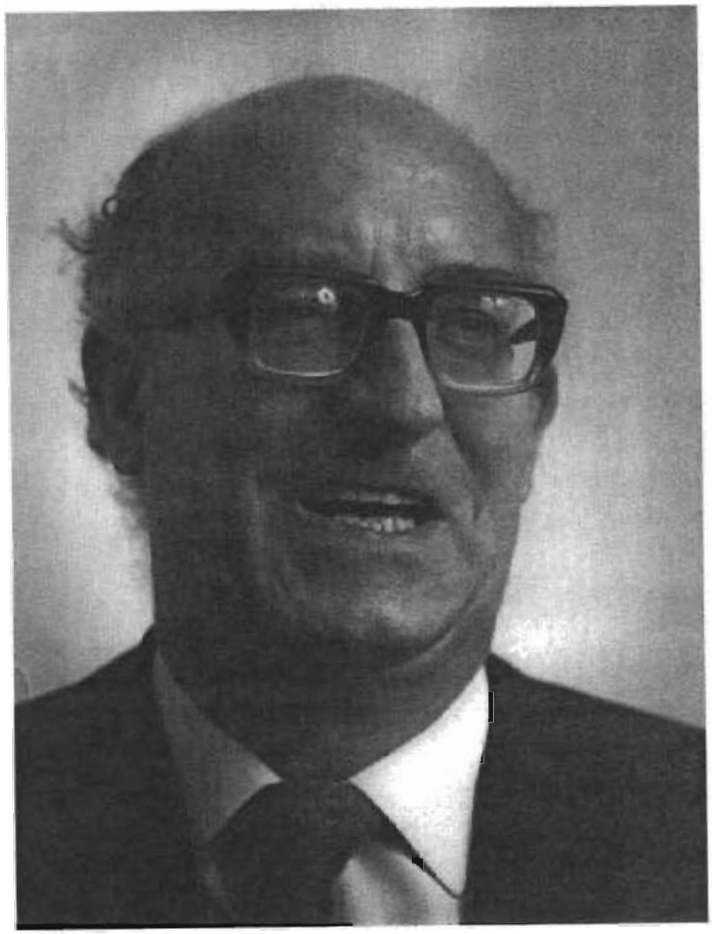

Brass Tacks

Essays in Medical Demography

A TRIBUTE TO THE MEMORY OF
PROFESSOR WILLIAM BRASS

Edited by

BASIA ZABA and JOHN BLACKER

'Birth, and copulation, and death.
That's all the facts when it comes to brass tacks:
Birth, and copulation, and death.'

T.S.Eliot, *Sweeney Agonistes*

THE ATHLONE PRESS
London & New York

CHAPTER THREE

Estimation of adult mortality from data on adult siblings

Ian M. Timæus, Basia Zaba and Mohamed Ali

Centre for Population Studies
London School of Hygiene & Tropical Medicine

Measurement of the mortality of adults in countries that lack effective death registration systems remains a challenge. No fully satisfactory approach exists though demographers have developed a series of ingenious methods that provide in part for this need. William Brass's contribution to this effort has been paramount, both in developing methods to rehabilitate incomplete data on deaths by age (Brass, 1975, 1979a, 1979b) and in pioneering techniques for estimating mortality indirectly from data on the survival of specific categories of relatives (Brass, 1961). Brass and Hill (1973) present the first fully developed method of this type for measuring adult mortality, based on data on the survival of respondents' parents. Brass's associates and others subsequently applied this strategy for the measurement of adult mortality to data on other categories of relatives, including spouses and siblings (e.g. Hill and Trussell, 1977; UN, 1983). The value of the approach was enhanced greatly once Feeney (1980), Brass and Bamgboye (1981) and Brass (1985) had developed methods for estimating the time period to which indirect measures of mortality apply in populations with changing mortality.

The original sibling method for measuring adult mortality has seldom been applied. Trials of it suggested that respondents tend to omit a substantial proportion of siblings from their reports (Blacker and Brass, 1983; Zaba, 1985). Hill and Trussell (1977) point out that in addition to recall errors, respondents might not ever have known about older brothers and sisters who died before the respondent was born or when he or she was very young. Wilson (1985) proposed that one could avoid this problem by asking only about living siblings and those who died in a fixed reference period before the enquiry. This approach has the dual advantages of excluding from the siblings of interest those who are most likely to be omitted from respondents' reports and of generating up-to-date information on mortality. On the other hand, reference period errors may seriously bias mortality estimates made from such data. Moreover, the information collected pertain to both adult siblings and those who are either still children or who died in childhood.

Brass's contribution to the use of reports on siblings for mortality measurement dates from 1989, when he and his collaborators (Graham *et al*, 1989) proposed the sisterhood method for measuring maternal mortality. They suggest that questions should be asked about sisters who survived to age 15 and whether they died subsequently – respondents are then asked whether these adult deaths were associated with pregnancy or childbirth. Restriction of the scope of the questions about survival to adult sisters has two benefits. Firstly, as with Wilson's approach, omission of dead sisters should be far less of a problem than with the original method. Secondly, the resulting data reflect adult mortality alone because those sisters who died as children are excluded from both the numerator and denominator. Because death in the teenage years is uncommon, misclassification of the ages of siblings around age 15 has little impact on the estimated proportion of siblings who died after that age.

Reflecting recent concern about maternal mortality, the sisterhood questions have been asked in numerous surveys (e.g. Graham *et al*, 1989; David *et al*, 1991; Rutenberg and Sullivan, 1991; Hernandez *et al*, 1994; O'Brien *et al*, 1994; Walraven *et al*, 1994; Shahidullah, 1995; Simons *et al*, 1996). By analysing the information on all sisters who have died since age 15, not just the maternal deaths, it should be possible to use these survey data to estimate all-cause adult women's mortality. If data on brothers have been collected, they provide equivalent information on adult men's mortality. Until now, however, no simple method of producing such estimates has existed. This chapter presents such a method.

THEORETICAL BASIS OF THE METHOD

Calculation of the proportion of siblings alive for a given age of respondent requires a model of the distribution of age differences between the respondent and all the other children borne by the respondent's mother. Two approaches have been developed: the first, used by Hill and Trussell (1977), assumes that all mothers experience the age-specific fertility rates of the general population. Thus, a sibling age distribution can be derived as a convolution of the fertility distribution. The second, which is the basis of the Graham *et al* (1989) sisterhood method, assumes that the distribution of age differences of siblings can be represented by a normal distribution. This section discusses the merits of both of these approaches. The term 'fertility distribution' is used here to denote the distribution of age-specific fertility rates in the whole female population, information which is widely available for most populations. The term 'birth distribution' is used to denote the distribution of times since first birth to subsequent births – such distributions are not widely reported in the literature but can be calculated from birth history data.

The first of these two approaches is based on an equation for the proportion of siblings alive by age of respondent in a stable population described by Goodman *et al* (1974). A theoretical attraction of this approach, if the models are developed in the context of stable population structure, is that it yields age distributions of surviving siblings that are consistent with the population age structure when aggregated over all ages of respondent (Zaba, 1987). However conformity at this aggregate level does not ensure that the ages of siblings are realistic for a population in which age patterns of fertility and completed family sizes vary between women, even if the patterns of variation do not change over time.

Keyfitz (1977) has shown that, if family sizes vary, the average size of the family that children are born into, C, must be larger than the average family size borne by women, M, since families of childless women are not represented at all in reports obtained from the children's generation, whilst families of size n are reported n times each. More precisely, $C = M + \sigma^2/M$ where σ^2 is the variance of the distribution of mother's family sizes (Preston, 1976). If the probability of giving birth at any age is independent of whether the woman has given birth at any other age, family sizes will follow a Poisson distribution and the variance of this distribution, σ^2, equals its mean, M. In this special case, a child's average sibship size would be one more than the mothers' average family size. Since larger families would tend to have a wider spread of ages at birth, one might expect that the variance of mothers' birth distributions as reported by the children to be larger than the birth distribution variances measured for all mothers.

The Poisson distribution is not a particularly good approximation for the birth distribution of individual women: the nine-month period of gestation and the period of postpartum amenorrhoea ensure that virtually no children are born in the year before or the year after the birth of a respondent. On the other hand, once a woman enters a stable union and starts bearing children, she is likely to keep up a relatively high tempo of childbearing until she reaches a desired family size or the union breaks up. Unless she enters a new partnership, she then becomes much less likely to bear children. This means, for example, that in a population in which the mean age at first birth is, say, 22 years of age, the few women who bear children at 15 are more likely to also bear children at 18 than other women. For these reasons, Brass has maintained that an adjusted form of the negative binomial distribution is a better representation of birth distributions than the Poisson (Brass, 1958, 1970; Farahani, 1981). If there is a wide scatter of ages at the start of childbearing, but a relatively narrow range of differences between age at first and last birth, as is typical in low fertility populations, we would expect the variance of the birth distribution to be considerably lower than the variance of the fertility schedule. It is the former distribution that determines the sibling age difference distribution.

Graham *et al* (1989) assume that the distribution of sibling age differences can be represented by a normal curve with mean zero and a variance of 80 years-squared. This assumption considerably simplifies the process of estimating the proportion of siblings who remain alive, but is difficult to justify on theoretical grounds. If the distribution of time since first birth to all subsequent births has a variance σ^2, then the age differences between siblings drawn from a cohort of mothers with completed childbearing would have a variance of $2\sigma^2$ and zero mean, but would only be normal if the mother's birth distribution itself was normal. A normal distribution with twice the birth distribution variance constitutes a reasonable approximation for the sibling age difference distribution if the birth distribution is peaked (i.e. $\sigma^2 < 35$), but is less satisfactory for representing sibling age difference distributions in the case of flat birth distributions, such as occur in natural fertility populations. From the discussion of family size, it is clear that the relationship between the variance of the birth distribution and that of the fertility schedule may differ from one population to the next. Therefore, knowledge of the overall shape of the fertility schedule is insufficient to determine the variance of the sibling age difference distribution.

A further problem arises when we consider the distribution of sibling age differences in a growing or shrinking population, rather than a stationary one, which would be equivalent to the cohort considered above. Goldman (1978) proved that, in a growing population, an individual selected at random from those whose mothers have completed childbearing has more younger siblings ever-born than older ones. The opposite is true in a shrinking population. Without repeating her formal mathematical proof here, one can understand this intuitively by considering respondents currently aged 40, all of whose mothers have completed childbearing. In a growing stable population, relatively more of these respondents will have young mothers (say those currently aged under 65 if they have survived) than in a stationary population because, at the time of their birth, there would have been more women aged under 25 than in the corresponding stationary population. But, if the respondents are children of young mothers, they are more likely to have younger than older siblings because their mothers have more childbearing before them than behind them. This means that the distribution of sibling age differences is not symmetrical: its mean lies below zero in a growing population and the opposite is true in a shrinking population. More precisely, if the variance of the underlying birth distribution is σ^2, then the mean of the sibling age distribution lies at approximately $-r\sigma^2$, where r is the population growth rate. Thus, even if all women experience the same age-specific fertility, the variance of the sibling age distribution in a growing population would still be slightly less than twice the variance of the fertility distribution and the distribution would be positively skewed. The opposite features characterize this distribution in shrinking populations.

Our model synthesizes the approaches of Brass and his colleagues (Graham *et al*, 1989) and Hill and Trussell (1977). We follow the former by deriving an expression for the proportion of siblings alive among those who have survived to age 15 in terms of the distribution of differences in ages between the respondent and their siblings. But, like Hill and Trussell, we relate this to a schedule of age-specific fertility, albeit modified to make it resemble a birth distribution more closely. This enables us to allow for the fact that that the age difference distribution is only symmetrical in a stationary population. The next section uses data collected in fertility surveys to examine the relationship between sibling age difference distributions and the age-specific fertility schedules in real populations and to determine empirically the range of variances found in sibling age difference distributions.

Using the probability approach developed by Goodman *et al* (1974), one can show that in a stable population the number of siblings ever-born z years before a respondent currently aged a is given by $\theta(a,z)$:

$$\theta(a,z) = \int_{\alpha}^{\beta} e^{-r(y-a)} f(y) l(y) f(y - z) dy \qquad \text{(for } z \geq 0\text{)} \qquad (1)$$

$$\theta(a,z) = \int_{\alpha}^{\beta} e^{-r(y-a)} f(y) l(y - z) f(y - z) dy \qquad \text{(for } z < 0\text{)} \qquad (2)$$

where equation 1 gives the number of older siblings, equation 2 the number of younger siblings, and:

$l(x)$ = life-table survivorship to age x,
$f(x)$ = the respondents' mothers' probability of giving birth at age x,
r = the growth rate in a stable population,
y = the age of the mother at the birth of the respondent,

and integration is over all ages at childbearing α to β.

The proportion of siblings still alive among those who lived to age 15 for respondents in a five-year age group, x to $x+5$ is given by:

$${}_5S^{15-}{}_x = \frac{\int_x^{x+5} l(a) \int_{15-a}^{\beta-a} \theta(a,z) l(a+z) dz da}{l(15) \int_x^{x+5} l(a) \int_{15-a}^{\beta-a} \theta(a,z) dz da} \quad \text{for } x \geq 15 \qquad (3)$$

EMPIRICAL EVIDENCE ON AGE DIFFERENCES BETWEEN SIBLINGS

Ideally, we would like to investigate the relationship between fertility distributions and birth distributions, and then between birth distributions and sibling age difference distributions, $\theta(a,z)$. This enables us to assess whether the latter relationship conforms to our theoretical predictions and allows us to discover empirical relationships between the variances of the fertility and birth distributions, and between indicators of the timing of fertility and the variance in completed family size. Unfortunately, birth history data collected in fertility surveys do not provide enough information to investigate all these relationships. Such data only furnish a complete picture of cross-sectional age-specific fertility for the year of the survey: earlier years are affected progressively by truncation of the fertility data at older ages. On the other hand, complete birth distributions by time since first birth are only available for the oldest women, with complete fertility. These women are not representative of the whole cohort as the experience of dead mothers is omitted. One can construct a complete distribution of the older siblings ever-born of children born in the survey year. Information on the distribution of younger siblings, however, can only be obtained for children born in the years preceding the survey. As one moves backward in time, these children are progressively less representative of all children born in earlier years, because of truncation of the data on older women and mortality selection. If the population under study was experiencing unchanging age-specific fertility, various extrapolation procedures could be used to fill in the missing information. But, in the face of evidence of changing fertility and contamination of retrospective data by dating errors, we do not consider this a useful approach. Instead, we examine the relationship between current age-specific fertility distribution and the distribution of older siblings, $\theta(0,z)$, omitting the intermediate goal of studying the relationship between both of these distributions and mothers' birth distributions. We can also ascertain how close the distributions of older siblings are to a truncated normal.

We base this investigation on birth history data from 12 World Fertility Survey (WFS) studies conducted in the 1970s. The surveys were selected to provide information on a range of populations with moderate to high fertility drawn from diverse parts of the developing world. The countries selected are listed in Table 3.1, together with estimates of their total fertility at the time when the WFS surveys were conducted.

Figure 3.1 shows the distribution by year of previous births for mothers who gave birth in the year before the 1975 survey in Thailand, a country that is fairly typical of those studied. A three-year moving average line calculated from the data points is also shown. Superimposed on this graph is part of a normal curve, shown cross-hatched, whose variance ($\sigma^2 = 102$)

Table 3.1 *Fertility and sibling age difference distributions in 12 World Fertility Surveys*

Country	Total fertility	Variance of the fertility distribution	Variance of the best normal fit to the sibling age difference distribution
Trinidad and Tobago	3.30	40.33	50.00
Sri Lanka	3.75	42.33	43.13
Thailand	4.63	51.10	67.16
Indonesia	4.73	51.68	82.76
Egypt	5.26	50.75	74.39
Ecuador	5.32	55.19	77.35
Dominican Republic	5.71	43.81	48.76
Tunisia	5.85	51.88	62.44
Morocco	5.90	60.60	97.79
Sudan (North)	6.03	58.92	85.23
Mexico	6.20	47.05	89.51
Senegal	7.15	56.97	77.76

is twice the variance of the age-specific fertility schedule and that has its mean at $-r\sigma^2/2 = -0.053$ (the theoretical mean point of the sibling age difference distribution in a growing population). The growth rate, $r = 0.103$ per cent, was calculated from the growth in total births during the 15 years before the survey (United Nations, 1989). The height of this curve depends on the level of fertility and was fixed by scaling it to obtain the best possible fit to the observed data in the range 2 to 22 years before the index birth[1] Clearly this curve is a poor fit to the observed data, with a variance that is far too wide. In Thailand, as anticipated, the variance of the sibling age difference distribution is considerably less than twice the variance of the population's fertility distribution. The normal curve shown in bold in the same figure was fitted to the data by a non-linear least-squares minimization procedure that allowed both the height of the curve and its variance to change, subject to the same constraint on the mean, again fitting at ages 2 to 22 years. This best-fitting curve has a variance of 67 years-squared and a mean at -0.034 years. Fitting a distribution to the data on older siblings with a mean that is below zero allows largely for the asymmetry of the sibling age difference distribution. Therefore, we can assume that the variance of this distribution is almost the same as the variance of the complete birth distribution.

The relationship in Thailand between the variance of the best-fitting normal and the variance that would be obtained if the birth distribution was normal with the same variance as age-specific fertility is fairly typical of the twelve populations that we studied. The data are shown in Table 3.1 and the relationship across the twelve populations is illustrated in Figure 3.2. As expected,

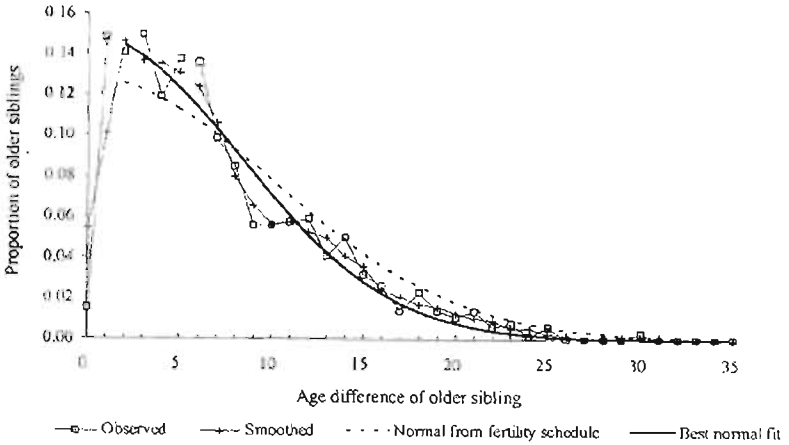

Figure 3.1 Normal fits to the sibling age difference distribution, Thailand

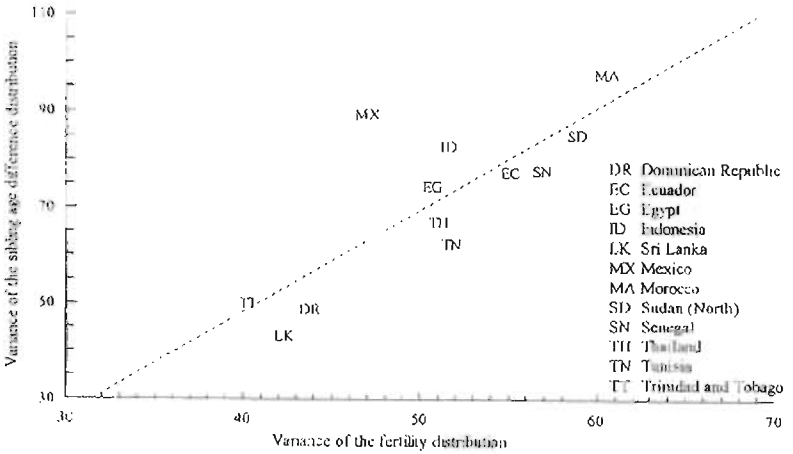

Figure 3.2 Relationship between the variances of the fertility distribution and the sibling age difference distribution in twelve World Fertility Surveys

the variance of the normal fit to the sibling age difference distribution is less than twice that of the fertility distribution in every population. Although the relationship between the variances of these two distributions is a loose one, they are clearly correlated positively. Ignoring the gap in the centre of the distributions, these results suggest that, in populations with moderate to high fertility, sibling age difference distributions have variances that range from about 40 to about 100 years-squared, averaging about 70. Typically,

therefore, the variance of sibling age difference distributions is less than the estimate of 80 adopted by Graham *et al* (1989).

To model full sibling age difference distributions for the purpose of calculating the proportion alive of adult siblings, we adapt an existing fertility model to represent birth distributions. The Relational Gompertz model proposed by Brass (1974, 1981) relates two fertility schedules:

$$-\ln(-\ln(F(x))) = \alpha + \beta.-\ln(-\ln(F_s(x)))$$

where $F(x)$ is the proportion of fertility occurring by age x. The α parameter of this model largely affects the location of the fertility distribution while β largely affects its spread. We use the model in conjunction with the standard fertility distribution, $F_s(x)$, proposed by Booth (1984). To allow for the absence of very short birth intervals in human populations, when generating $\theta(a,z)$ using equations 1 and 2 we set $\theta(a,0) = 0$ and $\theta(a,1)$ and $\theta(a,-1)$ to 40 per cent of the model values. The value of 40 per cent reproduces the average of the ratios $\theta(0,1)/\theta(0,2)$ in the twelve WFS populations. Allowing for the gap in the middle of the sibling age difference distribution in this way, raises the range of variances that should characterize the final age difference distributions from 40 to 100 years-squared to about 45 to 110.

As the theoretical discussion suggests, substitution of Booth's standard fertility distribution into equations 1 and 2 in combination with a range of mortality schedules and growth rates generates distributions of age differences between siblings with larger variances than are typical of the twelve populations on which we have WFS data. We therefore used a set of Gompertz models with a mean value of β of more than one to represent the birth distributions (as β increases the variance of the model distributions decreases). The final set of parameters selected is shown in Table 3.2. As a loose relationship exists between the level of fertility and the variance of the fertility distribution (see Table 3.1), we include no broad sibling age difference distributions ($\sigma^2 > 99$ years^2) for the low growth populations and no narrow ones ($\sigma^2 < 55$ years^2) for the high growth populations. The variances of the sibling age difference distributions that result range from 45.4 to 112.6 years-squared with a mean of 77.8. The widest and narrowest distributions of sibling age differences produced in this way are shown in Figure 3.3. They are close in shape to the twelve observed distributions also shown on Figure 3.3 and, allowing for some sampling and reporting errors in the empirical data, more or less span the range of variation of the latter. Small differences remain. Some of the WFS distributions are very highly peaked at two years (e.g. Dominican Republic, Sri Lanka) while others have a rather flat top (e.g. Morocco, Senegal, Sudan). This reflects variation in the mean length of birth intervals between the populations. Thus, by

Table 3.2 *Models used to simulate the proportion of siblings still alive among those who lived to age 15*

Model	*Parameter*	*Values*
Mortality – relational logit model	α	−1.0
Brass (1971) General Standard		−0.6
(35.3 < e_0 < 74.1, mean 55.3)		−0.2
		0.2
	β	0.7
		1.1
Maternity function – relational Gompertz model	α	−0.5
		−0.2
Booth (1984) standard		0.1
(25.0 < m < 30.8, mean 27.5)		0.4
	β	1.0 (r = 0.03)
		1.15
		1.4
		1.8 (r = 0.01)
Age structure – stable population model	r	0.01
		0.03

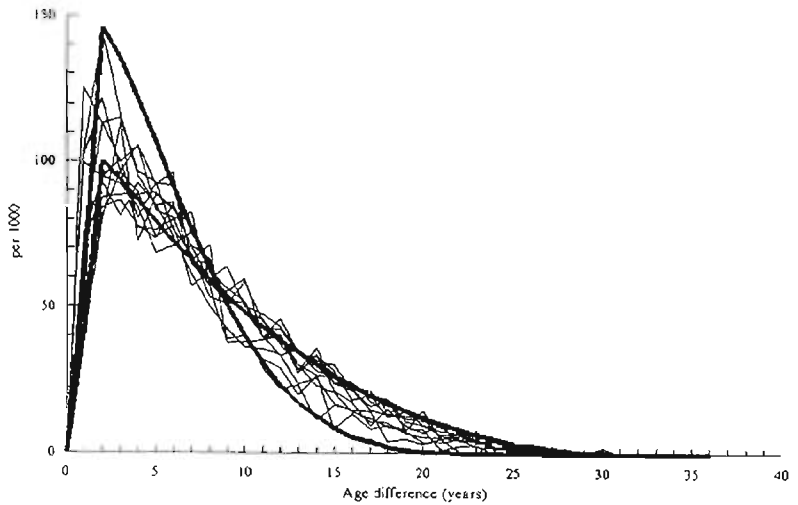

Figure 3.3 Model and observed distributions of the ages of older siblings

combining a range of gaps in the sibling age difference distributions of different width with a slightly different set of Relational Gompertz models, we could probably reproduce the empirical data even more closely. We judge that the gain in the precision of our method for estimating adult mortality would be small and, as students of Brass, restrain ourselves from unwarranted perfectionism.

COEFFICIENTS FOR ESTIMATING LIFE TABLE SURVIVORSHIP

To simplify the estimation of life-table survivorship from data on the survival of adult siblings, we propose a simple regression model of their relationship. We estimate the regression coefficients from data on 192 simulated populations generated from model mortality and birth distribution schedules and the relationships specified in equations 1 to 3. Mortality is represented by relational model life-tables based on Brass's (1971) General Standard. The birth distributions are represented by the Relational Gompertz models of fertility discussed in the previous section. Variation in age structure is represented by the use of two growth rates. These parameters are listed in Table 3.2. Where appropriate, they are the same as those used by Timæus (1991a,b, 1992) to derive coefficients for the estimation of adult mortality from data on orphanhood. Therefore, the adult sibling method should yield estimates of mortality consistent with those from these variants of the orphanhood method.

The adult siblings of respondents aged less than 20 years are, on average, substantially older than the respondents. The relationship between life-table survivorship and the proportion of adult siblings surviving to young respondents is very sensitive to the distribution of age differences between siblings and respondents. The siblings of respondents aged 20 or more years, however, are only slightly older than the respondents on average and the relationship between the respondents' and their siblings' ages varies less. Thus, a close relationship exists between the proportion of siblings still alive among those who survived to age 15 in an age group $n-5$ to $n({}_5S_{n-5}^{15+})$ and life-table survivorship from age 15 to n (l_n/l_{15}). We model this relationship by the regression equation:

$$\frac{l_n}{l_{15}} = \beta_0(n) + \beta_1(n)\,{}_5S_{n-5}^{15+} + \varepsilon \qquad (4)$$

The coefficients of the fitted models are supplied in Table 3.3. It is worth noting that the intercept terms tend to offset the coefficients for the survival of adult siblings. Thus, the adjustments that need to be made to convert the proportion of adult siblings still alive into life-table survivorship from age 15 are small. Despite the wide range of assumptions made about fertility, mortality and age structure to generate these regression equations, the fitted models have high R^2 values and small coefficients of variation. Thus, the relationship between life-table survivorship and the survival of adult siblings

Table 3.3 *Coefficients for the estimation of adult survivorship, l(n)/l(15), from the survival of adult siblings, ${}_5S_a^{15+}{}_5$*

Age (n)	$\beta_0(n)$	$\beta_1(n)$	R^2	Coefficient of variation[a]
25	−0.0003	1.0011	0.9818	0.0042
30	−0.1546	1..560	0.9950	0.0034
35	−0.1645	1.1660	0.9981	0.0029
40	−0.1388	1.1406	0.9984	0.0035
45	−0.1140	1.1168	0.9985	0.0042
50	−0.1018	1.1066	0.9986	0.0052

[a] Root mean squared error divided by the mean value of l(n)/l(15)

is very robust to variation between populations in their other demographic characteristics. Inspection of the residuals confirms that, apart from moderate heteroscedasticity in older age groups of respondents, the regression models adequately represent the data.

TIME LOCATION OF ADULT SIBLING ESTIMATES

The adult sibling method yields a separate estimate of survivorship from the data on each five-year age group of respondents aged between 20 and 50 years. Each estimate reflects the survivorship of a cohort of siblings from age 15 to just less than the mid-point of the respondents' age group. Estimates based on data supplied by younger respondents equal period mortality at a date relatively close to the time that the data were collected. Estimates based on data supplied by older respondents equal period mortality at a somewhat earlier date.

Brass has demonstrated that the time at which orphanhood- and widowhood-based estimates of adult mortality equal period mortality is unaffected by the rate of change in adult mortality and has developed methods for estimating this time (Brass and Bamgboye 1981: Brass, 1985). The structure of the relationships on which the adult sibling method is based is very similar to that underlying the widowhood method in a population where all men and all women marry at about age 15. Thus, existing time location methods for indirect estimates of adult mortality can be applied to mortality estimates obtained from data on adult siblings.

Brass (1985) shows that the time to which indirect estimates of adult mortality apply is:

$$T = \frac{N}{2}\left(1 - \frac{\ln {}_5S_x}{3} + \frac{\ln\frac{80 - M - N}{80 - M}}{3}\right) \qquad (5)$$

where:

T = the time in years before the date when the data were collected.
N = the mean time for which the respondent's relatives are exposed to the risk of dying,
M = the mean age of the relatives at the onset of exposure to the risk of dying,
${}_5S_x$ = the proportion of the relatives remaining alive in the age group x to $x+5$ years.

This equation states that T is somewhat less than halfway through the mean period of exposure and depends on the level of mortality as indicated by the survival of the relatives compared with mortality in a population in which the survivorship function is linear and everyone dies by age 80.

In the adult sibling method, M, the age at which exposure begins, is exactly 15 years for every sibling. The asymmetry of the sibling age difference distribution means that, in a growing population, they are on average slightly younger than the respondents. This age difference varies between about zero and 1.75 years in those populations in which one is likely to want to apply the method. We suggest using a central value of 0.8 years in all applications. Thus, the duration of exposure, N, becomes $(n-2.5–0.8)-15$, where n is still the upper limit of the age group of respondents. Because M is fixed at 15 years, equation 5 can be simplified for each age group to a linear equation of the form:

$$T = \beta_0(n) - \beta_1(n)\ln {}_5S_{n-5}^{15+}$$

These equations for estimation of the time location of life table indices based on data on adult siblings are presented in Table 3.4.

Table 3.4 *Coefficients for the estimation of the time location (T) of life-table indices based on the survival of adult siblings, ${}_5S_a^{15+}{}_5$*

Age (n)	$\beta_0(n)$	$\beta_1(n)$
25	3.23	1.12
30	5.46	1.95
35	7.52	2.78
40	9.38	3.62
45	11.00	4.45
50	12.32	5.28

BIASES DUE TO AIDS MORTALITY

The HIV/AIDS epidemic is producing a massive increase in adult mortality in large parts of Eastern and Southern Africa and some other parts of the developing world. This has highlighted the importance of monitoring levels and trends in adult mortality. It has also made this task considerably more difficult. The HIV epidemic poses two problems for indirect methods of estimating mortality based on the survival of relatives. Firstly, both the sexual and vertical routes of transmission produce significant selection biases in data collected in surveys on the survival of relatives. Secondly, the incidence of HIV infection is concentrated among young adults. Thus, populations with significant AIDS mortality have very different age patterns of mortality both from other populations and from the model life-tables used to derive coefficients for converting data on survival of relatives into measures of life-table survivorship.

Focusing first on the selection biases, almost all the mothers of infected children are infected themselves. A disproportionate number of the children of infected women are also infected, as are the spouses of infected adults. Thus, more children, parents and partners of infected individuals die than other people of the same age. When questions are asked about the survival of women's children, about orphanhood or widowhood, the HIV-positive are less likely to have surviving relatives to report on them than the rest of the population.

Methods have been developed that allow one to make an approximate adjustment for the selection bias in data on the survival of women's children and on orphanhood in populations subject to mortality from AIDS (Timæus and Nunn, 1997). These methods require the analyst to have some idea of the severity of the HIV epidemic in a population. However, the widowhood method for estimating mortality seems beyond salvage. A major advantage of the adult sibling method, compared with these existing methods of measuring adult mortality, is that it is free of selection biases arising from direct transmission of the virus. Some residual bias will remain. In particular, the risk of infection tends to vary markedly between localities and siblings often live close to each other. The impact of this, however, will be relatively small.

Bias in the regression coefficients used to estimate life table survivorship remains more of a problem. With respect to equation 3, it is the change in the age pattern of mortality experienced by the siblings as a result of AIDS that is of concern, not the impact of the epidemic on the sibling age difference distribution. For one thing, the latter distribution would only begin to change 15 years after AIDS mortality became significant. Secondly, the main factor shaping this distribution is the age pattern of childbearing, rather than mortality or age structure. Finally, the regression coefficients are affected only moderately by the characteristics of this distribution.

No data on the survival of adult siblings exist for populations that are both subject to AIDS mortality and have reliable mortality statistics. Thus, we can only assess the biases in sibling estimates of adult mortality by combining simulated and actual data. We do this by calculating the proportions of adult siblings still alive by evaluating equation 3, combining our model of the sibling age difference distribution, $\theta(a,z)$, with survivorship data from an AIDS-affected population. Having done this, we use the regression coefficients in Table 3.3 (which were derived for an AIDS-free population) to estimate life-table survivorship from the proportions of siblings alive and compare the results with the actual survivorship values used as input.

The mortality data used for this exercise refer to women. They come from the Medical Research Council study in the Masaka district of Uganda. This prospective study has collected high-quality mortality data on a population of about 10,000 that has been under demographic surveillance since the end of 1989 (Nunn *et al*, 1997; Timæus and Nunn, 1997). Nearly 12 per cent of women of childbearing age are infected with the HIV virus. Death rates among infected women are an order of magnitude higher than those among the seronegative population. The mortality data used here are based on the first five years of surveillance. They are combined with a $\theta(a,z)$ distribution generated using the mean values of the model parameters shown in Table 3.2. Thus, the proportions of adult sisters alive are calculated allowing for high mortality due to AIDS among respondents but assuming that their mothers were unaffected by AIDS mortality.

The results of the analysis are shown in Table 3.5. The level of mortality prevailing in the study population during the first half of the 1990s was very high. At this level, half those surviving to age 15 would die before their 50th birthday. Because the sibling age difference distribution used to generate the proportions of sisters alive is an average one, almost all of the differences between estimated survivorship in the second column of the table and actual survivorship, shown in the third column, reflect the unusual age pattern of mortality in this population. Despite this unusual age pattern of mortality, the adult sibling method produces estimates of survivorship that are close to the actual values for Masaka district. The estimates based on data which represent the reports that would be expected from respondents aged 20–24 years and more than 40 years are extremely accurate. However, those based on data for respondents aged 25–39 years overestimate survivorship. This is because the regression coefficients fail to allow for the concentration of AIDS deaths in this age range. Estimates of sibling survivorship across narrower age ranges (15 to 25 years) and wider ones (15 to 45+ years) are more accurate because the proportions still alive reflect mortality in age cohorts that have experienced both peak and lower AIDS mortality.

Table 3.5 *Errors in adult sibling estimates of survivorship in a female population affected by AIDS*

Age (n)	Proportion of adult sisters alive	Estimated survivorship from 15 to n	Actual survivorship from 15 to n	Relative error (%)	Level of mortality (α)	Estimated survivorship from 15 to 50	Relative error (%)
25	0.8521	0.8527	0.8646	−1.4	0.6757	0.5070	2.1
30	0.7986	0.7686	0.7245	6.1	0.7565	0.4909	−1.2
35	0.7306	0.6874	0.6585	4.4	0.8347	0.4767	−4.0
40	0.6592	0.6131	0.5965	2.8	0.8700	0.4708	−5.2
45	0.5946	0.5500	0.5558	−1.0	0.8306	0.4774	−3.9
50	0.5396	0.4953	0.4968	−0.3	0.7335	0.4953	−0.3

To use sibling estimates of adult survivorship to monitor mortality trends, it is necessary to fit a model life-table to the estimates for specific age ranges and use it to extrapolate to an index referring to a common range of ages. The final three columns of Table 3.5 present the results of extrapolating to survivorship from 15 to 50 years, ${}_{35}p_{15}$, in this way. The α values indicate the level of mortality in the 1-parameter family of relational model life-tables based on Brass's (1971) General Standard. The proportions of siblings alive were generated using a single period life-table rather than by simulating an AIDS epidemic of growing severity. Thus, except for random fluctuations in the number of deaths by age in the Masaka district study population, one would expect all the estimates to yield the same values of α and of ${}_{35}p_{15}$. Insofar as each series differs systematically by age, it is because the mortality models used to derive the regression coefficients are inappropriate for populations with significant AIDS mortality.

The results are surprising: the final series of estimates of survivorship from 15 to 50 years, ${}_{35}p_{15}$, remain fairly accurate. Those obtained from respondents aged 25 to 34 are more accurate than the estimates of l_n/l_{15} on which they are based. Errors due to the failure to allow for the impact of AIDS on the mortality schedule in first calculating the coefficients and then extrapolating to a common measure of survivorship largely cancel out. Further modelling suggests that this finding is robust to variation in background mortality and choice of a mortality standard. It would not necessarily hold, however, in populations where the demography of the AIDS epidemic, in particular the mean age at death from AIDS, is very different from Masaka. Unfortunately, we know too little about trajectories of mortality change in developing country populations affected by AIDS to investigate fully the sensitivity of the finding to such factors. Nevertheless, estimates of ${}_{35}p_{15}$ obtained from the adult sibling method probably represent relatively robust indices for the monitoring of mortality trends as the AIDS epidemic develops. As with other indirect methods, if successive sets of data are collected for the same population, checks on the consistency of the results for periods when they overlap provide a powerful indication of the robustness of our assumptions.

APPLICATIONS

Figure 3.4 presents estimates of the probability of surviving from age 15 to 50 in three countries that have collected data on the survival of siblings in a recent DHS survey: Peru, Morocco and Zimbabwe. Each plot presents a series of six indirect measures of survivorship for men and women estimated using the adult sibling method. The calculations involved in producing these estimates for Peru are shown as an example in Table 3.6.

The plots also include direct estimates for the periods 0–4 and 5–9 years before the recent DHS surveys calculated using the information on ages and dates of death that these DHS surveys collected in the form of sibling histories (Rutenberg and Sullivan, 1991). The plots for Peru (Timæus, 1995) and Morocco (Timæus, 1991b) also include previously published estimates of survivorship from 15 to 50 for earlier dates. These estimates are based on data on lifetime orphanhood and orphanhood before and since first marriage collected in an earlier DHS survey. In Zimbabwe, the earliest series of estimates are based on the orphanhood data collected in the 1982 Census

Table 3.6 *Calculation of the probability of surviving from age 15 to 50 from data on the survival of adult siblings, Peru Demographic and Health Survey, 1991–2*

Age group of respondents (n−5 to n−1)	Living siblings aged 15+	Dead siblings aged 15+	Proportion of siblings alive (${}_5S_n^{n-5}$)	Survivorship from age 15 to n (${}_{n-15}p_{15}$)	Mortality level (α)	Time location (T)	Survivorship from age 15 to 50 (${}_{35}p_{15}$)
(a) Brothers							
15–19	4539	134	0.9713				
20–24	5713	157	0.9733	0.9740	−0.646	1988.7	0.870
25–29	5642	199	0.9659	0.9620	−0.700	1986.5	0.880
30–34	5074	229	0.9568	0.9511	−0.750	1984.4	0.890
35–39	4481	301	0.9371	0.9300	−0.708	1982.4	0.882
40–44	3472	233	0.9371	0.9326	−0.883	1980.7	0.912
45–49	2533	276	0.9017	0.8961	−0.786	1979.1	0.896
(b) Sisters							
15–19	4543	84	0.9818				
20–24	5488	102	0.9818	0.9825	−0.865	1988.8	0.909
25–29	5430	139	0.9750	0.9725	−0.880	1986.5	0.911
30–34	4809	182	0.9635	0.9590	−0.848	1984.4	0.906
35–39	4290	206	0.9542	0.9495	−0.894	1982.5	0.914
40–44	3533	218	0.9419	0.9379	−0.929	1980.7	0.919
45–49	2339	260	0.9000	0.8941	−0.775	1979.1	0.894

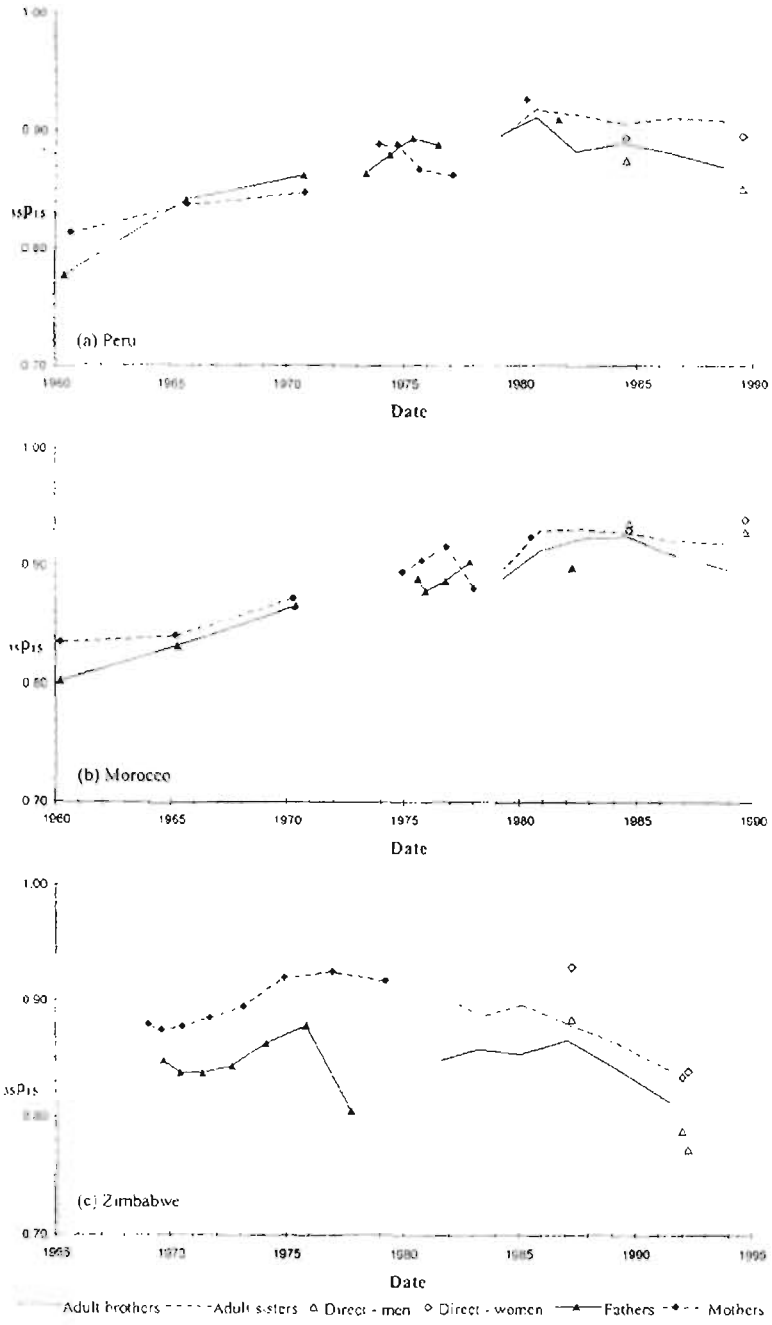

Figure 3.4 Estimates of survivorship from 15 to 50 years, Peru, Morocco and Zimbabwe

(Zimbabwe, 1985). Moreover, direct estimates are also presented for Zimbabwe based on the question about deaths in the last year asked in the 1992 Census (unpublished tables). The latter data have not been adjusted as evaluation using the growth balance and Preston–Coale methods (United Nations, 1983) yielded no clear evidence of underreporting of deaths. Figure 3.4c shows that the Census-based and most recent direct sibling estimates agree very closely.

In Peru, the indirect estimates based on data on adult siblings suggest higher adult survivorship in the late 1980s than the direct estimates based on the same data. The indirect measures may be biased slightly by failure of respondents to report some siblings who died before the reference period for which the direct measures are calculated or by misdating of siblings' deaths. Nevertheless, the two sets of estimates are fairly close. Both suggest that adult women's chances of survival stagnated and adult men's survivorship fell during the late 1980s. Furthermore, the earlier of the indirect estimates of adult mortality obtained from sibling data are very close to the most recent orphanhood-based estimates, which were derived from data on orphanhood since first marriage.

The indirect and direct estimates calculated from data on siblings in Morocco are also in fairly close agreement. Once again, both sets of estimates suggest that the rise in adult survivorship tapered off in the 1980s. Moreover, the earlier estimates made using the adult sibling method are similar to those obtained for the beginning of the 1980s from data on orphanhood since marriage. The indirect estimates for men indicate slightly lower survivorship than the direct estimates. It is possible that it is the latter which are less accurate, perhaps because they have larger sampling errors, as according to them the slight but long-standing excess in adult men's mortality disappeared in the 1980s.

The different series of adult mortality estimates for Zimbabwe also agree quite well. The most recent orphanhood estimates are based on respondents who are still children and may overestimate survivorship somewhat because of the adoption effect. For men, these estimates are erratic. The earlier direct and the indirect adult sibling estimates are consistent with the orphanhood-based estimates. The more recent estimates document a marked decrease in adult survivorship between the first half of the 1980s and the early 1990s. This no doubt reflects the initial rise in AIDS mortality in Zimbabwe. According to the direct estimates, the fall in adult survivorship began later and has been more abrupt than is indicated by the adult sibling method. It was suggested in the last section, that the adult sibling method should perform reasonably well in populations with a high *level* of mortality from HIV/AIDS. Like any indirect method based on the overall proportion of relatives who have died spread out over many years, however, it will tend to smooth out abrupt reversals in the trend in mortality. This seems to have

happened in Zimbabwe. Nevertheless, while the most recent adult sibling estimates overestimate adult survivorship in the early 1990s slightly, the method does successfully reveal that a substantial increase in adult mortality has occurred.

DISCUSSION

In this chapter we have investigated the characteristics of the distribution of the age differences between an index individual and his or her siblings as a basis for the development of a simple regression-based method for estimating adult survivorship from data on the survival of a respondent's adult siblings. The data required to implement this method of estimation are straightforward and can be collected in a single-round household survey. Only two questions are required: 'How many living siblings do you have aged 15 years or more?' and 'How many of your siblings died after surviving to age 15 years or more'. Usually, however, one would ask about brothers and sisters separately to improve the quality of the data and to obtain mortality estimates by sex. In addition, one can ask about the residence of siblings to further improve data quality and obtain information that can be used to measure migration (Zaba, 1985). Such data on the survival of adult siblings already exist for many populations as a by-product of efforts to measure maternal mortality by the sisterhood method (Graham *et al*, 1989). They can now be exploited to measure all-cause mortality.

Our initial appraisal of the performance of the adult sibling method suggests that it has several advantages over existing methods for measuring adult mortality in countries with limited and defective data. Firstly, the relationship between the proportion of adult siblings alive and life-table survivorship is a close one that varies little between populations with different patterns of ages at childbirth and of mortality. If the data are accurate, they should yield rather precise estimates. Secondly, the estimates are efficient in terms of the statistical precision of the proportions because data will typically be collected on considerably more siblings than respondents. Of course, siblings tend to have shared risks of dying. Thus, the effective sample size lies somewhere in between the number of respondents and number of siblings reported on. Nevertheless, there is a real gain in sample size. Thirdly, theoretical considerations, simulated data and the application of the method to data for Zimbabwe, all suggest that the adult sibling method should perform better in populations with significant mortality from HIV/AIDS than existing indirect techniques for the estimation of adult mortality. On the one hand, any selection bias resulting from shared risks of infection should be small, while, on the other, the estimation procedure is relatively unaffected by the unusual age pattern of mortality that develops in populations where HIV infection is prevalent.

The initial applications of the adult sibling method presented here confirm that it is a useful new way of measuring adult mortality. Of course, like any indirect method, the results do not yield the detailed information on time trends and age patterns of adult mortality that are provided by accurate direct data. On the other hand, collecting accurate direct data on adult mortality has proved a difficult challenge (Timæus, 1991c). In particular, collecting sibling histories, as has been done in several Demographic and Health Surveys involves a lengthy and, therefore, expensive series of questions. It is unlikely to be successful in enquiries conducted with less well-trained field staff or less experienced professional staff. The simpler questions required to apply the indirect adult sibling method can be used more widely.

Our investigation of the characteristics of sibling age difference distributions and how they relate to the overall fertility distribution presented here is of potential value in other applications. Most obviously, it provides a means for refining the procedure used to estimate maternal mortality from data on sisters who die while pregnant or during the postpartum period (Graham *et al*, 1989). It also bears on the existing literature on the mathematical demography of kinship.

Preston (1996) has recently suggested that indirect methods of estimating mortality have been:

> one of the great achievements of demography during the post-war period. They raised the status of demographers in the international health arena, because demographers alone appeared to have the tools needed to measure a central component of a population's health.

Both the original vision underlying these methods and much of the detailed research involved in their development originate with William Brass. Preston goes on to argue that 'improved methods for the assessment of the mortality of adults remain an important piece of unfinished business for demography'. In another paper in the same volume, Brass (1996) concludes that 'There appear to be no insuperable problems to prevent the development of improved techniques for indirect estimation', referring in particular to the issues raised by the increase in the number of deaths from AIDS.

This paper makes a contribution to the enterprise referred to by Brass and Preston. Indirect methods of estimating mortality will always provide less specific measures than detailed, regularly collected direct data. Even the few results presented in this chapter, however, undermine the argument that indirect methods have nothing to contribute to the monitoring of levels and trends in adult mortality. So long as funding remains restricted, administrative capacity limited, and a substantial part of the world's population illiterate, indirect techniques will remain one important approach to the measurement of both child and adult mortality.

In developing the adult sibling method, we have tried to follow the principles that have guided Brass's own research. The method is straightforward to apply and is based on simple questions that avoid making onerous demands on the memories of those who respond in demographic enquiries. As with other indirect methods, however, the technique's apparent simplicity is underlain by careful mathematical description and empirical study of the relationships between the indices of interest. In essence, however, at least for those trained by William Brass, the development of a new indirect technique of estimation is 'normal' science (Kuhn, 1962). Any particular originality in the research described here can be attributed to the man who first established that such things are possible.

REFERENCES

Blacker, J. G. C. and Brass, W. (1983) 'Experience of retrospective enquiries to determine vital rates', in L. Moss and H. Goldstein (eds) *The Recall Method in Social Surveys*, London: University of London Institute of Education, pp. 48–61.

Booth, H. (1984) 'Transforming Gompertz's function for fertility analysis: the development of a standard for the relational Gompertz function'. *Population Studies*, 38, pp. 495–506.

Brass, W. (1958) Models of birth distributions in human populations. *ISI Bulletin*, 36, pp. 165–78.

Brass, W. (1961) 'The construction of life tables from child survivorship ratios', in *Proceedings of the International Population Conference 1961*. Liège: International Union for the Scientific Study of Population, pp. 294–301.

Brass, W. (1970) 'Outlines of a simple birth distribution model for the study of systematic and chance components of variation'. Paper presented at the third conference on the mathematics of population. Chicago, July.

Brass, W. (1971) 'On the scale of mortality', in W. Brass (ed.) *Biological Aspects of Demography*. London: Taylor & Francis, pp. 69–110.

Brass, W. (1974) 'Perspectives in population prediction: illustrated by the statistics of England and Wales'. *Journal of the Royal Statistical Society, Series A*, 137, pp. 532–83.

Brass, W. (1975) 'Estimating mortality from deficient registration data', in W. Brass (ed.), *Methods for Estimating Fertility and Mortality from Limited and Defective Data*, Chapel Hill, NC: International Program of Laboratories for Population Statistics, pp. 117–23.

Brass, W. (1979a) 'Evaluation of birth and death registration using age distributions and child survivorship data'. *Asian and Pacific Census Forum*, 5(3), pp. 9–11 & 20.

Brass, W. (1979b) 'A procedure for comparing mortality estimates calculated from intercensal survival with the corresponding estimates from registered deaths'. *Asian and Pacific Census Forum*, 6(2), pp. 5–7.

Brass, W. (1981) 'The use of the Gompertz relational model to estimate fertility'. *International Population Conference, Manila*, vol 3, Liège: International Union for the Scientific Study of Population, pp. 345–61

Brass, W. (1985) 'Further simplification of time location estimates for survivorship of adult relatives reported at a survey', in W. Brass (ed.), *Advances in Methods for Estimating Fertility and Mortality from Limited and Defective Data*, London: Centre for Population Studies Occasional Paper, London School of Hygiene & Tropical Medicine, pp. 11–16.

Brass, W. (1996) 'Demographic data analysis in less developed countries, 1946–1996'. *Population Studies*, 50, pp. 451–67.

Brass, W. and Bamgboye, E. A. (1981) *The Time Location of Reports of Survivorship. Estimates for Maternal and Paternal Widowhood and the Ever-Widowed*, London: Centre for Population Studies Research Papers, 81–1, London School of Hygiene & Tropical Medicine.

Brass, W. and Hill, K. (1973) 'Estimating adult mortality from orphanhood', in *International Population Conference, Liège, 1973*, vol 3, Liège: International Union for the Scientific Study of Population, pp. 111–23.

David, P., Kawar, S. and Graham, W. (1991) 'Estimating maternal mortality in Djibouti: an application of the sisterhood method'. *International Journal of Epidemiology*, 20, pp. 551–57.

Feeney, G. (1980) 'Estimating infant mortality trends from child survivorship data'. *Population Studies*, 34, pp. 109–28.

Farahani, M. (1981) 'A model of fertility by birth order and duration of marriage', unpublished PhD thesis, University of London.

Goldman, N. (1978) 'Estimating the intrinsic rate of increase of a population from the average numbers of younger and older sisters'. *Demography*, 15, pp. 499–521

Goodman, L. A., Keyfitz, N. and Pullum, T. W. (1974) 'Family formation and the frequency of various kinship relationships'. *Theoretical Population Biology*, 5, pp. 1–27.

Graham, W., Brass, W. and Snow, R. (1989) 'Estimating maternal mortality: the sisterhood method'. *Studies in Family Planning*, 20, pp. 125–35.

Hernandez, B., Chirinos, J., Romero, M. and Langer, A. (1994) 'Estimating maternal mortality in rural areas of Mexico: the application of an indirect demographic method'. *International Journal of Obstetrics and Gynecology*, 46, pp. 285–89.

Hill, K. and Trussell, T. J. (1977) 'Further developments in indirect mortality estimation'. *Population Studies*, 31, pp. 313–33.

Keyfitz, N. (1977) *Applied Mathematical Demography*, New York: Wiley.

Kuhn, T. S. (1962) *The Structure of Scientific Revolutions*, Chicago: Chicago University Press.

Nunn A. J., Mulder D. W., Kamali, A. *et al.* (1997) 'Mortality associated with HIV-1 infection over five years in a rural Ugandan population: cohort study' *British Medical Journal*, 315, pp. 767–71.

O'Brien, J., Wierzba, T., Knott, S. and Pikacha, J. (1994) 'Measuring maternal mortality in developing Pacific Island countries: experience with the sisterhood method in the Solomon Islands'. *New Zealand Medical Journal*, 107, pp. 268–9

Preston, S. H. (1976) 'Family sizes of children and family sizes of women' *Demography*, 31, pp. 105–14.

Preston, S. H. (1996) 'Population studies of mortality'. *Population Studies*, 50, pp. 525–36.

Rutenberg, N. and Sullivan, J. (1991) 'Direct and indirect estimates of maternal mortality from the sisterhood method', in *Demographic and Health Surveys World Conference, 5–7 August, 1991, Washington, D.C.*, vol 3. Columbia, MD: IRD/Macro International, pp. 1669–96.

Shahidullah, M. (1995) 'The sisterhood method of estimating maternal mortality: the Matlab experience'. *Studies in Family Planning*, 26, pp. 101–6.

Simons, H., Wong L., Graham, W and Schkolnik, S. (1996) 'Experience with the sisterhood method for estimating maternal mortality', in I. M. Timæus, J. Chackiel and L. Ruzicka (eds), *Adult Mortality in Latin America*, Oxford: Clarendon Press, pp. 108–20.

Timæus, I. M. (1991a) 'Estimation of adult mortality from orphanhood before and since marriage'. *Population Studies*, 45, pp. 455–72.

Timæus, I. M. (1991b) 'Estimation of mortality from orphanhood in adulthood'. *Demography*, 27, pp. 213–27

Timæus, I. M. (1991c) 'Measurement of adult mortality in less developed countries: a comparative review'. *Population Index*, 57, pp. 552–68.

Timæus, I. M. (1992) 'Estimation of adult mortality from paternal orphanhood: a reassessment and a new approach'. *Population Bulletin of the United Nations*, 33, pp. 47–63.

Timæus, I. M. (1995) 'New estimates of the decline in adult mortality since 1950', in I. M. Timæus, J. Chackiel and L. Ruzicka (eds), *Adult Mortality in Latin America*, Oxford: Clarendon Press, pp. 87–107.

Timæus, I. M. and Nunn, A. J. (1997) 'Measurement of adult mortality in populations affected by AIDS: an assessment of the orphanhood method'. *Health Transition Review*, vol 7 (supplement 2), pp. 23–43.

UN (United Nations) (1983) *Indirect Techniques for Demographic Estimation*. ST/ESA/SER.A/81. New York: United Nations.

United Nations (1989) *World Population Prospects, 1988*. ST/ESA/SER.A/106. New York: United Nations.

Walraven, G., Mkanje, R. J., van-Roosmalen, J., *et al.* (1994) 'Assessment of maternal mortality in Tanzania'. *British Journal of Obstetrics and Gynaecology*, 101, pp. 414–17.

Wilson, S. E. (1985) *The Estimation of Recent Levels of Adult Sibling Mortality*, DrPH dissertation, University of North Carolina at Chapel Hill, Ann Arbor: UMI Dissertation Services.

Zaba, B. (1985) *Measurement of Emigration using Indirect Techniques*, Liège: Ordina Editions for IUSSP/CELADE.

Zaba, B. (1987) 'The indirect estimation of migration: a critical review'. *International Migration Review*, 21, pp. 1395–445.

Zimbabwe, Central Statistical Office (1985) *Main Features of the Population of Zimbabwe*. Harare: CSO.