be considerable, it cannot be captured from trials as reported in the literature. Observational studies on clinical outcomes, on the other hand, often show important improvements in survival, but usually cannot link these to specific interventions. For example, it is apparent from the substantial reductions in mortality from common surgical conditions that non-urgent surgery has become much safer, but this is likely to reflect many different advances in anaesthesia, infection control and surgical technique. Similarly, there have been substantial improvements in long-term survival for many treatable cancers, but again the available studies rarely link this to a specific innovation.⁵³

An alternative interpretation therefore is that innovations in healthcare did lower mortality from many of the conditions that we studied but occurred in a much more diffuse way than we assumed in our study design, which relied on discrete changes in mortality trends. We had considerable difficulty in finding information on when specific innovations were introduced. It was easier for pharmaceutical innovations, but even there we found examples of off-label prescribing prior to licensing. More importantly, the introduction of a particular innovation is only one factor that leads to increasing effectiveness. The determinants of increasing effectiveness can be represented by a three-dimensional diagram (figure 2), with innovation on one axis, coverage (or diffusion) on a second and quality with which it is applied on a third. Take hypertension: although the first drugs became available in the 1950s (innovation⁵⁴), the threshold for treatment has changed, with progressively lower levels of blood pressure seen as justifying initiation of treatment,⁵⁵ which has simultaneously been extended to those at older ages (coverage⁵⁶). The 'rule of thumb' that treatment was only required when systolic pressure was '100 plus your age' is now viewed as incorrect and obsolete. Over this time, the quality of care has also improved, with newer classes of drugs having fewer side effects, and thus improved compliance,⁵⁷ as well as the introduction of combined treatment that maximises the

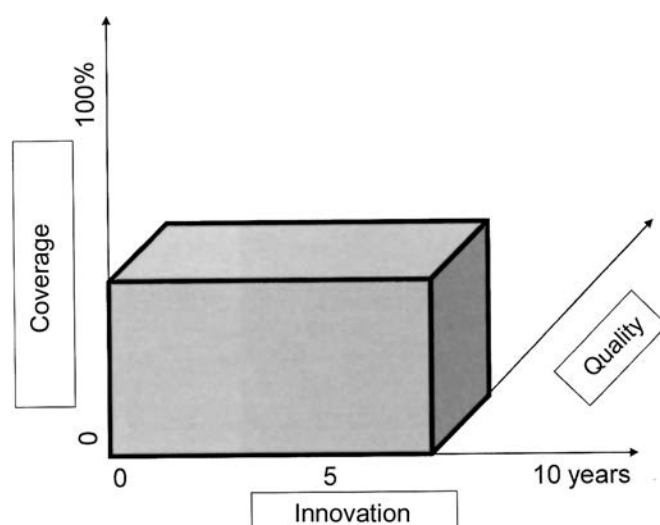


Figure 2 Dimensions of healthcare improvements.

intended effects but minimises the unintended ones,⁵⁸ and the introduction of practice guidelines which have systematised detection and treatment of hypertension.⁵⁹ We only studied one of the three factors, with considerable difficulty, and in hind-sight, it is therefore not surprising that we were rarely able to link reductions in mortality to specific healthcare interventions.

Even if, as we believe, improvements in healthcare have contributed importantly to lowering mortality from these conditions, rates of mortality from amenable conditions must be used with caution as indicators of healthcare effectiveness in international comparisons, as our data show (last two columns of table 1). Countries with relatively high mortality rates from these conditions in 2005 often already had relatively high mortality before the innovations were introduced, as shown by the positive correlations between recent mortality levels and the mortality levels before a favourable change in trend. Only in the case of Hodgkin's disease do recent mortality levels seem primarily to reflect the steepness of decline. This implies that variations in current absolute levels of mortality in countries are likely to be determined, to some extent, by variations in the incidence of diseases and their risk factors, which may mask variations in healthcare effectiveness. In other words, at the end of a process of gradual diffusion (and increasing experience in their use along the way) of new interventions, the residual level of mortality probably reflects the influence of other factors that already prevailed before the introduction of the innovation.

It may therefore be useful to consider two other quantitative measures of mortality from conditions amenable to healthcare intervention. The first is the timing of a favourable change in mortality trends that can plausibly be linked to advances in treatment. In some cases, a clear change in trend could be seen, and the timing of these changes has been illustrated for three conditions in table 2. Sweden was clearly ahead of many other countries in bringing down breast cancer mortality (which could be associated with its early adoption of screening in regional pilot studies) and in bringing down cerebrovascular disease mortality (plausibly associated with effective detection and treatment of hypertension). Estonia, on the other hand, was only able to adopt many new interventions after independence in 1991. This indicator may reflect the responsiveness of healthcare systems to innovation.

Table 2 Possible alternative measures of amenable mortality, illustrated for four candidate conditions

	HIV/AIDS	HIV/AIDS	Hodgkin's disease	Hodgkin's disease	Breast cancer	Breast cancer	Cerebrovascular disease	Cerebrovascular disease	Cerebrovascular disease
	Timing of start of mortality decline, by gender (95% CIs)	% Mortality decline per year since start of mortality decline, by gender (95% CIs)	Timing of start of mortality decline, by gender (95% CIs)	% Decline per year since 1984 (95% CIs)	Timing of start of mortality decline (95% CIs)	% Mortality decline per year since favourable change (95% CIs)	Timing of start of mortality decline due to detection and treatment of hypertension, by gender (95% CIs)	% Decline per year since start of mortality decline due to detection and treatment of hypertension, by gender (95% CIs)	Timing of start of mortality decline due to intensive management, by gender (95% CIs)
UK	M: 1996 (95 to 96) F: 1996 (96 to 97)	M: 18.0 (8.9 to 26.1) F: 0.6 (−4.9 to 5.9)	NA	M: 4.4 (3.9 to 4.9) F: 3.5 (3 to 3.9)	F: 1989 (88 to 89)	F: 2.3 (2.1 to 2.4)	M: 1984 (84 to 84) F: 1984 (84 to 84)	3.5 (3.3 to 3.7) 3.0 (2.8 to 3.2)	M: 2003 (03 to 03) F: 2003 (03 to 03)
Germany	M: 1995 (94 to 95) F: 1995 (95 to 95)	M: 23.7 (11.5 to 34.2) F: 14.3 (7.7 to 20.5)	NA	M: 7.2 (6.6 to 7.8) F: 6.0 (5.3 to 6.7)	F: 1986 (86 to 87)	F: 0.4 (0.2 to 0.5)	M: 1996 (96 to 97) F: 1995 (95 to 96)	5.8 (5.4 to 6.2) 4.7 (4.3 to 5.0)	M: no change F: no change
Spain	M: 1996 (96 to 96) F: 1996 (96 to 96)	M: 24.8 (14.3 to 34.1) F: 22.8 (13.6 to 30.9)	NA	M: 3.7 (3.2 to 4.2) F: 2.5 (2.1 to 3.0)	F: 1993 (93 to 94)	F: 2.1 (1.9 to 2.4)	M: 1984 (83 to 85) F: 1984 (84 to 84)	4.5 (4.2 to 4.7) 4.6 (4.4 to 4.8)	M: no change F: 2003 (03 to 03)
France	M: 1995 (95 to 95) F: 1995 (95 to 95)	M: 30.0 (23.0 to 36.4) F: 26 (20.2 to 31.4)	NA	M: 3.8 (3.3 to 4.3) F: 2.9 (2.3 to 3.5)	F: 1995 (94 to 97)	F: 1.4 (1.1 to 1.7)	M: 1986 (86 to 86) F: 1981 (81 to 81)	4.9 (4.6 to 5.1) 4.8 (4.6 to 5.0)	M: no change F: no change
Sweden	M: 1996 (95 to 96) F: 1996 (96 to 96)	M: 19.6 (13.3 to 25.5) F: 15.5 (6.8 to 23.3)	NA	M: 4.5 (3.4 to 5.5) F: 3.8 (2.4 to 5.2)	F: 1972 (72 to 72)	F: 0.8 (0.8 to 0.9)	M: 1974 (74 to 75) F: 1976 (75 to 76)	1.7 (1.6 to 1.9) 1.9 (1.7 to 2.0)	M: 2000 (99 to 00) F: 2002 (02 to 03)
The Netherlands	M: 1995 (94 to 95) F: 1995 (95 to 95)	M: 22.2 (16.7 to 27.4) F: 14.2 (10.1 to 18)	NA	M: 5.0 (4.2 to 5.9) F: 4.1 (3.0 to 5.1)	F: 1991 (90 to 93)	F: 1.5 (1.2 to 1.7)	M: 1991 (90 to 93) F: 1992 (91 to 92)	2.5 (2.2 to 2.9) 1.9 (1.4 to 2.4)	M: 2002 (02 to 03) F: 2002 (02 to 03)
Estonia	M: no decline F: no decline	M: no decline F: no decline	No data	No data	F: 2000 (99 to 02)	F: 1.8 (−0.8 to 4.3)	M: 1994 (93 to 95) F: 1993 (93 to 94)	5.0 (4.4 to 5.7) 6.4 (4.9 to 5.9)	M: 2003 (03 to 03) F: 2003 (03 to 03)

The % mortality decline per year since start of mortality decline has been obtained by spline regression models. We restricted these models to have only one turning point in the year where mortality starts to decline. For the period between this year and 2005, the slope and its CI have been estimated.
F, female, M, male; NA, not applicable (no start of mortality decline within the study period).

The second alternative measure is the rate of mortality decline. Mortality decline could be calculated either on an annual basis or cumulatively over a fixed time interval and could be expressed either as a percentage of the original mortality level or in absolute terms. Sometimes large variations in speed of mortality decline are observed, and while these could also be due to variations in the speed with which population exposure to risk factors for amenable conditions changes, they are less likely to be biased by such confounding factors than national mortality levels. This indicator may reflect the speed with which healthcare systems achieve a substantial coverage of those in need. Table 2 shows some illustrative results for the four conditions, using percentage annual decline as an indicator. Spain had a relatively slow mortality decline for Hodgkin's disease and Germany for breast cancer, suggesting low speed of implementation of the respective innovations. This interpretation gains some support from a study of changes in amenable mortality in the different parts of the UK during a period when their relative funding levels were changing.¹³ It is also consistent with research on the USA showing a failure to achieve gains seen in other industrialised countries.²⁰

In conclusion, in the current state of knowledge, between-country differences in the absolute levels of mortality from amenable conditions cannot be recommended for use in routine surveillance of healthcare performance. These levels are likely to reflect the influence of many different factors and do not support simple judgements. It is more likely that alternative measures of mortality from amenable conditions, such as the timing of a favourable change in mortality or the scale and pace of mortality decline, are better measures of healthcare performance, although both require further investigation. The main lesson from this study is, however, just how scarce is the evidence on, first, the ability of many common healthcare innovations to reduce mortality (which is, after all, one of their main objectives) and, second, on the timing, diffusion and

What is already known on this subject

There is a great need for healthcare performance indicators. Mortality from conditions amenable to healthcare as been proposed as indicator, but its validity is unknown.

What this study adds

Changes in mortality from conditions amenable to healthcare do not often coincide with the introduction of healthcare innovations. Levels of mortality from conditions amenable to healthcare are likely to reflect incidence and risk factors.

Policy implications

Levels of mortality from conditions amenable to healthcare cannot be recommended for use in routine surveillance systems. Other measures of amenable mortality, reflecting change over time, may be better indicators of healthcare performance.

coverage of healthcare interventions in Europe. Given the importance of healthcare to the well-being of European citizens and to economic growth in Europe, it seems remarkable that so little effort has gone into collection of comparative data on the uptake of health innovations and the quality of care among Member States, a situation that contrasts markedly with the large amounts of data collected in other sectors. Until these deficiencies are remedied, it will not be possible to quantify precisely the contribution of healthcare to population health, even though it undeniably exists.

Acknowledgements We thank the other steering committee members (Katrin Lang and Wolfgang Hellmeier) as well as the advisory board (chair: Prof. W.W. Holland) of this project for their valuable contributions.

Contributors JPM conceptualised the study, obtained funding, supervised the analyses and wrote the paper. RH and CL performed the mortality trend analyses. BK performed the literature review and the Delphi study. IP collected data for the Netherlands and coordinated the study. GR and EJ calculated adjustment factors for coding changes and collected data for France. RW analysed differences in introduction of healthcare innovations and collected data for Sweden. KP collected data for Estonia. JA collected data for Spain. MM conceptualised the study and performed the literature review and the Delphi study. All authors commented on draft versions of the paper and approved the final version.

Funding This paper summarises the results of the project 'Avoidable mortality in the European Union: Towards better indicators for the effectiveness of health systems (AMIEHS)' which was co-funded by the European Union within the framework of the Public Health Program (project number 2007106). The funding body played no role in the design, writing or decision to publish this manuscript.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- Nolte E, McKee M, Wait S. Research on health, health system and service evaluation. In: Bowling A, Ebrahim S, eds. *Handbook of Health Research Methods: Investigation, Measurement and Analysis*. Buckingham: Open University Press, 2005:12–43.
- Rutstein DD, Berenberg W, Chalmers TC, et al. Measuring the quality of medical care. A clinical method. *New Engl J Med* 1976;**294**:582–8.
- Holland WW. The "avoidable death" guide to Europe. *Health Policy* 1986;**6**:115–17.
- Holland W. Avoidable death as a measure of quality. *Qual Assur Health Care* 1990;**2**:227–33.
- Charlton JR, Holland WW, Lakhani A, et al. Variations in avoidable mortality and variations in health care. *Lancet* 1987;**1**:858.
- Charlton JR, Hartley RM, Silver R, et al. Geographical variation in mortality from conditions amenable to medical intervention in England and Wales. *Lancet* 1983;**1**:691–6.
- Mackenbach JP, Bouvier-Colle MH, Jougle E. "Avoidable" mortality and health services: a review of aggregate data studies. *J Epidemiol Community Health* 1990;**44**:106–11.
- Holland WW, ed. *European Community Atlas of Avoidable Mortality*. Oxford: Oxford University Press, 1988.
- Holland WW, ed. *European Community Atlas of Avoidable Death*. 2nd edn. Oxford: Commission of the European Communities/Oxford University Press, 1991.
- Jozan P, Prokhorshkas R. *Atlas Of Leading And 'Avoidable' Cause Of Death In Countries of Central and Eastern Europe*. Budapest: Hungarian Central Statistical Office Publishing House, 1997.
- Department of Health. *The NHS Outcomes Framework 2012/13*. London: Department of Health, 2010.
- Commonwealth Fund. *State Scorecard*. 2011. <http://www.commonwealthfund.org/Maps-and-Data/State-Data-Center/State-Scorecards.aspx> (accessed 22 Dec 2011).
- Desai M, Nolte E, Karanikolos M, et al. Measuring NHS performance 1990–2009 using amenable mortality: interpret with care. *J R Soc Med* 2011;**104**:370–9.
- Andreev EM, Nolte E, Shkolnikov VM, et al. The evolving pattern of avoidable mortality in Russia. *Int J Epidemiol*. 2003;**32**:437–46.
- Albert X, Bayo A, Alfonso JL, et al. The effectiveness of health systems in influencing avoidable mortality: a study in Valencia, Spain, 1975–90. *J Epidemiol Community Health* 1996;**50**:320–5.
- Nolte E, McKee M. Variations in amenable mortality—trends in 16 high-income nations. *Health Policy* 2011;**103**:47–52.
- Nolte E, McKee M. Measuring the health of nations: analysis of mortality amenable to health care. *BMJ* 2003;**327**:1129.
- Nolte E, Scholz R, Shkolnikov V, et al. The contribution of medical care to changing life expectancy in Germany and Poland. *Soc Sci Med* 2002;**55**:1905–21.
- Logminiene Z, Nolte E, McKee M, et al. Avoidable mortality in Lithuania: 1991–1999 compared with 1970–1990. *Public Health* 2004;**118**:201–10.
- Nolte E, McKee CM. Measuring the health of nations: updating an earlier analysis. *Health Affairs* 2008;**27**:58–71.
- Mackenbach JP. Health care expenditure and mortality from amenable conditions in the European Community. *Health Policy* 1991;**19**:245–55.
- OECD. *Health Care Systems: Efficiency And Policy Settings*. Paris: OECD, 2010.
- Mackenbach JP, Stronks K, Kunst AE. The contribution of medical care to inequalities in health: differences between socio-economic groups in decline of mortality from conditions amenable to medical intervention. *Soc Sci Med* 1989;**29**:369–76.
- Kunitz SJ, McKee M, Nolte E. State political cultures and the mortality of African Americans and American Indians. *Health Place* 2010;**16**:558–66.
- Baburin A, Lai T, Leinsalu M. Avoidable mortality in Estonia: exploring the differences in life expectancy between Estonians and non-Estonians in 2005–2007. *Public Health*. 2011;**125**:754–62.
- Stirbu I, Kunst AE, Bopp M, et al. Educational inequalities in avoidable mortality in Europe. *J Epidemiol Community Health* 2010;**64**:913–20.
- Stirbu I, Kunst AE, Bos V, et al. Differences in avoidable mortality between migrants and the native Dutch in The Netherlands. *BMC Public Health* 2006;**6**:78.
- Nolte E, McKee M. *Does Health Care Save Lives? Avoidable Mortality Revisited*. London: Nuffield Trust, 2004.
- Plug I, Hoffmann R, Mackenbach J. *AMIEHS Avoidable Mortality in the European Union: Towards Better Indicators for the Effectiveness of Health Systems. Volume 1: Final Report*. Rotterdam: Department of Public Health, Erasmus MC, 2011.
- Rey G, Aouba A, Pavillon G, et al. Cause-specific mortality time series analysis: a general method to detect and correct for abrupt data production changes. *Popul Health Metr* 2011;**9**:52.
- Black N, Murphy M, Lamping D, et al. Consensus development methods: a review of best practice in creating clinical guidelines. *J Health Serv Res Policy* 1999;**4**:236–48.
- Lundgren JD, Phillips AN, Vella S, et al. Regional differences in use of antiretroviral agents and primary prophylaxis in 3122 European HIV-infected patients. EuroSIDA Study Group. *J Acquir Immune Defic Syndr Hum Retrovirol* 1997;**16**:153–60.
- Fischl MA, Richman DD, Grieco MH, et al. The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. A double-blind, placebo-controlled trial. *New Engl J Med*. 1987;**317**:185–91.
- Concorde Coordinating Committee. Concorde: MRC/ANRS randomised double-blind controlled trial of immediate and deferred zidovudine in symptom-free HIV infection. *Lancet* 1994;**343**:871–81.
- Delta Coordinating Committee. Delta: a randomised double-blind controlled trial comparing combinations of zidovudine plus didanosine or zalcitabine with zidovudine alone in HIV-infected individuals. *Lancet* 1996;**348**:283–91.
- Lin DY, Fischl MA, Schoenfeld DA. Evaluating the role of CD4-lymphocyte counts as surrogate endpoints in human immunodeficiency virus clinical trials. *Stat Med* 1993;**12**:835–42.
- Phillips AN, Grabar S, Tassie JM, et al. Use of observational databases to evaluate the effectiveness of antiretroviral therapy for HIV infection: comparison of cohort studies with randomized trials. EuroSIDA, the French Hospital Database on HIV and the Swiss HIV Cohort Study Groups. *AIDS* 1999;**13**:2075–82.
- Munoz A, Gange SJ, Jacobson LP. Distinguishing efficacy, individual effectiveness and population effectiveness of therapies. *AIDS* 2000;**14**:754–6.
- Hamers FF, Downs AM. The changing face of the HIV epidemic in western Europe: what are the implications for public health policies? *Lancet*. 2004;**364**:83–94.
- Devita VT Jr, Serpik AA, Carbone PP. Combination chemotherapy in the treatment of advanced Hodgkin's disease. *Ann Intern Med* 1970;**73**:881–95.
- Frei E 3rd, Gehan EA. Definition of cure for Hodgkin's disease. *Cancer Res* 1971;**31**:1828–33.
- Reece DE, Phillips GL. Intensive therapy and autotransplantation in Hodgkin's disease. *Stem Cells* 1994;**12**:477–93.
- Duffy SW, Tabar L, Olsen AH, et al. Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the breast screening programme in England. *J Med Screen* 2010;**17**:25–30.
- Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2011;(1):CD001877.
- Autier P, Boniol M, Gavin A, et al. Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. *BMJ* 2011;**343**:d4411.
- McGuire WL. Hormone receptors: their role in predicting prognosis and response to endocrine therapy. *Semin Oncol* 1978;**5**:428–33.
- Hamilton M, Thompson EM, Wisniewski TK. The role of blood-pressure control in preventing complications of hypertension. *Lancet* 1964;**1**:235–8.
- Stroke Unit Trialists' Collaboration. Collaborative systematic review of the randomised trials of organised inpatient (stroke unit) care after stroke. *BMJ* 1997;**314**:1151–9.
- The Cardiac Arrhythmia Suppression Trial (CAST) Investigators. Preliminary report: effect of encainide and flecainide on mortality in a randomized trial of

- arrhythmia suppression after myocardial infarction. *New Engl J Med* 1989;**321**:406–12.
50. **The International Chronic Granulomatous Disease Cooperative Study Group.** A controlled trial of interferon gamma to prevent infection in chronic granulomatous disease. *New Engl J Med* 1991;**324**:509–16.
51. **Yudkin JS,** Lipska KJ, Montori VM. The idolatry of the surrogate. *BMJ* 2011;**343**:d7995.
52. **Britton A,** McKee M, Black N, *et al.* Choosing between randomised and non-randomised studies: a systematic review. *Health Technol Assess* 1998;**2**:i–iv, 1–124.
53. **Verdecchia A,** Guzzinati S, Francisci S, *et al.* Survival trends in European cancer patients diagnosed from 1988 to 1999. *Eur J Cancer* 2009;**45**:1042–66.
54. **Hamdy RC.** Hypertension: a turning point in the history of medicine...and mankind. *South Med J* 2001;**94**:1045–7.
55. **Anon.** MRC trial of treatment of mild hypertension: principal results. Medical Research Council Working Party. *Br Med J (Clin Res Ed)* 1985;**291**:97–104.
56. **Amery A,** Birkenhager W, Brixko P, *et al.* Mortality and morbidity results from the European Working Party on High Blood Pressure in the Elderly trial. *Lancet* 1985;**1**:1349–54.
57. **Hansson L,** Lindholm LH, Ekblom T, *et al.* Randomised trial of old and new antihypertensive drugs in elderly patients: cardiovascular mortality and morbidity the Swedish Trial in Old Patients with Hypertension-2 study. *Lancet* 1999;**354**:1751–6.
58. **Corrao G,** Nicotra F, Parodi A, *et al.* Cardiovascular protection by initial and subsequent combination of antihypertensive drugs in daily life practice. *Hypertension* 2011;**58**:566–72.
59. **Barker WH,** Mullooly JP, Linton KL. Trends in hypertension prevalence, treatment, and control: in a well-defined older population. *Hypertension* 1998;**31**:552–9.