Policy and Practice

On measuring inequalities in health

Michael Wolfson[1] & Geoff Rowe[2]

Abstract In a recent series of papers, Murray et al. have put forward a number of important ideas regarding the measurement of inequalities in health. In this paper we agree with some of these ideas but draw attention to one key aspect of their approach — measuring inequalities on the basis of small area data — which is flawed. A numerical example is presented to illustrate the problem. An alternative approach drawing on longitudinal data is outlined, which preserves and enhances the most desirable aspects of their proposal. These include the use of a life course perspective, and the consideration of non-fatal health outcomes as well as the more usual information on mortality patterns.

Keywords Health status; Socioeconomic factors; Mortality/trends; Longevity; Outcome assessment(Health care); Analysis of variance; Longitudinal studies; Models, Statistical (*source: MeSH*).

Mots clés Etat sanitaire; Facteur socioéconomique; Mortalité/orientations; Longévité; Evaluation résultats (Santé); Analyse variance; Etude longitudinale; Modèle statistique (*source: INSERM*).

Palabras clave Estado de salud; Factores socioeconómicos; Mortalidad/tendencias; Longevidad; Evaluación de resultado (Atención de salud); Análisis de varianza; Estudios longitudinales; Modelos estadísticos (*fuente: BIREME*).

Voir page 560 le résumé en français. En la página 560 figura un resumen en español.

Introduction

In a series of recent papers, Murray et al. (*1–3*) have put forward a number of important ideas regarding the measurement of inequalities in health. The third paper, by Gakidou et al., gave rise to some debate (*4*). Subsequently, WHO published the *World health report 2000* (*5*) which began putting these ideas into practice. The purpose of the current paper is to explore some of them and offer a number of suggestions.

The core question is what should be meant by inequalities in health. Gakidou et al. "define health inequality to be variations in health status across individuals in a population... which allows us to perform cross-country comparisons and study the determinants of health inequality... WHO is interested in measuring health inequality as a distinct dimension of the performance of health systems" (*3*). They then go on to focus on the scalar "healthy lifespan" as the individual-level health measure of interest.

As they observe, healthy lifespan can be estimated by a modified sequence of age-specific survival probabilities (or more precisely, a mortality hazard function). Their key modification of this standard notion is to combine risks of mortality with risks of disease onset and progression, characterized by an age profile of a summary health index, in order to form a "health survivorship function" which in turn underlies estimates of healthy lifespan.

Later in the paper, Gakidou et al. observe that the healthy lifespans of the members of a living population are intrinsically unobservable (since their lives have still to finish unfolding, and we cannot see the future). But, they claim, "... the distribution of health risks can be reasonably approximated...", and the distribution of these risks determine (subject to elements of chance) the distribution of ultimately realized healthy lifespans. As a result, any feasible programme of measurement of health inequalities of this sort requires the measurement of health risks.

They appeal in general to methods based on subdividing the population into groups. "Inevitably", they observe, "this will underestimate the distribution of health expectancy (i.e. healthy lifespan) in the population even if the groups are perfectly non-overlapping in terms of their individual health expectancies. The more refined the groupings the more we will approximate the true underlying distribution of health expectancy. Small area analyses hold out the promise of being one of the most refined methods for revealing the underlying distribution of health expectancy in a population."

This focus on health risks, and how they can be observed, raises serious measurement issues. Our

[1] Assistant Chief Statistician, Statistics Canada, Analysis and Development Field, 26-K, R.H. Coats Building, Ottawa, K1A 0T6, Canada (email: wolfson@statcan.ca). Correspondence should be addressed to this author.

[2] Senior Analyst, Socioeconomic Modelling Group, Statistics Canada, Ottawa, Canada.

basic concern is that efforts to describe health inequalities in terms of health survivorship *risks*, based on inferences from small area data, are highly problematic, desirable as the approach may be. However, alternative approaches, based on longitudinal data, are much more promising.

Defining inequalities

We begin with two basic points.

First, inequality is a property of a population. It assumes that each member of the population has an attribute, "health", that is unidimensional and measurable on a cardinal scale, such as income.

Second, the term "health inequalities" is used in two quite different ways. Formally, we can think of health as one among a number of relevant attributes of an individual, so that when thinking about inequality of health in a population, it can be represented by a multivariate or multidimensional joint distribution. These dimensions include health as well as other attributes such as income, gender, ethnicity, and education. One way to think of health inequality is then in terms of a univariate (or unconditional or marginal) distribution — some individuals have high health, others have low health, and health inequality is intended to indicate the extent of dispersion of "health" within the population. The second concept is in terms of a bivariate (or conditional) distribution — whether those with high health also have high income, for example.

Gakidou et al. focus on the univariate or marginal approach, much like analyses of income inequality. In this case, the focus is on a single "health" attribute margin in the multivariate joint distribution characterizing a population. The main question is how to characterize the "shape" of the univariate density function describing how "health" is distributed in the population. In this case, there is an extensive literature in economics that can inform the choice of inequality or polarization measure (6, 7).

In the other bivariate or conditional case, the focus of attention is on at least two variables for each member of a population: "health", and some measure or indicator of socioeconomic status (SES), such as income. The concerns in this literature are the pervasive observation in many populations of a strong positive correlation of income and health. We agree with Braveman et al. (4) that the assessment of this kind of bivariate distribution (or the conditional distribution of health given income, or SES more generally) is fundamental to the discourse on health inequalities.

Gakidou & King (8) give the misleading impression that the choice is between invididual-level and group-level approaches to health inequality measurement, with individual-level inequality in their terms equivalent to what we are calling the univariate approach. Group-level inequality in their terms is the only concept that deals with socioeconomic factors. However, it is entirely feasible to have both univariate and bivariate notions of health inequality at the individual level — in other words based on individual microdata from a representative sample. They also suggest that the individual-level approach is "more of a purely public health perspective", implying that a bivariate or group level approach is not central to public health. This is clearly a contentious claim.

In order to study health inequalities in the bivariate (or multivariate) sense, it is necessary to have a scalar index of individual health analogous to income. Consequently, it is also logically possible and ultimately practical to study univariate health inequalities as well.

We return to these considerations of univariate and bivariate (or marginal and conditional) health inequalities later. The prior issue raised by Gakidou et al. is how to construct the univariate index of each individual's health. The simplest approach is to conduct a census or a survey of a representative sample of individuals in a population, and to elicit from each one — preferably by means of a well-structured series of questions and a sophisticated methodology, such as the McMaster Health Utility Index (9) — a univariate cardinal index of their health status at the moment.

However, Gakidou et al. want to be more ambitious than this. They want instead to base assessments of health inequality on some notion of inequality in individuals' life chances — i.e. to include the entire life cycle, and also to include in their concept of life chances both length of life and health-related quality of life — in short, *health* expectancy.

We certainly agree with them about the great importance of combining both length of life and health-related quality of life during a person's lifetime. However, for this specific discussion, the inclusion of health status, or non-fatal health outcomes, complicates matters unnecessarily. As a result, we assume that all that matters in assessing health inequality is length of life. Equivalently, we assume that each individual's health-related quality of life is perfect from birth to the moment of death.

Measuring univariate inequality with small area data

While the notion of individual-level life (health) expectancy is highly appealing as an index of each individual's health, life chances are not directly observable at the individual level. Moreover, heterogeneity in life chances is not easily observable even at the level of populations. This can be demonstrated by a simple numerical example. Suppose we have two countries, with large populations. In country A, everyone has an identical mortality risk of 0.5 over a given time span (e.g. a decade), i.e. perfect equality in health as defined by Gakidou et al., bearing in mind that for simplicity, we are ignoring health status or, equivalently, assuming that it is perfect up to the moment of death. In country B, however, half the population has a mortality risk of 0.25, while the other

half has a mortality risk of 0.75 over the same time span. This is clearly a situation of high univariate health inequality. Nevertheless, by construction, both countries have the same overall mortality rate. This thought experiment in part recalls the discussion by Vaupel et al. of heterogeneous frailty (10). This should be an ideal case for the use of something like small area data to reveal the much higher inequality in health in country B as compared to country A, as proposed by Gakidou et al. and illustrated in Fig. 2.3 of the *World health report 2000* (5).

Now suppose that we do something even better than the standard combination of the census and death registration statistical systems typically used to estimate small area mortality rates. Instead of only juxtaposing death counts and the population at risk, we conduct surveys in both of these hypothetical countries, and simply collect individual identification. We then follow all survey respondents prospectively over a decade to observe whether they die. As a hypothetical numerical experiment, the resulting pattern of deaths is easily simulated, based on the posited equal and unequal mortality rates in the two countries. We can now apply the suggestion of Gakidou et al. and partition each sample into subgroups, and compare observed mortality rates across these subgroups.

One might expect that the variance in observed mortality rates across the subgroups in country B would be significantly higher than in country A, thereby revealing its much greater inequality in underlying health (actually mortality) risks. However, Table 1 shows negative results.

Table 1 has eight columns in four pairs (1 through 4) corresponding to four replicates of the sample for each hypothetical country (A and B). Table 1 also has seven rows, corresponding to different ways of randomly partitioning the samples, in this case 50 000 in each country, into groups of equal size. These range from 2000 groups (implying 25 observations per group) to 25 groups (implying 2000 observations per group). These partitions are analogous to small area observations ranging in size from census tracts to the state or province level in each country. Within each cell of the table, the standard deviation of the mortality rates across the partitions is shown.

Perhaps surprisingly, there are no observable differences between the two countries for any level of (the analogue of) geographic aggregation, or any of the replicates.[a] This kind of numerical simulation was

---

[a] We thank an anonymous reviewer for showing that a similar result can be derived algebraically, assuming a large enough sample. If the mean probability of mortality is the same in both a homogeneous and a heterogeneous population, then the expected values of the variances of mortality risk will also be identical (within the simplified framework of our simulations, e.g. binomial mortality risk and no age dependence). This is a more general version of the conclusion we reached through simulation. The simulations are still useful in demonstrating that, even without a large sample assumption, we should not expect to find excess variability in mortality from thoroughly mixed heterogeneous populations.

Table 1. **Hypothetical standard deviations of mortality rates**

Number of groups	A1	B1	A2	B2	A3	B3	A4	B4
2000	0.0992	0.0992	0.1004	0.1007	0.1015	0.0996	0.1013	0.1022
1000	0.0704	0.0718	0.0728	0.0709	0.0699	0.0693	0.0725	0.0734
500	0.0501	0.0487	0.0490	0.0503	0.0493	0.0519	0.0509	0.0494
250	0.0338	0.0365	0.0349	0.0341	0.0369	0.0361	0.0365	0.0323
100	0.0213	0.0228	0.0208	0.0213	0.0239	0.0209	0.0224	0.0220
50	0.0189	0.0130	0.0164	0.0149	0.0155	0.0140	0.0201	0.0131
25	0.0133	0.0106	0.0140	0.0120	0.0101	0.0126	0.0104	0.0093

also run using a sample size of 1 million, as well as with mortality risks of one tenth of the levels in the simulations shown (i.e. 0.05 versus a 50–50 mixture of 0.025 and 0.075). There was still no difference between A and B.

An intuitive way of accounting for this result is as follows. Let us first consider two urns, A containing a large number of identical gray balls, and B containing a 50–50 mixture of pure black and pure white balls. If we drew any sample of balls from urn A, the colour of the set of balls would be the same shade of gray. However, in a sequence of random samples of balls from urn B, standing at a distance so the individual colours merge, some samples would be lighter gray, and others darker gray. So far, it would seem that our urn analogy supports the Gakidou et al. approach. However, the colours of balls drawn from the urns correspond to the mortality *risks*, not to observed mortality rates. The reason we do not see an analogous difference in Table 1 between country A and country B is that the results in Table 1 are not observations of mortality risks per se. Rather they are observations of the outcomes of stochastic processes whose parameters are the mortality risks in question. The "noisiness" of these stochastic mortality processes is essentially obscuring the very real (by design) differences between the two (stochastic) mortality processes.

The objective of the approach of Gakidou et al. is to define health inequality in terms of risks. But these risks are inherently unobservable; only their impacts can be seen, and the numerical example shows that differences in impact, even when the risk distributions are very different, can be invisible.

This result has strong implications. If we are unable to detect even this blatant kind of univariate health inequality in populations where measurement is simple and perfect, what chance is there in the much more complex realities of imperfect measurement, more detailed indicators of health, and more subtle kinds of heterogeneity?

One reaction to this numerical thought experiment is that it necessarily obscures the underlying heterogeneity (inequality) in mortality risks because the sample has been *randomly* partitioned to form the analogue of geographic areas. The alternative would be to partition the samples by some "marker" of mortality risk. The use of such a

marker, whether at the individual or the geographic area level, however, would breach the objective of revealing heterogeneities in individuals' health that are *independent* of any systematic factors. This independence is required in the Gakidou et al. approach, since the focus is explicitly on health inequality in the univariate sense, unlike studies such as that of Ross et al., where the focus is on the correlation of health and SES at the level of subnational geographic regions (*11*).

Related conceptual and methodological difficulties arise in using mothers rather than geographic areas as a method for grouping individuals (infants in this case) to explore heterogeneities in mortality risk. A key problem is if the chances of being in the sample (mothers by number of children ever born) and risks are related. An otherwise unobserved mortality risk (e.g., from contaminated water supplies in certain villages) may affect both the survival of live-born children and the extent to which mothers successfully carry their pregnancies to term. If this happened, larger "lower mortality" families would be observed more frequently. This in turn would bias estimates of heterogeneities ("inequalities") in infant mortality rates derived from samples of mothers from whom were collected data on the numbers of surviving children compared to the number of births. At the extreme end of risks, an important proportion of women might remain childless. In other words, the sample would be reduced to zero in just those cases where the greatest mortality risks might have been observed.

From the viewpoint of measuring *univariate* health inequalities, these are unfortunate conclusions. Measuring health inequalities in terms of heterogeneities in health *risks* using mother-specific infant mortality patterns, and small geographic area data as a form of partitioning — methods suggested by Gakidou et al. (*3*), with the former applied as one of the key international rankings in the *World health report 2000* (*5*) — appears not to work.

Of course, small area mortality analysis remains intrinsically important, and highly relevant in a bivariate or multivariate context for assessing health inequalities. Indeed, we have used precisely such analyses to raise fundamental questions about the determinants of health in Canada and the USA (*11*).

Cross-sectional inequality

There are several alternative approaches for assessing health inequalities in the univariate sense intended by Gakidou et al. The simplest and most strightforward is what we might call direct or cross-sectional. Here, we continue with the current more limited but widespread approach where a representative sample of the living population is surveyed and asked structured and standardized questions about their current health status. The distributions of one or other type of response can then be analysed univariately using standard statistical methods (e.g. means, variances, quantiles), or the kinds of inequality and polarization measures developed in the income inequality literature.

This approach is most likely to yield rather unsurprising results — that some people are sicker than others. And levels of health will probably be highly correlated with age, and therefore not very informative. However, within age and sex groups, it would be of interest to track trends in the prevalence of excellent or very good health in comparison to fair or poor health, for example, in a population over time, or across populations in different countries.

In addition, if the surveys also captured socioeconomic characteristics of the respondents (e.g. income, education), this straightforward approach would support bivariate analyses of the kinds advocated by Braveman et al. (*4*). One could track the extent to which higher income individuals were also healthier, and whether this association was stronger in one country than another.

However, this approach does not consider individuals over their lifetimes, one of the main objectives of the Gakidou et al. analysis. This necessarily requires longitudinal data, or piecing together data from different age groups to form a life-cycle or cohort perspective, or both.

Mortality inequality with longitudinal data

We can begin by considering a second broad approach. It draws on longitudinal data analysis, and is more ambitious in terms of its data requirements. In order to illustrate this group of measures, let us assume the following conditions for a number of countries: a population census is conducted every 10 years which collects SES data, among other things; there is complete and accurate death registration; and all deaths are linked back to their corresponding census record. As a result, it is possible on a regular basis to estimate a mortality hazard as a function of age, sex, and a range of baseline SES covariates. Given the institutional and legal possibility of undertaking this kind of data linkage, most developed countries would have more than adequate sample sizes for very rich mortality hazard estimation.

With these data, a range of health (more simply, mortality) inequality measures is possible. The most direct (multivariate) indication of health inequality, in this case, would simply be the extent to which the SES covariates were statistically significant. If these covariates were not significant, we could conclude that the health dynamics at work in the country's population, which include not only health care but a range of other health determinants, were "colour blind" to SES, so that individuals' health risks were independent of socioeconomic status.

However, it is likely that the SES covariates would be statistically significant. One could then construct a further set of indicators. In order to develop the basic idea, we can start with the easiest, though not the most reasonable measure — the dispersion in estimated mortality *risks* for the population given the actual dispersion in its SES covariates. In this case, we would start with the estimated hazard function, and then evaluate it for each individual's vector of covariates observed in the population. We could then use the resulting set of predicted hazards (e.g. point estimates of the five-year mortality risks) as the population distribution of individual-level health risks. In other words, we take the estimated equation plus a sample of individuals' SES characteristics. Then, one at a time, we plug these SES characteristics into the equation and compute each individual's chances of dying over the next five years.

The resulting distribution of these predicted hazards could then be summarized by a statistic such as the variance, an inequality measure such as the Gini coefficient, or a polarization measure. This is surely the most straightforward indication of the overall inequality in a population's health risks. However, such a measure would confound the extent, for example, of income inequality, with the strength of the SES association with mortality risk. In other words, this approach mixes together the dispersion in incomes from the observed distribution of SES characteristics in the population, with the magnitudes of the coefficients on the SES characteristics in the equation for mortality risks. This approach would also understate inequality insofar as the hazard regression did not fit the data exactly.

As a result, a more informative (conditional) measure could be defined, drawing on an analogue to age standardization. We could estimate separate mortality hazard functions for the same population at several points in time, or for several different populations, such as different countries. In other words, we have two or more estimated equations relating SES to mortality risk, one for each of the time periods or populations we wish to compare. We could also define a "reference population" distribution of SES covariates, analogous to a reference distribution of the population by age and sex in age standardization.

It is then straightforward to evaluate estimated mortality risks for a large representative sample of points from this reference distribution of SES covariates, for each of the separately estimated hazard functions over time or across populations. We just have to plug the same standard or reference sample of SES characteristics into the mortality risk equation for country A, then for country B, and so on. Each resulting distribution of mortality risks is the outcome of interest. For the same posited reference distribution of SES covariates, we can imagine one country having a narrower distribution of mortality risks than another. In this case, the former country could be said to have less health inequality, as a result of having a weaker link between the standardized distribution of SES covariates and subsequent mortality.[b]

If we are still interested in health inequalities in the univariate or marginal sense preferred by Gakidou et al., rather than in the conditional sense just described, we might imagine instead focusing on the distribution of "residuals" to the hazard regressions in each period or each country. In other words, "controlling for" a vector of SES covariates, we could see whether the hazard regressions for one country fit the data better than another. In effect, we are seeking to isolate the unobserved heterogeneity in mortality risks, after first removing the systematic part of the variations associated with an agreed and commonly defined set of SES covariates. This variation in mortality risks that is left over — the unobserved heterogeneity — is then identified with some sort of "intrinsic" health. This seems to be the underlying concept of univariate health being sought by Gakidou et al. If this leftover variation is smaller in country A than in country B (even if income inequality is higher in country A, or even if the strength of the association between income and mortality risk is higher in country A), we might then claim that (non-systematic) univariate health inequality is lower in A than in B.

However, this is a weak strategy empirically because residual variation is likely also to be influenced by the omission of important covariates as well. To give a simple example, suppose that the regressions include age and income as covariates, but not education. If education is also strongly and independently predictive of subsequent mortality, as is typically the case, then the omission of this variable in the regressions would make it impossible to ascribe any differences in the goodness of fit of the regressions solely to differences in (non-systematic) univariate or "intrinsic" health inequality.

Using an estimated synthetic cohort

The largest source of dispersion in mortality risks would be associated with age. This leads to a third alternative set of measures — one that combines longitudinal data analysis with life table concepts. The basic idea is to form a cohort of complete individual health life cycles, and then use this estimated synthetic cohort as the basis for computing a variety of health inequality measures.

In the first and most basic instance, the fact that mortality risks are predominantly associated with age could be accommodated by using only age as a covariate in the hazard regressions — indeed, simply constructing mortality rates — and then using these

---

[b] There is a question of the likely auto-correlation of these SES covariates over the life cycle. The simplest approach in this case would be to posit some standard scenario. But these auto-correlations too could be measured using longitudinal surveys. Their dynamics, as time-varying covariates, would then have to be simulated as well. This is precisely what POHEM is designed to do.

rates to construct a period life table. If the comparison is across populations, the country with the lowest inequality would then be the one with the lowest variance in age-at-death, or, equivalently, life length, i.e. a "rectangularized" survival curve. This is essentially the approach used by Le Grand (12).

A considerably more sophisticated life table style or cohort approach would be needed to come close to Gakidou et al.'s notion of health expectancy (but still only life expectancy for now, to simplify the discussion) as the basis for measuring health inequality. Again we can assume no more than a series of decennial population censuses with standardized, internationally comparable SES questions, and complete mortality follow-up. To do health expectancy, we would also need longitudinal follow-up of health status, not just mortality follow-up. But instead of positing some general reference distribution of SES covariates, which by its nature will include age, we build up a statistical description of mortality risks over the full life course. We do this by chaining together in sequence a series of age-specific mortality hazards, say by five-year age groups plus, as is conventional in abridged life tables, a separate infant mortality rate. Additionally, we can assume that the hazard regressions are fit in a way that also yields an estimated distribution of unobserved heterogeneity.

We start with a synthetic cohort of, say, 1 million individuals at birth, as in the radix of a life table. We then expose these newborn individuals, one at a time, to the observed mortality risks in the 0 to 1 year age group, conditional on a random draw from the posited distribution of SES covariates applicable to this age group (e.g. based on parents' SES). The process is then repeated for the survivors to age 1 year. Each is exposed to the mortality risks estimated for the 1 to 5 year age group, again conditional on a random draw from their distribution of SES covariates. Then the process is applied to the survivors entering the 5 to 10 year age group, and so on.

We could also construct this synthetic cohort using estimates of the distribution of unobserved heterogeneity, which for convenience we can refer to as "frailty" or "resilience". Each individual at the start of the simulation used to construct the cohort life table would be tagged with a frailty or resilience index, drawn from the estimated distribution. This index would then also be plugged into the estimated hazard equation to determine, in each age group, that individual's chance of dying.

Note that standard multi-state life table methods are impractical for constructing synthetic cohorts based on these kinds of complex multivariate transition probability functions. Such life tables would ultimately need millions of columns to represent the combinatorial explosion of possible states. Fortunately, microsimulation methods such as those developed for Statistics Canada's POpulation HEalth Model (POHEM) are more than adequate (13).

These processes will generate distributions of life lengths (or in a fuller analysis, health-adjusted life lengths), conditional on the posited age- (and sex-) specific distributions of SES covariates.[b] It will also be possible to construct two versions, one with and the other without an explicit account of the estimated pattern of frailty or resilience (i.e. unobserved heterogeneity). The latter distribution would be narrower, and the difference can reasonably be identified with the notion of Gakidou et al. of univariate health inequality.

Given the assumed data, the set of hazard regressions, and the microsimulation apparatus for constructing life table measures as just described, at least in thought experiment mode, we now have the raw material to consider the concept of health inequality both in a full life-cycle framework and in terms of health risks, exactly as suggested by Gakidou et al.

We have just noted what seems to be the most appropriate definition of univariate health inequality in this analytical context, based on the difference in the dispersion in expected life lengths depending on whether or not unobserved heterogeneity is taken into account. It is also possible to extend this framework to estimate a set of age- (and sex-) specific hazards for country A and for country B. It would also be possible to estimate the distributions of SES covariates for the two countries, and to posit some sort of reference SES covariate distribution. It would then be possible to construct various distributions of expected life lengths for the two countries (or time periods in the same country).

More precisely, we now have the following ingredients: four sets of hazard regressions, for populations A and B, and with and without explicit estimates of unobserved heterogeneity — namely $\{H'_A\}$ and $\{H'_B\}$ with, and $\{H_A\}$ and $\{H_B\}$ without unobserved heterogeneity, and two sets of SES covariates (SES_A and SES_B) plus a reference set (SES^*, say). For any given $\{H\}$ and SES, the microsimulation apparatus sketched above generates a distribution of life lengths L based on the set of hazard regressions $\{H\}$ and SES.

To begin, if the distribution of L ($\{H'_A\}$, SES_A) is more dispersed than L ($\{H'_B\}$, SES_B), this is an overall indication that A has a more unequal distribution of life lengths than B. However, this seems not to be Gakidou et al.'s notion of univariate or marginal health inequality (Chris Murray, personal communication).

Of course, we can compare SES_A with SES_B to see which country is more unequal in terms of SES. We can also compare (by inspecting coefficients) $\{H_A\}$ and $\{H_B\}$ to see which country has the strongest association between SES and mortality. But this would be rather tedious and complex. A summary approach would compare the two distributions of hypothetical life lengths, L ($\{H_A\}$, SES^*) with L ($\{H_B\}$, SES^*) using a standard or reference SES distribution, SES^*. The country with the widest distribution of expected life lengths would then be the one with the strongest association between mortality and SES — an indication of health inequality in the bivariate or conditional sense.

Finally, the best way to approach univariate health inequality in the sense apparently desired by Gakidou et al. in this context would be to compare the difference between L ($\{H'_A\}$, SES*) and L ($\{H_A\}$, SES*) with the difference between L ($\{H'_B\}$, SES*) and L ($\{H_B\}$, SES*). The country with the greatest difference in the distribution of expected life lengths with and without an explicit account of unobserved heterogeneity — our estimate of frailty or resilience — would be the one with the most inequality in health in the univariate sense, over and above any systematic relation with SES factors.

Extensions and conclusions

We see merits in measures of health inequality in both this univariate or marginal sense, and in the bivariate or conditional sense. However, it is quite possible that the general public and those concerned with public policy will be most interested in the latter notion of health inequality — the "social patterning" of variations in health. One reason is that the covariates give some clues as to the causes of inequality, and possible areas of intervention.

Of course, all of the longitudinal or life cycle indicators just sketched are quite ambitious with respect to current data availability. At the same time, they point in a different direction than Gakidou et al. for health data development. Our approach places primary emphasis on mortality follow-up specifically, and on longitudinal data collections more generally. This is in contrast to Gakidou et al.'s idea of developing small area mortality rate data.

Small area data are certainly of intrinsic interest, and they can generate hypotheses insofar as they show significant patterns such as correlations between life expectancy and unemployment rates (14) or between mortality rates and income inequality (11). However, small area mortality data also suffer from a number of problems. Particularly with small areas in terms of population, migration may be a problem, if the circumstances or character of the decedents in a given place are not representative of the population currently living there. The smaller numbers of deaths is also likely to lead to problems of statistical stability. And the likelihood that place and SES factors are correlated means that any variations are likely to be systematically related to SES, and therefore not indicative of the kinds of univariate heterogeneities Gakidou et al. are seeking to assess.

In addition, geographic areas for which data are available are usually defined for political or administrative purposes and, as such, they will tend to blur mortality differentials — since geographic areas defined to maximize mortality (and health) homogeneity would probably have different boundaries. Even with significant observed differences in mortality rates across small areas, these "ecological" differences will typically underestimate individual-level mortality heterogeneity because various differ-

ences between persons are likely to be replicated in each unit of any geographic partition of the population (for example, there will often be a mixture of smokers and non-smokers in each unit). To be easily compared, geographic areas should be similar in more ways than simply population size. For example, the population density of areas may make a difference (risks from infectious disease, traffic accidents, etc.). These examples amount to a range of serious problems with strategies building on small area mortality data.

On the other hand, development of mortality follow-up, and of longitudinal data more generally, seems much more promising. Let us continue with the thought experiment based on a census-mortality follow-up plus microsimulation-based life table approach as just sketched. The key idea is that once a health inequality indicator is based on a model, it is quite straightforward to pose and rigorously answer "what if" questions. A large class of such questions serves to generalize the epidemiological notion of attributable fraction. In effect, we construct two estimates of the distribution of expected life lengths. The first or "base case" uses the observed or posited standard distribution of SES covariates. The second "snips" the connection of one particular covariate to the estimated morality outcome, and hence the estimated distribution of life lengths.

For example, the coefficients relating income to mortality in the hazard regressions for each age could be set arbitrarily to zero, essentially "snipping" the (direct) connection between income and mortality. The resulting differences between the base case and the "snipped" hypothetical distribution of life lengths then provides an indication of the importance of the snipped SES covariate in generating health inequalities in the given society — e.g. the amount of health (actually life length) inequality attributable to income inequality. A version of this kind of analysis is given at the end of Wolfson's article (7) where, as a rough estimate, it was concluded that about one-fifth of the SES gradient in life expectancy in Canada could be attributed to differences in smoking prevalences by SES.

Gakidou et al. raise a significant set of questions. Theirs is a very important objective: to seek a measure of health, for purposes of inequality analysis, that is comprehensive, combining both life length and health status over the life course, and defining these in terms of risks. The main problem is in the proposed methods for estimating these risks. They suggest the use of small area data. However, as shown above, this strategy will not work. However other methods, based on longitudinal follow-up data combined with micro-simulation-based life table analysis can substantially meet their ultimate objectives. ∎

Conflicts of interest: none declared.

Résumé

A propos de la mesure des inégalités de santé

Dans une récente série d'articles, Murray et al. ont présenté un certain nombre d'idées importantes concernant la mesure des inégalités de santé. Nous partageons ici leur point de vue sur nombre de questions, mais nous attirons l'attention, en illustrant notre démonstration par un exemple numérique, sur les défauts d'un des aspects majeurs de leur approche – la mesure des inégalités sur la base de données géographiquement limitées. Nous décrivons une autre approche qui fait appel à des données longitudinales et qui permet de conserver, tout en les améliorant, les aspects les plus séduisants de leur proposition, à savoir l'intégration sur la vie entière et la prise en compte aussi bien des issues non fatales que des données usuelles sur les profils de mortalité.

Resumen

Sobre la medición de las desigualdades en salud

En una serie de artículos recientes, Murray et al. exponen diversas ideas importantes respecto a la medición de las desigualdades en salud. En este artículo se coincide con muchos de sus razonamientos, pero se señala a la atención un aspecto fundamental de su método – la medición de las desigualdades a partir de datos de áreas pequeñas – que presenta fallos. Se da un ejemplo numérico para ilustrar el problema. También se esboza un método alternativo basado en datos longitudinales, con el que se mantienen y potencian los aspectos más convenientes de su propuesta. Entre ellos figuran la perspectiva del ciclo vital y la consideración de los problemas de salud no mortales, así como de la información habitual sobre las pautas de mortalidad.

References

1. **Murray CJL**. *US county patterns of mortality by race: 1965–1994.* See: www.hsph.harvard.edu/organizations/bdu/papers/usbodi/index.html, 1998.

2. **Murray CJL, Gakidou EE, Frenk J**. Health inequalities and social group differences: what should we measure? *Bulletin of the World Health Organization*, 1999, **77**: 537–543.

3. **Gakidou EE, Murray CJL, Frenk J**. Defining and measuring health inequality: an approach based on the distribution of health expectancy. *Bulletin of the World Health Organization*, 2000, **78**: 42–54.

4. **Braveman P, Krieger N, Lynch J**. Health inequalities and social inequalities in health. *Bulletin of the World Health Organization*, 2000, **78**: 232–235.

5. *The world health report 2000 – health systems: improving performance.* Geneva, World Health Organization, 2000.

6. **Cowell FA.** *Measuring inequality.* London, Philip Allen, 1977.

7. **Wolfson MC.** Divergent inequalities — theory and empirical results. *Review of Income and Wealth*, December 1997.

8. **Gakidou E, King G**. *An individual-level approach to health inequality: child survival in 50 countries. GPE discussion paper series, No. 18.* Geneva, World Health Organization, 2000 (unpublished document).

9. **Torrance GW.** Utility approach to measuring health-related quality of life. *Journal of Chronic Diseases*, 1987, **40**: 593–600.

10. **Vaupel JW, Manton KG, Stallard E**. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 1979, **16**: 439–454.

11. **Ross N et al**. Income inequality and mortality in Canada and the United States. *British Medical Journal*, 2000, **320**: 898–902.

12. **Le Grand J**. Inequalities in health, some international comparisons. *European Economic Review*, 1987, **31**: 182–191.

13. **Wolfson M**. POHEM – a framework for understanding and modelling the health of human populations. *World Health Statistical Quarterly*, 1994, **47**: 157–176

14. **Statistics Canada**. Life expectancy. In: *Health reports: how healthy are Canadians? A special issue*, 2000, Vol. 11, No. 3, Catalogue No. 82-003-XPB.