

A NOTE ON ESTIMATING THE CAUSE OF DEATH STRUCTURE IN HIGH MORTALITY POPULATIONS

A. D. Lopez and T. H. Hull***

SUMMARY

Strategies for reducing mortality should be based on an informed assessment of the underlying structure of causes of death. In the absence of an adequate vital registration scheme, cause of death information is generally derived from fragmentary data which are often limited to specific sectors of the population. Using the previous mortality experience of industrialized countries at higher levels of mortality and the cause of death returns of contemporary developing societies with reasonably accurate and complete mortality data, a linear model has been developed to estimate the cause-structure of mortality in a population from knowledge of the crude death rate alone. Through an appropriate weighting of age-cause-specific mortality rates, emphasis is placed on those causes of death likely to be of greatest concern in high mortality populations.

Using the model parameters derived in this fashion, the technique is applied to estimate the cause-structure of mortality in Java. The resulting distribution of deaths differs markedly from that suggested by an older age-standard and is much more in accord with epidemiological opinion about likely causes of death based on rural health surveys, hospital records and other health data. Regression parameters for estimating the cause of death structure during infancy and early childhood are also shown and may be used independently of the all-age regression estimates.

With the widespread adoption of new indirect methods to estimate demographic parameters from limited data, much more has been learned about the level and trend of mortality in a number of developing countries. At the same time, systems for the collection of cause of death data, such as lay reporting schemes, have generally failed to provide a more reliable account of the underlying cause of death structure in these populations. Consequently, health policies designed to reduce overall levels of mortality continue to be based on fragmentary and often highly questionable evidence about the relative impact of certain causes of death.

The possibility exists, however, that the cause structure of mortality can be estimated with some confidence through an appropriate model relationship depending on a single parameter such as the general level of mortality prevailing in the population. Hull and Rohde (1978), for example, in their study of mortality in Java in the early 1970s, employed a series of linear relationships of the form

$$M_i = a_i + b_i M$$

to estimate the cause-specific death rate (M_i) from the overall crude death rate (M). The set of coefficients (a_i , b_i) used to define the model were originally derived by Preston (1976, p. 18) from the recorded mortality experience of 165 populations at various stages of development

using the age-standardized death rate as a summary index of the mortality schedule.¹

An important consideration in the formulation of the model is the choice of weights to be applied to the respective age and age-cause-specific death rates. In deriving his set of functional relationships, Preston chose as a standard the West female stable population age structure defined by an expectation of life of 65 years and an annual growth rate of 1 per cent. As a result, the standardized mortality rates, and hence the coefficients (a_i , b_i), are more representative of mortality patterns at the older ages than of the cause of death structure in a typical developing country. Clearly, if inferences are to be drawn about the relationship between M_i and M in a population with a comparatively high growth rate, they

TABLE 1. PERCENTAGE DISTRIBUTION OF STANDARD POPULATIONS
BY BROAD AGE GROUPS

Age last birthday	West female stable population with:	
	$e_0^0 = 45.0, r = 0.02$ (younger standard)	$e_0^0 = 65.0, r = 0.01$ (older standard)
0-4	15.4	10.0
5-14	24.3	18.2
15-44	43.8	43.1
45-64	12.8	19.4
65+	3.7	9.3

Source: Coale and Demeny (1966).

¹For a description of the populations comprising the basic data set, see Preston (1976), pp. 3 and 4.

*Division of Health Statistics, World Health Organization, Geneva, Switzerland.

**Department of Demography, Australian National University, Canberra, Australia.

TABLE 2. COEFFICIENTS FROM THE REGRESSION OF AGE-STANDARDIZED CAUSE-SPECIFIC MORTALITY RATES (M_i) ON THE AGE-STANDARDIZED DEATH RATE FOR ALL CAUSES COMBINED (M)

Cause of death category (i)	$c_0^0 = 45.0, r = 0.02$						$c_0^0 = 65.0, r = 0.01$					
	Males			Females			Both sexes			Males		
	a_i	b_i	a_i	b_i	a_i	b_i	a_i	b_i	a_i	b_i	a_i	b_i
Respiratory tuberculosis.....	-0.00040	0.09350 (0.00426)	-0.00029	0.09264 (0.00452)	-0.00034	0.09254 (0.00307)	-0.0011	0.1188 (0.0054)	-0.0007	0.1188 (0.0054)	0.1059 (0.0049)	
Other infectious and parasitic diseases	-0.00094	0.16661 (0.00610)	-0.00065	0.16272 (0.00583)	-0.00076	0.16153 (0.00432)	-0.0015	0.1458 (0.0062)	-0.0011	0.1458 (0.0062)	0.1398 (0.0052)	
Neoplasms	0.00104	-0.02633 (0.00230)	0.00079	-0.01039 (0.00167)	0.00091	-0.01770 (0.00147)	0.0024	-0.0569 (0.00390)	0.0016	-0.0569 (0.00390)	-0.0245 (0.00335)	
Cardiovascular diseases	0.00187	0.00622 (0.00682)	0.00126	0.03384 (0.00343)	0.00152	0.02331 (0.00400)	0.0048	-0.0316 (0.0160)	0.0033 (0.0160)	-0.0316 (0.0160)	0.0179 (0.0123)	
Influenza/pneumonia/bronchitis	-0.00126	0.25028 (0.00768)	-0.00080	0.22841 (0.00760)	-0.00100	0.23696 (0.00546)	-0.0026	0.2831 (0.0082)	-0.0016	0.2831 (0.0082)	0.2434 (0.0078)	
Diarrhoeal diseases	-0.00062	0.12598 (0.00723)	-0.00042	0.12538 (0.00707)	-0.00050	0.12361 (0.00507)	-0.0010	0.1030 (0.0066)	-0.0007	0.1030 (0.0066)	0.1041 (0.0060)	
Certain degenerative diseases.....	0.00022	0.01269 (0.00245)	0.00017	0.01164 (0.00224)	0.00019	0.01276 (0.00167)	0.0003	0.0206 (0.0050)	0.0003	0.0206 (0.0050)	0.0165 (0.0043)	
Complications of pregnancy ^b	—	—	0.00004	0.01854 (0.00071)	-0.00001	0.00734 (0.00086)	—	—	-0.0001	—	-0.0001 (0.0197)	
Certain diseases of infancy	0.00022	0.07457 (0.00468)	0.00020	0.06976 (0.00429)	0.00020	0.07323 (0.00317)	-0.0001	0.0447 (0.0033)	0.0000	0.0447 (0.0033)	0.0422 (0.0028)	
Violence ^c	0.00038	0.02370 (0.00453)	0.00021	0.00826 (0.00134)	0.00034	0.02242 (0.00335)	0.0006	0.0232 (0.0041)	0.0003	0.0232 (0.0041)	0.0041 (0.0014)	
All other and unknown causes (residual)	-0.00072	0.27315 (0.00879)	-0.00041	0.25952 (0.00813)	-0.00055	0.26442 (0.00598)	-0.0020	0.3475 (0.0153)	-0.0013	0.3475 (0.0153)	0.3307 (0.0131)	
Total ^d	-0.00001	1.00037	0.00002	1.00032	0.00000	1.00042	0.0002	1.0002	0.0000	1.0002	0.9998	

Source: Computed from data given in Preston, Keyfitz and Schoen (1972). Coefficients shown in right-hand panels are taken from Preston (1976, p. 18).

^aBased on death rates expressed as annual deaths per person.
^bCoefficients are only meaningful for the female population.
^cIncluding motor vehicle accidents.
^dThe coefficients a_i should sum to zero and the b_i should sum to one. Differences are due to rounding.

TABLE 3. ESTIMATED PERCENTAGE DISTRIBUTION OF DEATHS BY CAUSE FOR TWO STANDARD AGE STRUCTURES, JAVA, 1972

Cause of death category (i)	Percentage of all deaths attributable to ith cause under the West female stable population with:					
	$e_0^o = 45.0, r = 0.02$			$e_0^o = 65.0, r = 0.01$		
	Males	Females	Total	Males	Females	Total
Respiratory tuberculosis	6.7	7.1	6.8	4.7	5.3	5.0
Other infectious and parasitic diseases	10.5	11.3	10.8	4.7	5.7	5.2
Neoplasms	4.2	4.9	4.6	10.3	9.6	10.0
Cardiovascular diseases	13.1	12.8	13.0	28.9	26.7	27.9
Influenza/pneumonia/bronchitis	16.7	16.8	16.7	11.3	12.3	11.7
Diarrhoeal diseases	8.5	9.4	8.9	3.9	5.1	4.5
Certain degenerative diseases	2.7	2.4	2.6	4.1	3.9	4.0
Complications of pregnancy	—	1.6	0.7	—	1.2	0.6
Certain diseases of infancy	8.9	8.5	8.8	3.9	4.2	4.0
Violence	6.2	2.4	4.6	6.4	2.7	4.6
All other and unknown causes (residual)	22.5	22.8	22.5	21.8	23.3	22.5
TOTAL	100.0	100.0	100.0	100.0	100.0	100.0
Crude death rate	15.1	13.3	14.2	15.1	13.3	14.2

*Calculated as a weighted average of proportions for males and females.

should rather be based on a model of mortality which better reflects the impact of a younger age structure.²

In constructing their mortality tables, Preston, Keyfitz and Schoen (1972) also computed age-standardized death rates based on a younger age-standard, the West female stable population with $e_0^o = 45.0$ and $r = 0.02$. A summary of the two age structures is given in table 1. Of particular interest is the difference in the proportion of children aged less than 5 years, a reflection of the differential growth rate of the two populations. Since approximately one half of all deaths occurring annually in populations with a life expectancy of less than 50 years or so can be expected to occur at these ages, the greater weighting assigned to them in the compilation of the mortality index by the younger age standard should lead to a more representative description of the overall pattern of causes of death.³

Table 2 presents the schedule of coefficients (a_i, b_i) which define the set of linear relationships between cause-specific age-standardized death rates (M_i) and the overall level of mortality using the younger age structure as the standard. The model relationships have been estimated for males and females separately, as well as for both sexes combined. For comparative purposes, the coefficients determined by Preston on the basis of the older age standard are shown in the right-hand panels of the table.

The new set of b_i 's, which measure the proportionate contribution, on average, of each cause of death category to total mortality decline, serve to reinforce Preston's

(1976, pp. 19 and 20) conclusions about the relative importance of the various communicable diseases in reducing overall death rates. Only for the "residual" category of diseases do the b_i 's differ by more than 0.05 (i.e., a 5 per cent difference in the estimated average contribution of the residual group). The lower proportionate impact (25 per cent) of these conditions under the younger age standard is to be expected, however, since the majority of deaths in this category occurred at the older ages where a precise diagnosis would often have been extremely difficult given the limited knowledge about the pathology of disease and without modern diagnostic aids. Irrespective of which weighting system is used, sex differences in the b_i 's are relatively small, reflecting the roughly equivalent proportionate role of the various causes of death in reducing overall mortality levels for both males and females.

On the other hand, when the two sets of coefficients are used to estimate the cause of death structure at relatively high levels of mortality, markedly different patterns emerge. This is illustrated in table 3, which shows the estimated proportion of all deaths (expressed as a percentage) assigned to each cause category based on crude death rates of 15.1, 13.3 and 14.2 per thousand males, females and total population, respectively, as estimated for Java in 1972.⁴ It is immediately obvious that the distribution derived from mortality rates based on the younger age standard is much more consistent with the expected cause structure in a youthful, high mortality population. Specifically, the relative impact of diseases of infancy, the diarrhoeal diseases and other infectious and parasitic conditions is approximately twice as great under the new standard, each now claiming an estimated 10 per cent or so of all deaths. The class of specific respiratory ailments also accounts for a higher proportion of deaths, claiming one in six victims, compared to an estimated one

²This is clear from the composition of the total mortality rate. Thus the age-standardized mortality rate for a given population can be $M' = \sum m_x \cdot p'_x$ where $\{m_x\}$ is the schedule of age-specific death rates and $\{p'_x\}$ the proportion of the standard population aged x years. The crude death rate can similarly be viewed as a weighted average of age-specific death rates ($M = \sum m_x \cdot p_x$), the weights $\{p_x\}$ in this case being the actual population composition. Obviously, the closer the agreement between the two sets of weights $\{p_x\}$ and $\{p'_x\}$, the more appropriate will be the regression coefficients (a_i, b_i) for estimating the cause of death structure.

³See, for example, the appropriate national tables in Preston, Keyfitz and Schoen (1972) or the model schedules as estimated by Coale and Demeny (1966).

⁴The cause of death distributions shown in the table were derived from the crude death rate according to the linear model described earlier. The coefficients (a_i, b_i) in the equation are taken from table 3. In this way one can estimate the crude cause-specific death rate (M_i) for each cause of death from knowledge of the crude death rate (M) alone. Since the cause-specific death rates must sum to the overall crude death rate, the proportionate distribution of cause-specific mortality follows.

TABLE 4. REGRESSION PARAMETERS AND IMPLIED CAUSE OF DEATH DISTRIBUTIONS DURING INFANCY AND EARLY CHILDHOOD

Cause of death category (i)	Age 0		Percentage of deaths attributable to ith cause ^a	Age 1-4		Percentage of deaths attributable to ith cause ^a
	a _i	b _i		a _i	b _i	
Respiratory tuberculosis.....	-0.00016	0.00572	0.4	-0.00002	0.02113	2.0
Other infectious and parasitic diseases	-0.00274	0.11703	9.6	-0.00022	0.27130	26.1
Neoplasms	0.00010	-0.00010	0.1	0.00010	-0.00181	0.3
Cardiovascular diseases.....	-0.00084	0.01948	1.3	0.00002	0.01096	1.2
Influenza/pneumonia/bronchitis	-0.00217	0.20369	18.7	-0.00007	0.24409	24.1
Diarrhoeal diseases	-0.00255	0.18998	17.0	-0.00017	0.19509	18.8
Certain degenerative diseases.....	0.00002	0.00151	0.2	0.00003	0.00695	0.8
Certain diseases of infancy	0.01029	0.20779	28.8	—	—	—
Violence	0.00032	0.00800	1.0	0.00039	0.01050	2.9
All other and unknown causes (residual)	-0.00226	0.24690	22.9	-0.00006	0.24128	23.8
FOTAL	0.00001	1.00000	100.0	0.00000	0.99949	100.0

^aAs estimated for Java in 1972 assuming an infant mortality rate (both sexes combined) of 128 per thousand live births and a childhood mortality rate of 21.5 per thousand population aged 1-4 years.

in nine under the old standard. Conversely, the collective impact of the major chronic diseases (neoplasms, cardiovascular diseases and certain other degenerative diseases) has been more than halved, accounting for roughly one in five deaths rather than an improbable 42 per cent as suggested by the older standard.

The regression procedure could, of course, be further refined by estimating separate regression equations for broad age groups chosen so as to delineate between stages of life when the composition of the leading causes of death is likely to change. By aggregating the results one could then obtain an estimated cause of death distribution for the total population.⁵ As the basic data set is presented according to conventional age groupings (0-1, 1-5, 5-10, 10-15, etc. . .), specific regression equations can at least be estimated for two ages which at the same time account for a substantial proportion of all deaths—namely, infancy and early childhood (1-4 years).⁶ The two sets of regression coefficients derived from the 165 populations are shown in table 4, along with the estimated cause of death structure which they would imply given the level of mortality thought to prevail in Java in 1972 (this distribution is obtained in an identical manner to that for the total population as shown in table 3).

The estimated cause-specific distributions which result are largely consistent with what one might expect in a high mortality population. During infancy, three cause of death groups predominate, one of which, the class of disorders peculiar to the first four weeks of life, accounts for almost 30 per cent of all infant deaths. This category essentially comprises the various constitutional maladies

which are often present among the newborn together with such conditions as umbilical sepsis and including deaths due to injuries sustained during parturition. Diarrhoeal and respiratory diseases each claim a further 15 to 20 per cent of infant deaths, most of which could be expected to occur during the post-neonatal period when the malnutrition-respiratory-diarrhoeal triad emerges as the leading threat to infant survival. Much the same mechanism continues to operate throughout early childhood, with an additional substantial contribution (around one quarter of all deaths) due to the other infectious and parasitic diseases, such as measles and diphtheria, which are particularly common at these ages. Interestingly, some 10 per cent of infant deaths at this level of mortality are also estimated to have occurred from this class of diseases, a considerable proportion of which may well be due to neonatal tetanus. One may also note from the relative magnitude of the b_i coefficients that these same causes are largely responsible for effecting declines in mortality at these ages.

The fact that the results are roughly in accord with the cause structure of mortality expected to prevail in a population in the early stages of the epidemiological transition does not, of course, detract from the inherent weaknesses of this type of estimation procedure. To begin with, the empirical evidence used to derive the model relationships was largely drawn from the experience of the industrialized countries at various stages of development and what developing countries were included were almost exclusively from Central and Southern America. It is obviously highly contentious to suppose that similar structural relations exist in contemporary developing societies, given their particular climatic and environmental conditions and the diffusion of modern medical technology. Moreover, the cause of death categories employed are undoubtedly too broad to provide an adequate basis for formulating public health policy, a restriction necessarily imposed in order to improve the comparability of the basic data set. None the less, in countries where knowledge about the cause structure of mortality is extremely limited, these model relationships can at least indicate the order of magnitude of common fatal health concerns, which, in conjunction with simple epidemiological observation, should help to avoid the misappropriation of already scarce national health resources.

⁵Wherever possible, the structure estimated by such indirect methods should be reconciled with available epidemiological and clinical evidence of suitable reliability in order to ensure that the final estimates are roughly in accord with known mortality experience. For an example of this approach, see Hull, Lopez and Rohde (1981).

⁶For the remaining ages (5 years and over), age-standardized cause-specific death rates may be computed, for selected age-segments or for the entire age-span, using the data set of Preston, Keyfitz and Schoen. This procedure, in association with the estimates derived for infants and young children in this article, could be expected to lead to a more realistic appraisal of the underlying overall cause of death pattern in the population. Unfortunately, because the Preston-Keyfitz-Schoen data set was not available to the authors in machine-readable form, it has not been possible to carry out a more complete age-disaggregation in this analysis.

REFERENCES

- Coale, A. J. and P. Demeny (1966). *Regional Model Life Tables and Stable Populations*. Princeton, Princeton University Press.
- Hull, T. H. and J. E. Rohde (1978). Prospects for rapid decline of mortality rates in Java; a study of causes of death and the feasibility of policy interventions for mortality control. Working Paper Series No. 16. Yogyakarta, Indonesia, Population Institute, Gadjah Mada University.
- Hull, T. H., A. D. Lopez and Jon E. Rohde (1981). A framework for estimating causes of death in Indonesia. *Majalah Demografi Indonesia*, vol. 8, No. 15, pp. 77-125.
- Preston, Samuel H., Nathan Keyfitz and Robert Schoen (1972). *Causes of Death; Life Tables for National Populations*. Studies in Population. New York, Seminar Press.
- Preston, Samuel (1976). *Mortality Patterns in National Populations; With Special Reference to Recorded Causes of Death*. Studies in Population. New York, Academic Press.

كيفية الحصول على منشورات الأمم المتحدة

يمكن الحصول على منشورات الأمم المتحدة من المكتبات ودور التوزيع في جميع أنحاء العالم . استلم منها من السفارة التي تتعامل معها أو اكتب إلى : الأمم المتحدة ، قسم البيع في نيويورك أو في جنيف .

如何购取联合国出版物

联合国出版物在全世界各地的书店和经营处均有发售。请向书店询问或写信到纽约或日内瓦的联合国销售组。

HOW TO OBTAIN UNITED NATIONS PUBLICATIONS

United Nations publications may be obtained from bookstores and distributors throughout the world. Consult your bookstore or write to: United Nations, Sales Section, New York or Geneva.

COMMENT SE PROCURER LES PUBLICATIONS DES NATIONS UNIES

Les publications des Nations Unies sont en vente dans les librairies et les agences dépositaires du monde entier. Informez-vous auprès de votre librairie ou adressez-vous à : Nations Unies, Section des ventes, New York ou Genève.

КАК ПОЛУЧИТЬ ИЗДАНИЯ ОРГАНИЗАЦИИ ОБЪЕДИНЕННЫХ НАЦИЙ

Издания Организации Объединенных Наций можно купить в книжных магазинах и агентствах во всех районах мира. Наводите справки об изданиях в вашем книжном магазине или пишите по адресу: Организация Объединенных Наций, Секция по продаже изданий, Нью-Йорк или Женева.

COMO CONSEGUIR PUBLICACIONES DE LAS NACIONES UNIDAS

Las publicaciones de las Naciones Unidas están en venta en librerías y casas distribuidoras en todas partes del mundo. Consulte a su librero o diríjase a: Naciones Unidas, Sección de Ventas, Nueva York o Ginebra.

