RESEARCH

A new method for estimating recent adult mortality from summary sibling histories

Bruno Masquelier[1*], Ashira Menashe-Oren[1], Georges Reniers[2] and Ian M. Timæus[2,3]

[*]Correspondence: bruno.masquelier@uclouvain.be
[1]Center for Demographic Research, UCLouvain, Louvain-la-Neuve, Belgium
Full list of author information is available at the end of the article

Abstract

Background: In low- and middle-income countries with limited death registration statistics, adult mortality rates are commonly estimated through sibling survival histories (SSH). In full SSH, respondents are asked about either the age, or the age and time of death, of each of their siblings in turn. Full SSH allow direct mortality estimation but can be time-consuming to collect. In this study, we introduce a new indirect estimation method using summary SSH, requiring only a limited set of questions to produce recent mortality estimates.

Methods: We developed a set of 192 microsimulations representing a wide range of fertility and mortality patterns, and reconstructed summary SSH within these simulations as if they had been collected from adults aged 15-49. For each age group of respondents, we calculated coefficients that convert the proportion of adult siblings who died in the previous 5 years into age-specific mortality rates. We then evaluated the performance of this new method with real data, using 154 Demographic and Health Surveys.

Results: The new indirect method provides mortality rates that are consistent with direct estimates from full SSH. Across all DHS, the mean absolute percentage error in the risk of dying in adulthood (ages 15-49) is 6% for both men and women. In all but one survey, 95% confidence intervals around the direct and indirect estimates overlap. As with direct estimates of adult mortality from SSH, the indirect estimates remain, however, lower than those of the Population Division of the United Nations.

Conclusions: Summary questions on sibling survival can be included in censuses and rapid turn-around surveys for the measurement of recent adult mortality.

Keywords: Adult mortality; Indirect estimation; Sibling survival histories; Demographic and Health Surveys

Background

In many low- and middle-income countries, systems of civil registration and vital statistics are underdeveloped, and improvements in the completeness of death registration have been slow [1]. Population-based surveys and censuses thus remain key data sources for estimating mortality, especially among adults [2, 3]. Sibling survival histories (SSH) collected in several survey programs, most notably the Demographic and Health Surveys (DHS), have helped fill important data gaps since the 1990s.

In full SSH, women aged 15-49 are asked about all maternal siblings, and provide details on their gender, and age at the survey if still alive. For the deceased siblings, information is collected on age at death and years since death. All-cause mortality

rates are obtained by dividing the number of deaths in each desired period and age group by the total number of years of exposure. Additional questions enable one to identify cause-specific mortality such as pregnancy-related or violent deaths. Full SSH have been used to monitor mortality increases due to HIV/AIDS [4, 5], to measure the effectiveness of health programmes [6], or, to study excess mortality associated with violence [7–9]. Full SSH are also one of the primary sources of data on adult survival used by the *Global Burden of Disease* (GBD) study [10] and the *World Population Prospects* (WPP) [11].

The collection of full SSH data comes with an important caveat. Full SSH modules can be time-consuming; in Senegal, the median duration of an interview to collect full SSH and some basic socio-demographic information was 30 minutes [12]. Consequently, about half of Demographic and Health Surveys omit the sibling module, and census questionnaires never incorporated full SSH. Moreover, the full SSH module may be less suitable for mobile phone surveys, where brevity is important to limit interruptions due to network problems or respondent fatigue [13]. Summary SSH can address this caveat. Compared to full SSH, collecting summary SSH can reduce interview duration by approximately 10 minutes [14]. The framing of questions varies, but respondents are typically asked about the total number of sisters/brothers ever born, and the number of these sisters/brothers who have died. Mortality rates are indirectly derived from the proportions of surviving siblings, tabulated by the respondent's age. Indirect estimation is possible with data from DHS (using the summary data collected before the full SSH questions) and a selection of Multiple Indicator Cluster Surveys (MICS) and CDC's Reproductive Health Surveys. Some censuses have also collected summary SSH (e.g. Dominican Republic in 2002, Eswatini in 2007, and Lesotho in 2006). When first introduced, summary SSH were only asked of sisters and associated with an additional question regarding the number of sisters who died during pregnancy, childbirth, or within six weeks after the end of pregnancy. The proportions of adult sisters who died from pregnancy-related causes can be converted into pregnancy-related mortality rates using the "sisterhood method" developed by Graham et al. [15]. Timæus et al. in 2001 extended this method to all-cause mortality, including the estimation of men's mortality from data on brothers [16]. To apply their method, respondents should have been asked about the number of their siblings who survived to the age of 15 and how many of them are still alive. No additional information is required regarding the ages or timing of the death of siblings.

The indirect method has three significant limitations, all related to the timing of deaths. First, estimates cover long reference periods, with data from respondents aged 20-24 reflecting mortality experiences three years before data collection, and data from those aged 45-49 yielding estimates for a period up to 15 years prior to the survey. Second, the method assumes regular and unidirectional mortality trends, making it unsuitable for settings disrupted by conflict, disasters, or epidemics that may cause sudden increases in mortality. Third, there is an increased likelihood of omissions in respondents' reports because the calculation includes sibling deaths from the distant past [5].

To address these limitations we introduce a novel approach, focusing exclusively on recent deaths and building on the literature on synthetic cohorts for mortality

estimation [17, 18]. The method requires that the following three questions are asked:

1 How many of your sisters born to the same mother have reached age 15?
2 How many of these adult sisters have died?
3 How many of these adult deaths occurred during the last 5 years?

The same questions can be asked about brothers to estimate male mortality. The first two questions enable one to apply the method developed by Timæus et al. (2001) to reconstruct past trends [16]. The last question makes it possible to obtain the proportions of siblings who have died in the last five years, among all those who have reached adulthood, for each age group of respondents. These proportions can be converted into conditional probabilities of dying referring to the recent past.

In introducing this new method, we use a set of microsimulations to generate coefficients to convert reports on sibling deaths to mortality estimates. We show in detail how this method differs from the direct estimation method, and the original indirect method of Timæus et al. [16]. We evaluate the performance of both indirect approaches, comparing estimates obtained from 154 DHS, and in comparison to the WPP.

Existing indirect methods for summary sibling histories

Using theoretical models of stable populations, the proportion of surviving siblings can be expressed as a function of age patterns of mortality and fertility, and the population growth rate. For example, if we denote the current age of a respondent as a and the age interval between the respondent and their siblings as $x - y$, the number of surviving siblings born before the respondent is obtained as follows [19]:

$$S^{old,surv}(a) = \int_{\alpha}^{\beta} \int_{\alpha}^{x} m(y) e^{-rx} \, {}_{x}p_0 \, m(x) \, {}_{a+x-y}p_0 dx dy \tag{1}$$

where $m(x)$ refers to the fertility schedule, ${}_{x}p_0$ to the life table survivorship to age x and r to the growth rate in a stable population. A similar equation exists for surviving siblings born after the respondent. Hill and Trussell [20] used these expressions to convert the proportions of surviving siblings into life table survivorship, accounting for all deaths, including those that occurred in childhood. Through simulations of various stable populations, they generated 324 sets of proportions of surviving siblings by age group of the respondent (${}_{5}S_n$). These proportions were related to the underlying probabilities of survival using the following equation:

$${}_{n}p_0 = \beta_0(n) + \beta_1(n) {}_{5}S_n \tag{2}$$

Despite its potential, this indirect method has not been widely adopted, because the proportions are prone to errors as the respondent may omit siblings, especially those who died before the respondent reached maturity. A comparison of sibship sizes reported in SSH with sibship sizes expected from fertility trends suggested that around 15% of siblings were unreported in DHS, mostly due to omissions of siblings who died in childhood and/or during the childhood of the respondent [21].

What one assumes about the age differences between siblings is important. Graham et al. [15] assumed that the distribution of age differences between the respondent and her sisters remains constant across different ages of respondents, is normally distributed, and is centred around zero when the sibships are complete, i.e. when the mothers of the respondents have concluded their reproductive period. This assumption amounts to considering that the age of the respondents is our best approximation for the age of their sisters. However, Garenne and Friedberg [22] evaluated the performance of the sisterhood method through simulations and identified several issues: indirect estimates were consistently higher than direct estimates, and had a large margin of error. The pregnancy-related mortality rates inferred from the youngest respondents appeared particularly biased. One key source of this bias can be traced back to the assumption about age differences.

Timæus et al. [16] showed that the average difference between the age of an individual and the age of their siblings is zero only in stationary populations. In populations with a positive growth rate, this difference is negative, as respondents tend to have more younger siblings than older ones. Conversely, in populations with a negative growth rate, the difference is positive. In addition, Timæus et al. [16] highlighted that these distributions have varying standard deviations, whereas Graham et al. assumed that it was fixed [15]. Theoretically, in a cohort of women who have completed childbearing, the variance of the distribution of age differences between siblings should be approximately twice the variance of the distribution of intervals between the first birth and all subsequent births— referred to by Timæus and colleagues as the "birth distribution" [16]. However, this birth distribution is not easily observed, and its variance cannot be inferred from knowledge of the fertility schedule measured at the aggregate level. If there is a lot of variance around ages at first birth, the variance of the birth distribution will be smaller than the variance of the fertility schedule. Timæus et al. [16] used data from 12 World Fertility Surveys and confirmed that the variance of the birth distributions was smaller than twice that of the fertility schedule.

They then developed an indirect method for estimating all-cause mortality from proportions of siblings who survived to the time of data collection, among those who reached their 15th birthday (${}_5S_n^{15+}$). In contrast to the method proposed by Hill and Trussell (1977) [20], they excluded siblings who died in childhood from the numerator and denominator. We will refer to these proportions as "adult lifetime proportions" because the death of adult siblings might have happened at any point in time in the past. Using stable population theory, they expressed ${}_5S_n^{15+}$ as a function of mortality and fertility rates. They generated 192 different stable populations, and related the proportions of surviving siblings to survivorship probabilities with linear regression:

$$ {}_{n-15}p_{15} = \beta_0(n) + \beta_1(n) \times_5 S_{n-5}^{15+} \tag{3} $$

The β_0 and β_1 coefficients are provided in the appendix (Table S1). They tend to cancel each other out, suggesting that the adjustments to convert the proportions ${}_5S_{n-5}^{15+}$ into survival probabilities ${}_{n-15}p_{15}$ are small. The method should therefore

not be very sensitive to variations in fertility and mortality age schedules. Reports provided by those under age 20 are discarded, as siblings are on average older than the respondents aged 15-19, rendering the method quite sensitive to the choices made in modelling the distribution of age differences between a respondent and his or her siblings. Estimates obtained from older respondents (e.g. 45-49) refer to an earlier period than those provided by younger respondents (e.g. 20-24). In order to date the estimates, Timæus et al. (2001) [16] computed coefficients to estimate the time elapsed between the reference period and the survey, assuming a smooth and unidirectional change in mortality.

Data and methods

Sibling histories from Demographic and Health Surveys

We use 154 Demographic and Health Surveys that included a module on sibling mortality (appendix, Table S2). These surveys were conducted between 1992 and 2022 in 54 countries. SSH follow a similar structure to the full birth histories collected to measure fertility and child mortality but focus on the children of the respondent's mother. There are relatively few missing responses on gender, survival status, or age at the time of the survey in the SSH collected in DHS [23]. Respondents seem to find it more challenging to place deaths in time, resulting in a higher proportion of missing responses on the timing of deaths as well as heaping on the time since death (especially 10 years prior to the survey). Further, deaths that occurred more than 8-10 years before the survey are underreported more often than recent deaths [5, 24].

Direct and indirect calculations typically leave out the respondent, as the respondent is by definition a survivor. This does not introduce bias into the estimates, as long as the number of adult siblings alive in recent years is not associated with mortality. This is because the exclusion of the respondent in SSH is offset by the absence of data on sibships without survivors and the greater likelihood that a low-mortality sibship will be found multiple times in the sample [25, 26]. Other approaches to dealing with selection biases have been proposed [27, 28].

Direct estimation based on full sibling histories

Direct mortality estimation from full SSH consists of counting deaths and exposure time for a given reference period, by sex and age group [29]. Discrete-time regression models can be used to model trends or patterns by age [4]. Several R packages are available to facilitate direct estimation (e.g. `demogsurv` [1], `DHS.rates`[2]).

An example of direct calculation is provided in Table 1 for the 2015 DHS conducted in Zimbabwe. Estimates are presented for two reference periods: (1) 0-6 years before data collection, consistent with the published DHS reports, and (2) 0-4 years prior to data collection, which aligns with the duration for computing estimates using the new indirect method described below. The calculation is based on months of birth and death imputed by the DHS program from responses given to questions about age at the time of survey, or age at death and time since the death. The age-specific mortality rates (${}_nm_x$) are converted in risks of dying (${}_nq_x$) and

[1] https://github.com/mrc-ide/demogsurv
[2] https://cran.r-project.org/web/packages/DHS.rates/index.html

chained together to obtain the summary index ${}_{35}q_{15}$. Confidence intervals around the age-specific mortality rates and the probability ${}_{35}q_{15}$ are obtained using a stratified jackknife approach. The female ${}_{35}q_{15}$ probability for 0-6 completed years before the survey is 281.6‰ (95% CI: 254.5–307.7), which is consistent with the published report (282‰) [30]. The probability ${}_{35}q_{15}$ calculated for 0-4 years before the survey is slightly lower, estimated at 245.4‰ (95% CI: 216.8–272.9).

n	Age group	Nb. of deaths	Person-years	ASMR (‰)	${}_5q_n$ (‰)	95% CI (‰)
0-6 years before data collection						
15	15-19	30	15812	1.9	9.4	5.0–13.7
20	20-24	49	19840	2.5	12.3	8.6–16.2
25	25-29	113	22540	5.0	24.8	19.6–30.0
30	30-34	212	19692	10.8	52.4	44.2–60.5
35	35-39	180	13168	13.7	66.1	55.3–76.5
40	40-44	123	8046	15.3	73.6	59.6–86.9
45	45-49	83	4831	17.2	82.4	60.4–103.1
				${}_{35}q_{15}$ = 281.6‰ (95% CI: 254.5–307.7)		
0-4 years before data collection						
15	15-19	23	10590	2.2	10.8	5.1–16.0
20	20-24	32	13645	2.3	11.7	7.2–16.4
25	25-29	72	15908	4.5	22.4	16.1–28.5
30	30-34	141	14942	9.4	46.1	36.9–55.2
35	35-39	108	10054	10.7	52.3	39.9–64.1
40	40-44	85	6270	13.6	65.5	49.9–80.8
45	45-49	50	3699	13.5	65.4	43.8–87.3
				${}_{35}q_{15}$ = 245.4‰ (95% CI: 216.8–272.9)		

Table 1 Direct calculation of age-specific mortality rates from the 2015 Zimbabwe DHS, female mortality, 0-6 and 0-4 years prior to the survey

Indirect estimation based on adult lifetime proportions

The indirect method proposed by Timæus et al. (2001) [16] can be applied to DHS data, using only the information on the mean number of siblings born to the same mother who have reached 15 years of age and the number of these siblings who are still alive at the time of the survey. The calculation is illustrated in Table 2, again based on the 2015 Zimbabwe DHS. We used the West model of Princeton life tables to convert survivorship ratios into the summary index ${}_{35}q_{15}$ [31]. The average of values obtained from 20-29-year-old respondents (269.1 ‰) is only about 4% lower than the direct estimate referring to the 0-6 year period before the survey (281.6‰) and is contained in the corresponding 95% confidence interval (254.5–307.7‰). Indirect estimates suggest, however, that mortality remained fairly stable over time. For Zimbabwe, this is not plausible as mortality declined owing to the roll-out of antiretroviral therapy during this period [32]. This decline is not well reflected here, most likely due to the assumption of linearity of trends required to date the estimates. It is also possible that the conversion of the age-specific estimates of ${}_{n-15}q_{15}$ into the summary index ${}_{35}q_{15}$ using a standard mortality pattern introduced biases since this pattern does not capture excess adult mortality in populations with generalized HIV epidemics.

Indirect estimation based on proportions of adult siblings who have died in the last five years

We now introduce the new indirect approach that involves asking an additional question to identify recent deaths. We denote ${}_5S_n^{15+}(t)$ the proportion of surviving siblings at the time of the survey (t), calculated from reports of respondents aged

n	Age group	Sisters reaching age 15	Sisters surviving	${}_5S^{15+}_{n-5}$	${}_{n-15}p_{15}$	Years since survey	${}_{35}q_{15}$ (‰)	95% CI (‰)
25	20-24	2580	2441	0.946	0.947	3.3	245.0	208.2–281.0
30	25-29	3370	3055	0.906	0.893	5.7	293.3	259.5–326.3
35	30-34	3787	3361	0.887	0.870	7.9	259.1	231.9–285.7
40	35-39	3110	2648	0.851	0.832	10.0	256.0	231.0–280.7
45	40-44	2578	2074	0.805	0.785	12.0	262.0	236.9–287.1
50	45-49	1543	1145	0.742	0.719	13.9	281.1	250.3–312.0

Table 2 Indirect calculation of age-specific mortality rates from the 2015 Zimbabwe DHS, female mortality, 0-14 years prior to the survey

from n to $n+5$ years, only among siblings who have reached the age of 15. The change in this proportion, relative to the equivalent proportion calculated among the same respondents at time $t-5$, when they were aged $n-5$ to n, is indicative of the impact of adult mortality over the five years.

Theoretically, it should be possible to use the coefficients developed by Timæus et al. (2001) [16] in this context, after chaining the cohort changes in the proportions across age groups to refer to a synthetic cohort. This approach was first proposed by Zlotnik et al. (1981) [17] for indirect estimation based on child and parental survival, and could be extended to sibling histories. However, one would need to account for the fact that some siblings reach 15 years during the 5-year window period, and these siblings are less likely to die compared to those who already are 15 at time $t-5$. This complicates the construction of a synthetic cohort.

We, therefore, developed new coefficients that can be applied directly to the ratio $\frac{{}_5S^{15+}_{n}(t)}{{}_5S^{15+}_{n-5}(t-5)}$. Previous indirect sibling methods were developed using mathematical expressions that relate the proportions of surviving siblings to mortality patterns in stable populations. This approach allows some freedom in the generation of stable populations and in the choice of regression model specifications but lacks flexibility when using estimators that are difficult to express mathematically. Here, we still work with stable populations, but rather than exploiting their mathematical expression, we simulate them. Through microsimulations of stable populations, we can directly observe sibships and compute the required estimators [26].

We use SOCSIM, an open-source individual-level simulation program designed to simulate population dynamics, including demographic events such as births, deaths, marriage and divorce [33–36]. The waiting time for each event is stochastically determined using a competitive risk model that considers predefined demographic rates. SOCSIM operates as a closed model, meaning individuals can only enter the simulation through birth and exit through death. This closed model structure facilitates the identification of sibships, ensuring that all individuals born during the simulation have an identified mother. Respondents, selected from adults aged 15-49 at the end of the simulation, have their sibling survival histories reconstructed by identifying all individuals born to the same mother.

We model mortality using relational logit models, specifying four values for the α_m parameter and two values for the β_m parameter, based on Brass's general standard [37] (Figure S1). We model fertility using the Brass relational model, with four values of α_f capturing the age pattern of the fertility schedule, four values of β_f describing the spread of the fertility schedule [38], and the standard developed by Booth (1984) for populations with high fertility [39]. The initial age structure of the

population is calculated based on the survival curve, share of the fertility schedule, and two values of the growth rate (0.01 or 0.03). The initial population size is set to reach approximately 40,000 surviving members after 150 years of simulation, allowing sufficient time for the population to grow and eliminate individuals with no pre-established kinship relationships at the start.

The main parameter values for the simulations are detailed in Table 3. These are identical to those used by Timæus et al. (2001) [16] to estimate the relationship between mortality and sibling survival. The resulting 192 simulations encompass a wide range of demographic profiles, with life expectancies at birth ranging from 35.4 to 74.1 years and the mean age of the fertility schedule ranging from 25.5 to 31.3 years.

In the appendix (A.3), we show that the microsimulation model is well calibrated. Mortality rates recalculated from the simulated populations for the last 5 years are consistent with the life tables introduced as input parameters, and direct estimates of the age-specific mortality rates from the reconstructed SSH provide unbiased estimates of the life table probability ${}_{35}q_{15}$, the risk that a person aged 15 dies before reaching age 50 (Figure S3). We also apply the indirect method proposed by Timæus et al. (2001) [16] to the microsimulations and recover unbiased estimates of the probability ${}_{35}q_{15}$. This is expected since we constructed the simulations from the same stable populations as those used to develop the coefficients, but it also indicates that assumptions made by Timaeus et al. (2001) to approximate age differences between siblings are well captured by the simulations.

Mortality	Brass General Standard
α_m	-1.0, -0.6, -0.2, 0.2
β_m	0.7, 1.1
Fertility	Booth standard
α_f	-0.5, -0.2, 0.1, 0.4
β_f	1.0 ($r = 0.03$), 1.15, 1.4, 1.8 ($r = 0.01$)
Growth rate	
r	0.01, 0.03

Table 3 Parameters used to set up the microsimulations

Using the 192 populations, we used linear regression to predict life table survivorship from the aggregate reported proportions of adult siblings who were still alive at the time of the survey, among those who were alive 5 years before, as follows:

$$ {}_5p_n = \beta_0(n) + \beta_1(n) \times \frac{{}_5S_n^{15+}(t)}{{}_5S_{n-5}^{15+}(t-5)} \quad (4) $$

The β_0 and β_1 coefficients are provided in Table 4. Other specifications were explored, such as regressing on ${}_5p_{n-5}$, but this simple model represents greater variance and has the additional advantage of providing estimates for all age groups. When questions are asked among respondents aged 15-49, the resulting estimates of ${}_5p_n$ can be chained together to obtain the summary probability ${}_{35}q_{15}$ as $1 - \prod_{n=15}^{45} {}_5p_n$.

An example of the calculation is provided for the 2015 Zimbabwe DHS in Table 5. The ${}_{35}q_{15}$ probability obtained with this indirect method is 228.8‰ (95% CI: 205.2 - 252.5), which is only 7% lower than the direct estimate computed from full

n	Age group	β_0	β_1	R^2	CV
15	15-19	0.0535	0.9459	0.9200	0.0039
20	20-24	-0.2408	1.2404	0.9323	0.0049
25	25-29	-0.0752	1.0742	0.9320	0.0050
30	30-34	0.0169	0.9829	0.9116	0.0055
35	35-39	0.0151	0.9843	0.9002	0.0067
40	40-44	-0.0273	1.0268	0.9061	0.0078
45	45-49	-0.0332	1.0326	0.9126	0.0086

Table 4 Coefficients used to convert proportions of adult siblings who are still alive at the time of the survey, among those who were alive 5 years before

SSH (Table 1). These two estimates are not significantly different. There are, however, some notable deviations in the age-specific mortality rates. These are higher than the direct estimates when deduced from summary data provided by younger respondents, and lower when derived from older respondents, but the confidence intervals overlap for five of the seven age groups.

n	Age group	${}_5S_{n-5}^{15+}$ $(t-5)$	${}_5S_n^{15+}$ (t)	${}_5p_n$ (‰)	${}_5q_n$ (‰)	95% CI (‰)
15	15-19	0.982	0.963	981.1	18.9	10.0–27.8
20	20-24	0.968	0.946	971.4	28.6	20.6–36.6
25	25-29	0.936	0.906	965.4	34.6	26.2–43.0
30	30-34	0.914	0.887	970.8	29.2	22.7–35.8
35	35-39	0.883	0.851	964.2	35.8	27.9–43.7
40	40-44	0.843	0.805	953.2	46.8	36.4–57.3
45	45-49	0.787	0.742	939.5	60.5	44.9–76.2
				${}_{35}q_{15}$ (0-4y) = 228.8‰ (95% CI: 205.2–252.5)		

Table 5 Indirect calculation of age-specific mortality rates using proportions of adult siblings who are still alive at the time of the survey, among those who were alive 5 years before, 2015 Zimbabwe DHS, female mortality

Results

Indirect estimates derived from adult lifetime proportions or changes in proportions surviving calculated from a single survey

In this section, we analyze all DHS data with an SSH module available at the time of writing. We assess the performance of the two indirect methods, based either on adult lifetime proportions or on recent changes in the proportions surviving. While direct estimates serve as our benchmark, it is important to acknowledge that these might be affected by omission or misplacement of deaths, as well as inaccuracies in age reporting. A record linkage study conducted in the Niakhar Health and Demographic Surveillance System (in Senegal) revealed that respondents tended to underestimate the ages of living siblings, ages at the time of death, and the time elapsed since the deaths [40]. These reporting inaccuracies introduced downward biases in mortality estimates, although recent estimates (0-6 years before the survey) remained unaffected. Several studies also showed that under-reporting of deaths was more pronounced for deaths occurring further in the past [4, 24, 41]. Consequently, direct estimates cannot be regarded as a gold standard, even though those for the period immediately preceding data collection are more reliable. For these reasons, we also compare sibling estimates with those of the WPP, which partly rely on DHS data but incorporate additional survey and census data [11]. The WPP also factor in the expected relationship between child and adult mortality, although this is largely informed by the historical record in high-income countries.

Figure 1 displays a series of ${}_{35}q_{15}$ probability estimates for Cambodia, Senegal and Zimbabwe. These three countries were chosen as examples as they have several surveys with sibling histories, and represent diverse mortality patterns. In Cambodia, the direct estimates are relatively consistent across surveys and indicate that the ${}_{35}q_{15}$ probability decreased by more than threefold from the 1980s to the end of the 2010s (e.g. from 323‰ in 1996 to 102‰ in 2018 for males). Trends inferred from adult lifetime proportions suggest a more gradual decline in mortality, although the mortality levels for the 1990s appear implausibly high among men. Indirect estimates derived from changes in the proportion surviving are virtually identical to those computed from full SSH for the 5 years preceding each survey. In Senegal, the direct estimates remain relatively consistent across surveys, although they are below the WPP values. Indirect estimates obtained from adult lifetime proportions are quite erratic, with several surveys indicating an increase in mortality. Indirect estimates from recent changes in the proportions surviving align closely with those derived from full SSH. Finally, in Zimbabwe, trends have been severely disrupted by the HIV/AIDS epidemic. Direct estimates suggest a fourfold increase in the ${}_{35}q_{15}$ probability between the early 1980s and the early 2000s. Indirect estimates from adult lifetime proportions appear too low: among men, they peak at 265.7‰, compared to a peak of 492.5‰ in direct estimates (in 2002). To a lesser extent, recent levels obtained indirectly from changes in the proportions surviving are also lower than direct estimates at the height of the epidemic.

Figure 2 compares, for all DHS, the indirect estimates derived from adult lifetime proportions with the direct estimates from full SSH (upper panel) and the World Population Prospects 2022 (lower panel) [11]. We interpolated between direct estimates and between WPP values to obtain a reference for each indirect point estimate. These series are not fully consistent; if direct estimates are taken as a reference, the mean absolute percentage error (MAPE) is 18% for male mortality and 17% for female mortality. The indirect method based on adult lifetime proportions tends to provide higher mortality rates; the median ratio between indirect and direct estimates is 1.04 for males and 1.02 for females (Table 6). Although these ratios are close to one, these differences are significant for both sexes ($p < 0.05$, Wilcoxon signed-rank tests). The median ratios of indirect/direct estimates differ by region, being lowest in Eastern and Southern Africa (0.98 and 0.99), and highest in Latin America and the Caribbean (1.16) and South and Southeast Asia (1.16), two regions where adult mortality is relatively low and sibship sizes are smaller. Deviations are greater when indirect estimates from adult lifetime proportions are compared with mortality rates from the WPP. The MAPE values are 26% for males and 25% for females. The ${}_{35}q_{15}$ probabilities calculated from summary sibling histories tend to be much lower than in the WPP, with median ratios at 0.76 in males and 0.78 in females (Table 6). The median ratios decline with the age of respondents to reach 0.69 when estimates are derived from reports of respondents aged 45-49. By region, median ratios are lowest in West Africa, and in the few surveys conducted with sibling histories in North Africa, West Asia and Oceania. The pattern observed in West Africa is consistent with earlier comparisons made with previous revisions of the WPP [5, 41, 42] and these low ratios could be attributed either to poorer data quality in these regions or to systematic over-estimation of adult mortality in the

WPP.

	(a) Adult lifetime proportions		(b) Change in proportions surviving	
	Median ratio	IQR	Median ratio	IQR
	Compared with direct estimates from SSH			
Sex				
Males	1.04	0.93–1.20	0.97	0.93–1.02
Females	1.02	0.91–1.18	0.98	0.94–1.02
Age group				
15-19	-	-	1.09	0.90–1.42
20-24	1.05	0.91–1.25	1.26	1.06–1.54
25-29	1.04	0.93–1.18	1.08	0.94–1.26
30-34	1.02	0.94–1.17	0.85	0.74–1.01
35-39	1.08	0.94–1.23	0.91	0.79–1.09
40-44	1.03	0.92–1.18	0.90	0.77–1.02
45-49	0.99	0.88–1.13	1.00	0.87–1.18
Region				
South & Southeast Asia	1.16	1.06–1.30	1.04	0.99–1.09
Middle Africa	1.01	0.90–1.13	0.97	0.95–1.00
Eastern Africa	0.98	0.88–1.14	0.94	0.92–0.97
Southern Africa	0.99	0.84–1.19	0.94	0.91–1.00
Western Africa	1.01	0.92–1.12	0.98	0.95–1.02
Latin America & Car.	1.16	1.02–1.32	1.02	0.97–1.07
North Africa/West Asia/Oceania	1.13	0.95–1.28	1.06	1.02–1.09
	Compared with WPP 2022			
Sex				
Males	0.76	0.66–0.91	0.84	0.73–0.95
Females	0.78	0.67–0.93	0.83	0.72–0.97
Age group				
15-19	-	-	0.97	0.77–1.31
20-24	0.90	0.76–1.07	0.97	0.75–1.23
25-29	0.84	0.72–0.98	0.85	0.71–1.04
30-34	0.78	0.68–0.91	0.76	0.62–0.93
35-39	0.76	0.66–0.87	0.77	0.62–0.90
40-44	0.73	0.62–0.83	0.78	0.65–0.92
45-49	0.69	0.59–0.80	0.79	0.65–0.94
Region				
South & Southeast Asia	0.86	0.71–0.99	0.91	0.80–1.02
Middle Africa	0.81	0.70–0.95	0.96	0.82–1.02
Eastern Africa	0.77	0.66–0.90	0.82	0.74–0.96
Southern Africa	0.75	0.65–0.94	0.84	0.76–0.93
Western Africa	0.72	0.61–0.83	0.77	0.69–0.85
Latin America & Car.	0.87	0.74–1.04	0.87	0.78–0.99
North Africa/West Asia/Oceania	0.64	0.52–0.77	0.62	0.53–0.75

Table 6 Median ratios of indirect estimates of adult mortality from (a) adult lifetime proportions or (b) changes in proportions surviving, over direct estimates from SSH (upper panel) or WPP estimates (lower panel)
Note: for indirect estimates computed from adult lifetime proportions, all ratios refer to the probability ${}_{35}q_{15}$, while for estimates derived from changes in proportions surviving, the median ratios computed by age groups refer to age-specific mortality from age 15 to 49.

Figure 3 compares the indirect estimates obtained using the new method and those derived from direct calculations based on sibling histories or data from the WPP. This comparison is also detailed in the bottom panel of Table 6. Each of the 154 surveys yields only one point estimate in this series. We observe a high level of consistency between the indirect and direct estimates, with a mean percentage error of 6% for both men and women. The median ratio for the summary probability $35q15$ is 0.97 (IQR: 0.93-1.02) for brothers and 0.98 (IQR: 0.94-1.02) for sisters, and according to Wilcoxon signed-rank tests, these are not significantly different from 1. In all surveys except one for men (Zimbabwe 2005, shown in Figure 1), the 95% confidence intervals around the ${}_{35}q_{15}$ probability obtained directly overlap

with those calculated around the indirect estimate. However, the consistency of the estimates seems to vary with age, as can be seen from ratios calculated for probabilities of death per five-year interval in Table 6. Median ratios are higher than 1 up to age 30, followed by lower ratios between ages 30 and 45. It is difficult to establish whether these fluctuations reflect errors in age reporting affecting direct estimates, or, biases inherent in the model used for indirect estimation. By contrast, regional disparities remain modest and there is good agreement observed in West Africa. This suggests that age or dating errors in full SSH are not more pronounced in this region compared to others.

Finally, when the new indirect estimates are compared to recent mortality risks estimated in the WPP, consistently lower values are observed again in the DHS. The median ratios of DHS/WPP estimates are slightly higher than with adult lifetime proportions, at 0.84 (IQR: 0.73-0.95) for men and 0.83 (IQR: 0.72-0.97) for women (Table 6). These ratios decrease with age, and are particularly low in West Africa and in surveys conducted in North Africa, Western Asia, and Oceania.

Indirect estimates derived from changes in the proportion of adult siblings surviving between two surveys

Recent changes in the proportions of adult siblings remaining alive can also be constructed for a cohort of respondents in two successive inquiries separated by about five years. This approach circumvents the need for an additional question to identify recent deaths but requires that changes in the proportion of adult siblings surviving between two surveys conducted at an awkward interval are converted into survivorship between conventional five-year age groups.

Except in late old age, adult human mortality rises as an exponential function of age and can be represented by a Gompertz-Makeham model. Moreover, most of the variation in mortality between populations is accounted for by α, the level parameter of the model. The β, or shape, parameter of the model varies little between populations, even when they have very different levels of mortality [43]. This empirical finding provides the basis for a method for converting measures of cohort survivorship in between two surveys of a population conducted at an awkward interval into conventional measures of five-year survivorship. In essence, one can use an estimate of β to interpolate within, or extrapolate from, estimated survival over age intervals equivalent to the inter-survey interval to survival over a five-year age interval (see Appendix A.4).

Figure 4 presents estimates of adult mortality for three countries based on inter-survey changes in the proportions that remain alive of the adult siblings of cohorts of respondents. The intervals between the surveys range from a bit less than 4 years to a bit more than 7 years, with the exception of those for 1993-2005 in Senegal, which should probably be discounted. The inter-survey estimates are compared with results from other sibling-based estimation methods that were presented initially in Figure 1. In Cambodia, the estimates of women's mortality are close to those obtained by asking about deaths of adult siblings during the 5 years before the survey. The estimates of men's mortality, however, follow a sawtooth pattern, with only those for the interval between the 2010 and 2014 coinciding with the results from the other methods. This pattern of results might result from fluctuations in

the completeness of reporting of dead brothers from survey to survey, with the data for 2000 and 2021 being least complete and those for 2005 and 2014 perhaps most complete. In Senegal, the estimates for both men and women are erratic. Those for 2005-11 seem somewhat high and those for 2011-17 seem too low, suggesting that the reporting of dead siblings may have been most complete in 2011. In Zimbabwe, all the estimates are lower than those from those obtained from data on the survival of siblings during the 5 years before the survey except for that based on cohort changes in the proportions of respondents adult sisters that are alive in between the 2011 and 2015 surveys. This pattern of results might result from a steady improvement over time in the quality of reporting on dead siblings.

Discussion

In countries without reliable vital statistics, the inclusion of sibling histories in several survey programs made a crucial contribution to our understanding of adult mortality trends. More than 150 Demographic and Health Surveys have incorporated detailed questions about siblings, prompting respondents about the ages of surviving siblings, or, ages at death and the timing of deaths of those who died. Full SSH are an irreplaceable source of information, allowing comprehensive tests of data quality [44], and the detailed modelling of mortality trends, sex ratios and age patterns of adult mortality [4, 42]. Because of the time required for data collection, however, the full SSH module is not systematically included in all DHS and MICS, and it is not well suited to censuses or rapid turnaround surveys such as those conducted in complex humanitarian emergencies [45].

In this study, we evaluated the performance of a new indirect estimation method requiring only summary SSH. Compared with the series of questions initially proposed by Timæus and colleagues in 2001 [16], only one additional question is needed to identify recent deaths. The method can therefore be combined with the one based on adult lifetime proportions of surviving siblings, in contexts where mortality trends have been regular. Using these data, analysts can obtain both a trend (with the original coefficients) and a recent estimate (with the method proposed here).

When applied to microsimulations, the original indirect method based on adult lifetime proportions provides unbiased estimates of adult mortality (see appendix A.3). However, indirect estimates calculated from adult lifetime proportions in DHS data are systematically higher than direct mortality rates. There are several possible explanations for this. First, it could be related to the conversion of age-specific rates into the summary index ${}_{35}q_{15}$ using model life tables. In countries affected by HIV, using an age pattern affected by AIDS would reduce slightly the estimates of ${}_{35}q_{15}$. Second, direct estimates may be too low due to underestimation of the age of living siblings, ages at death, and the time since deaths [12, 40]. Third, the age patterns of fertility and mortality in survey data could differ from those used to compute the coefficients allowing to convert proportions into survivorship probabilities. In particular, in both the analytical computations and microsimulations, all women are exposed to the same fertility distribution, regardless of their parity and the interval since the last birth (apart from a minimal birth interval). It might be possible to improve the conversion of proportions of surviving siblings to risks of dying

by adding some predictors, such as an indicator of the dispersion of the fertility distribution, and this is an area for future research.

Two important limitations of the original indirect method will inevitably remain. First, estimates derived from adult lifetime proportions refer to a relatively distant past (3-15 years before the survey), and second, trends in mortality are assumed to be regular. To address these limitations, we introduced a new indirect method focusing on recent deaths. When applied to survey data, this method provided estimates that were highly consistent with those derived from the full SSH module. The mean percentage error between direct and indirect estimates was only 6% for both sexes, and the median ratios for the summary probability ${}_{35}q_{15}$ were close to one (0.97 (IQR: 0.93-1.02) for brothers and 0.98 (IQR: 0.94-1.02) for sisters). We detected significant differences between the two sets of ${}_{35}q_{15}$ values in only one survey out of 154 (and for male mortality only). When the length of the interview needs to be reduced, and the focus is on recent mortality, this indirect method provides a good alternative to the full SSH module. In addition to saving time, indirect estimates will be less sensitive to errors in the ages of siblings, their ages at death, and the timing of their deaths than direct estimates. The role of models in the estimation process is reduced because there is no need to extrapolate age-specific probabilities to the ${}_{35}q_{15}$ index using standard age patterns; they can be directly chained together. Finally, the method is applicable to countries affected by conflicts and epidemics, as it does not require the assumption that trends have been regular and unidirectional. The method is, however, reliant on the correct attribution of deaths to the last five years before the interview date.

We also showed that recent changes in the proportions of adult siblings remaining alive can be constructed for a cohort of respondents in two successive inquiries separated by about five years. It is possible to adjust data from surveys conducted at awkward intervals to refer to conventional 5-year age groups. However, the resulting estimates were more erratic than the other estimates made from data on sibings, presumably because of differences in sample characteristics, changes in the size and composition of sibships, selective mortality of respondents and/or different rates of misreporting. Nevertheless, in applications in which no questions on the timing of deaths have been asked, this approach may be the only way to obtain an up-to-date estimate of mortality from the summary data on siblings. In other applications, moreover, it may represent a useful cross-check on the other estimates and provide additional insights into the quality of the data.

Given the importance of sibling histories in the measurement of adult mortality worldwide, further research is needed on data quality, possible biases and estimation methods. Other techniques for quantifying mortality from summary data are also conceivable. For example, age differences between the respondent and her siblings could be imputed from a full sibling history collected in the past, or from regional distributions, as in the methods developed for under-five mortality [46, 47].

These future developments and the possibility of using the method proposed in this study should motivate the inclusion of summary questions on siblings in all DHS surveys and in other programs such as MICS. The global burden of mortality in early adulthood is currently concentrated in countries where the development of civil registration systems remains slow. It is therefore vital to improve the collection

of data on these adults from surveys and censuses in the interim, pending full registration coverage.

Appendix

A.1. Coefficients for the indirect calculation based on the number of adult siblings surviving after age 15

n	β_0	β_1	R^2	CV
25	-0.0003	1.0011	0.9818	0.0042
30	-0.1546	1.1560	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

Table S1 Coefficients used to convert proportions of surviving siblings (among those who survived to 15 years) to survival probabilities - Sc: [16]

A.2. DHS surveys used in this study

Country	Year	Sample size (women aged 15-49)	Region
Afghanistan	2015	29461	South Asia
Angola	2015-16	14379	Middle Africa
Burkina Faso	1998-99	6445	Western Africa
Burkina Faso	2003	12477	Western Africa
Burkina Faso	2010	17087	Western Africa
Burkina Faso	2021	17659	Western Africa
Benin	1996	5491	Western Africa
Benin	2006	17794	Western Africa
Benin	2017-18	15928	Western Africa
Bolivia	1994	8603	South America
Bolivia	2003	17654	South America
Bolivia	2008	16939	South America
Brazil	1996	12612	South America
Burundi	2010	9389	Eastern Africa
Burundi	2016-17	17269	Eastern Africa
Congo Democratic Republic	2007	9995	Middle Africa
Congo Democratic Republic	2013-14	18827	Middle Africa
Central African Republic	1994-95	5884	Middle Africa
Congo	2005	7051	Middle Africa
Congo	2011-12	10819	Middle Africa
Cote d'Ivoire	1994	8099	Western Africa
Cote d'Ivoire	2005	5183	Western Africa
Cote d'Ivoire	2011-12	10060	Western Africa
Cote d'Ivoire	2021	14877	Western Africa
Cameroon	1998	5501	Middle Africa
Cameroon	2004	10656	Middle Africa
Cameroon	2011	15426	Middle Africa
Cameroon	2018	13527	Middle Africa
Colombia	2015	38718	South America
Dominican Republic	2002	23384	Caribbean
Dominican Republic	2007	27195	Caribbean
Ethiopia	2000	15367	Eastern Africa
Ethiopia	2005	14070	Eastern Africa
Ethiopia	2011	16515	Eastern Africa
Ethiopia	2016	15683	Eastern Africa
Gabon	2000	6183	Middle Africa
Gabon	2012	8422	Middle Africa
Gabon	2019-21	11043	Middle Africa
Gambia	2013	10233	Western Africa

Gambia	2019-20	11865	Western Africa
Guinea	1999	6753	Western Africa
Guinea	2005	7954	Western Africa
Guinea	2012	9142	Western Africa
Guatemala	1995	12403	Central America
Guatemala	2014-15	25914	Central America
Haiti	2000	10159	Caribbean
Haiti	2005-06	10757	Caribbean
Haiti	2016-17	14371	Caribbean
Indonesia	1994	28168	Southeast Asia
Indonesia	1997	28810	Southeast Asia
Indonesia	2002-03	29483	Southeast Asia
Indonesia	2007	32895	Southeast Asia
Indonesia	2012	45607	Southeast Asia
Jordan	1997	5548	West Asia
Kenya	1998	7881	Eastern Africa
Kenya	2003	8195	Eastern Africa
Kenya	2008-09	8444	Eastern Africa
Kenya	2014	31079	Eastern Africa
Cambodia	2000	15351	Southeast Asia
Cambodia	2005	16823	Southeast Asia
Cambodia	2010	18754	Southeast Asia
Cambodia	2014	17578	Southeast Asia
Cambodia	2021-22	19496	Southeast Asia
Comoros	2012	5329	Eastern Africa
Liberia	2007	7092	Western Africa
Liberia	2013	9239	Western Africa
Liberia	2019-20	8065	Western Africa
Lesotho	2004	7095	Southern Africa
Lesotho	2009	7624	Southern Africa
Lesotho	2014	6621	Southern Africa
Morocco	1992	9256	North Africa
Morocco	2003-04	16798	North Africa
Madagascar	1997	7060	Eastern Africa
Madagascar	2003-04	7949	Eastern Africa
Madagascar	2008-09	17375	Eastern Africa
Mali	1995-96	9704	Western Africa
Mali	2001	12849	Western Africa
Mali	2006	14583	Western Africa
Mali	2012-13	10424	Western Africa
Mali	2018	10519	Western Africa
Myanmar	2015-16	12885	Southeast Asia
Mauritania	2019-21	15714	Western Africa
Malawi	1992	4849	Eastern Africa
Malawi	2000	13220	Eastern Africa
Malawi	2004	11698	Eastern Africa
Malawi	2010	23020	Eastern Africa
Malawi	2015-16	24562	Eastern Africa
Mozambique	1997	8779	Eastern Africa
Mozambique	2003	12418	Eastern Africa
Mozambique	2011	13745	Eastern Africa
Nigeria	2008	33385	Western Africa
Nigeria	2013	38948	Western Africa
Nigeria	2018	41821	Western Africa
Niger	1992	6503	Western Africa
Niger	2006	9223	Western Africa
Niger	2012	11160	Western Africa
Namibia	1992	5421	Southern Africa
Namibia	2000	6755	Southern Africa

Namibia	2006-07	9804	Southern Africa
Namibia	2013	9176	Southern Africa
Nepal	1996	8429	South Asia
Nepal	2006	10793	South Asia
Nepal	2016	12862	South Asia
Peru	1991-92	15882	South America
Peru	1996	28951	South America
Peru	2000	27843	South America
Peru	2004-06	19090	South America
Peru	2007-08	22558	South America
Papua New Guinea	2016-18	15198	Pacific Islands
Philippines	1993	15029	Southeast Asia
Philippines	1998	13983	Southeast Asia
Rwanda	2000	10421	Eastern Africa
Rwanda	2005	11321	Eastern Africa
Rwanda	2010	13671	Eastern Africa
Rwanda	2014-15	13497	Eastern Africa
Rwanda	2019-20	14634	Eastern Africa
Sierra Leone	2008	7374	Western Africa
Sierra Leone	2013	16658	Western Africa
Sierra Leone	2019	15574	Western Africa
Senegal	1992-93	6310	Western Africa
Senegal	2005	14602	Western Africa
Senegal	2010-11	15688	Western Africa
Senegal	2017	16787	Western Africa
Sao Tome and Principe	2008-09	2615	Middle Africa
Eswatini	2006-07	4987	Southern Africa
Chad	1996-97	7454	Middle Africa
Chad	2004	6085	Middle Africa
Chad	2014-15	17719	Middle Africa
Togo	1998	8569	Western Africa
Togo	2013-14	9480	Western Africa
Timor-Leste	2009-10	13137	Southeast Asia
Timor-Leste	2016	12607	Southeast Asia
Tanzania	1996	8120	Eastern Africa
Tanzania	2004-05	10329	Eastern Africa
Tanzania	2010	10139	Eastern Africa
Tanzania	2015-16	13266	Eastern Africa
Tanzania	2022	15254	Eastern Africa
Uganda	1995	7070	Eastern Africa
Uganda	2000-01	7246	Eastern Africa
Uganda	2006	8531	Eastern Africa
Uganda	2011	8674	Eastern Africa
Uganda	2016	18506	Eastern Africa
South Africa	1998	11735	Southern Africa
South Africa	2016	8514	Southern Africa
Zambia	1996	8021	Eastern Africa
Zambia	2001-02	7658	Eastern Africa
Zambia	2007	7146	Eastern Africa
Zambia	2013-14	16411	Eastern Africa
Zambia	2018	13683	Eastern Africa
Zimbabwe	1994	6128	Eastern Africa
Zimbabwe	1999	5907	Eastern Africa
Zimbabwe	2005-06	8907	Eastern Africa
Zimbabwe	2010-11	9171	Eastern Africa
Zimbabwe	2015	9955	Eastern Africa

Table S2: Surveys included in this study, sample sizes and region

Note: We did not have access to datasets of surveys conducted in Eritrea (1995), Mauritania (2000) and Yemen (1997 and 2013). We excluded the SSH data from Peru (2009, 2010, 2012) because of missing data on imputed dates of birth and dates of death. We also excluded surveys from Bolivia (1989), Dominican Republic (1996), Egypt (1988), Ghana (1993), India (1999, 2006), Pakistan (2006), Sudan (1990) because the SSH data was not in a standard format.

A.3. Calibration of the microsimulation set

We calibrate the microsimulation model by calculating mortality in the simulated populations for the last 5 years and confronting these estimates with the life tables which were introduced as input parameters (Figure S3). As about 290 adult deaths (15-49) were produced over the last 5-year period in each simulation (IQR: 119-439), the recalculated ${}_{35}q_{15}$ probabilities are not identical to those in the input life tables, due to random fluctuations. However, these probabilities are unbiased. The median ratio between these probabilities and those in the input life tables is 1.00. The mean percentage deviation in ${}_{35}q_{15}$ is 5.4%. The coefficient of variation around ${}_{35}q_{15}$ is 6.8% on average. In 180 simulations (94%), the ${}_{35}q_{15}$ values from the input life tables are contained in the 95% confidence intervals.

By selecting only women aged 15-49 years at the end of the simulations and reconstructing their mother's birth history, mortality can be estimated directly from the full SSH, again for the last 5 years. Sibling-based estimates are included in Figure S3. About 466 recent adult deaths are mentioned in each set of SSH, which is higher than the actual number of deaths because most sisters are referenced multiple times. Despite this repetition of some sibships, the direct method based on full SSH provides unbiased estimates; the median ratio between the sibling-based and population-based probabilities ${}_{35}q_{15}$ is 1.00. The mean absolute percentage deviation of the sibling-based estimates is 5.8% when taking the population-based rates as reference. The coefficient of variation around the probability ${}_{35}q_{15}$ is 5.9% on average when computed from SSH. In 173 simulations (90%), the "true" probabilities calculated on the whole population are contained in the 95% confidence intervals around the sibling-based estimates.

We also applied the indirect method developed by Timæus et al (2001) [16] to the microsimulated SSH. Figure S4 displays the probabilities ${}_{35}q_{15}$ obtained indirectly from reports of respondents aged 45-49, and the corresponding probabilities in the life table used to generate the simulations. The coefficients provide unbiased estimates of the risks of dying underlying these simulations. The median ratios of indirect to direct estimates of ${}_{n-15}q_{15}$ are all close to 1, and decline only slightly with age (1.02 in respondents aged 20-24 to 0.98 in respondents aged 45-49). Across all age groups, the mean absolute deviation between indirect estimates of ${}_{n-15}q_{15}$ and the corresponding input life table values is only 8.5%.

A.4. Estimating survivorship between conventional five-year age groups from changes in the proportion of adult siblings surviving between two surveys conducted at an awkward interval

If two surveys have been conducted at an awkward interval, one can estimate the survivorship of the siblings of cohorts of respondents over that interval by tabulating the data from one of the surveys for unconventional age groups. This note

suggests a method for working back from such cohort data to conventional measures of mortality. It requires that the results of the second survey are tabulated for unconventional age groups that correspond to conventional five-year cohorts in the first survey. For example, in the case of a 5.75 year interval between the inquiries, one can produce a series of estimates of ${}_5p_n$ from the following cohort measures of sibling survival:

$$\frac{{}_5S^{15+}_{n+0.75}(t)}{{}_5S^{15+}_{n-5}(t-5.75)} \tag{5}$$

The approach that we adopt is to adjust this ratio up or down so that it approximates to the proportion of the adult siblings of respondents aged $n-5$ to n that survive for exactly 5 years. Once this has been done, one can estimate mortality using the synthetic cohort methods proposed in this paper.

Define the gompit of life table survivorship by duration x from any base age n as:

$${}_xz_n = \ln(-\ln({}_xp_n)) \tag{6}$$

Gompertz survivorship is equally log mortality since $-\ln({}_xp_n) = x.{}_xm_n$.

The Gompertz-Makeham model describes mortality as the combination of an age-independent extrinsic component, λ, and an intrinsic component that is an exponential function of age:

$$\mu_z = \lambda + \alpha e^{\beta z} \tag{7}$$

Integrating this expression for μ_z between any two adult ages n and $n+x$:

$$-\ln({}_xp_n) = \lambda x + \frac{\alpha}{\beta}(e^{\beta x} - 1) \tag{8}$$

so that:

$${}_xz_n = \ln\left(\lambda x + \frac{\alpha}{\beta}(e^{\beta x} - 1)\right) \tag{9}$$

Systematic variation across populations in ${}_xz_n$ is captured largely by variations in extrinsic mortality, λ, and in the level or shape parameter of the Gompertz model, α, which has a multiplicative effect on the mortality rates. The scale parameter β, which is equivalent to the slope of the log mortality schedule by age, varies far less between either populations or, Vaupel has suggested, individuals [43]. Instead, Vaupel hypothesizes, differences in population-level and individual frailty established prior to adulthood, that are captured by α, largely account for both differential mortality and the decelerating rate of increase in mortality with age in late old age due to the selection out of the population of its frailest members.

Expanding the right-hand side of this expression for Gompertz survivorship as a Puiseux series:

$${}_xz_n = \ln\left((\alpha+\lambda)x\right) + \frac{\alpha\beta}{2(\alpha+\lambda)}x + \frac{\alpha\beta^2(\alpha+4\lambda)}{24(\alpha+\lambda)^2}x^2 \ldots \tag{10}$$

where the terms in x^2 and higher powers of x can usually be ignored. Thus, Gompertz survivorship at duration x, offset by log duration, is an approximately linear function of duration:

$$ {}_xz_n - \ln(x) \approx a + bx \tag{11}$$

where $a = \ln(\alpha + \lambda)$ and $b = \alpha\beta/2(\alpha + \lambda)$. For baseline ages, n, at which intrinsic mortality, α, is considerably greater than extrinsic mortality, λ, the slope parameter b will approach 0.5β. Thus, in any population i:

$$ {}_xz_n(i) - \ln(x) \approx a(i) + bx \tag{12}$$

with b, like β, remaining close to constant across a wide range of baseline ages, n, life expectancies at birth, the two sexes, and different families of model life tables.

If this model fits data on human mortality adequately, it provides the basis for a straightforward way of estimating five-year survivorship from data on lifetime survivorship collected at awkward intervals. If the interval separating the surveys is x years and the survivorship of the siblings of cohorts of respondents has been estimated by tabulating the data from the second survey for unconventional age groups:

$$ {}_x\hat{z}_n(i) = {}_xz_n(i) - \left(\ln\left(\frac{x}{5}\right) + \beta(x-5)\right). \tag{13}$$

For example, if two inquiries have been conducted 5 years 9 months apart, one can estimate the gompits of the five-year survivorship probabilities by subtracting $0.14 + 0.75\beta$ from the gompits of survivorship over the interval of 5.75 years.

While this relationship applies in the life table, because:

$$ {}_5p_n \approx \frac{{}_5S_n^{15+}}{{}_5S_{n-5}^{15+}} \tag{14}$$

one will obtain almost the same final estimate by applying the adjustment to the ratio of the proportions of their adult siblings remaining alive reported by a cohort of respondents in two successive inquiries and then predicting ${}_5p_n$ as one would by adjusting an estimate of ${}_xp_n$ directly.

We estimated β for ages, n, of 20 to 45 years at five-year intervals by fitting a regression model to the four regional families of Princeton model life tables by sex with life expectancies at birth of 50, 60, 70, and 80 years. For each n, we calculated ${}_3p_{n-5}$, ${}_4p_{n-5}$ $\cdots {}_7p_{n-5}$ to estimate how survival from a conventional five-year age group into an unconventional one over an awkward interval relates to a conventional ratio. We then predicted ${}_xz_n$ from age, life expectancy at birth (LEB), sex and their first-order interactions, treated as factor variables, and the duration of exposure, x, measured at integer durations of 3 to 7 years but treated as a continuous variable:

$$ {}_xz_n - \ln(x) = \text{Age}_i + \text{Sex}_j + \text{LEB}_k + \text{Age}_i.\text{Sex}_j + \text{Age}_i.\text{LEB}_k + \text{Sex}_j.\text{LEB}_k + \beta x \tag{15}$$

As one would predict, based on our discussion of the Gompertz-Makeham model of mortality, the results suggest that x is close to linearly related to ${}_xz_n - \ln(x)$.

Sub-groups for which values of β are predicted	β for adjusting sibling survival into an unconventional age group ${}_xz_n = \ln\left(-\ln\left(\frac{{}_5S^{15+}_{n-5+x}(t)}{{}_5S^{15+}_{n-5}(t-x)}\right)\right) - \left(\ln\left(\frac{x}{5}\right) + \beta(x-5)\right)$
Overall	0.025
e_0 (50,60,70,80)	0.018 — 0.031
Initial age ($n = 20$ to 45, by 5)	0.012 — 0.046
Sex	0.023 — 0.026
Princeton regional family	0.020 — 0.026

Table S3 Adjustment factor (β) for converting survivorship of siblings over awkward durations, x, of between 3 and 7 years to survivorship for exactly 5 years.

Neither including a quadratic term in x nor estimating the effect of duration using a series of dummy variables for x significantly improves the fit of the model. Moreover, as Vaupel posited [43], inspection of the significance of the coefficients of the interaction terms in an exhaustive set of models, together with their log-likelihoods and AICs, fails to provide any evidence that β varies systematically by age, sex, life expectancy or regional family of Princeton model life tables. Table S3 presents the β coefficient and indicates the extent to which its estimated value varied across different subgroups of life tables.

Thus, in the case of two surveys conducted 5.75 years apart, the adjustment to be subtracted from the gompits of the proportions of siblings surviving between the surveys to obtain measures for survivorship between conventional five-year age groups is $0.14 + 0.025 \times (5.75 - 5) = 0.159$. For two surveys conducted 4.25 years apart with the data from the second survey tabulated for unconventional age groups starting at $n - 0.75$, one would subtract $-0.1625 + 0.025 \times (4.25 - 5) = -0.181$, which is to say add 0.181.

Abbreviations
CDC - Centers for Disease Control and Prevention
DHS - Demographic and Health Surveys
GBD - Global Burden of Disease
MICS - Multiple Indicator Cluster Surveys
SSH - Sibling survival histories
WPP - World Population Prospects

Ethics approval and consent to participate
We conducted a secondary analysis of publicly available and anonymized data from Demographic and Health Surveys. Ethical approval and consent to participate were obtained in the countries and permission to download and analyze the datasets was granted by DHS Program. See `https://dhsprogram.com/methodology/Protecting-the-Privacy-of-DHS-Survey-Respondents.cfm` for DHS informed consent statements.

Consent for publication
Not applicable.

Availability of data and materials
All data are publicly available through `www.dhsprogram.com`. We used the `survey` and `demogsurv` packages of the R statistical software. See `https://github.com/mrc-ide/demogsurv` for `demogsurv`. All R scripts are available upon request.

Competing interests
The authors declare that they have no competing interests.

Funding
This study was supported by the Bill & Melinda Gates Foundation, as part of the *Rapid Mortality Surveillance during COVID-19* (RaMMPS) project (INV-023211).

Authors' contributions
BM and TM carried out the analysis and drafted the manuscript. All authors interpreted the results, edited the manuscript and approved the final version.

Author details

[1]Center for Demographic Research, UCLouvain, Louvain-la-Neuve, Belgium. [2]Department of Population Health, London School of Hygiene & Tropical Medicine, London, UK. [3]Centre for Actuarial Research, University of Cape Town, Cape Town, South Africa.

References

1. Mikkelsen, L., Phillips, D.E., AbouZahr, C., Setel, P.W., de Savigny, D., Lozano, R., Lopez, A.D.: A global assessment of civil registration and vital statistics systems: monitoring data quality and progress. Lancet **386**(10001), 1395–1406 (2015). doi:10.1016/s0140-6736(15)60171-4
2. Hill, K., Choi, Y., Timæus, I.M.: Unconventional approaches to mortality estimation. Demographic Research **13(12)**, 281–300 (2005). doi:10.4054/demres.2005.13.12
3. Moultrie, T., Dorrington, R., Hill, A., Hill, K., Timæus, I., Zaba, B. (eds.): Tools for Demographic Estimation. International Union for the Scientific Study of Population, Paris (2013)
4. Timæus, I.M., Jasseh, M.: Adult mortality in Sub-Saharan Africa: evidence from demographic and health surveys. Demography **41**(4), 757–772 (2004). doi:10.1353/dem.2004.0037
5. Masquelier, B., Reniers, G., Pison, G.: Divergences in mortality trends in sub-Saharan Africa: survey evidence on the survival of children and siblings. Population Studies **68**(2), 161–177 (2014). doi:10.1080/00324728.2013.856458
6. Bendavid, E., Holmes, C., Bhattacharya, J., Miller, G.: HIV development assistance and adult mortality in Africa. Journal of the American Medical Association **307**(19), 2060–2067 (2012). doi:10.1001/jama.2012.2001
7. de Walque, D., Verwimp, P.: The demographic and socio-economic distribution of excess mortality during the 1994 genocide in Rwanda. Journal of African Economies **19**(2), 141–162 (2010). doi:10.1093/jae/ejp029
8. Obermeyer, Z., Murray, C., Gakidou, E.: Fifty years of violent war deaths from Vietnam to Bosnia: analysis of data from the world health survey programme. British Medical Journal **336**, 1482–1486 (2008). doi:10.1136/bmj.a137
9. Wagner, Z., Heft-Neal, S., Wise, P.H., Black, R.E., Burke, M., Boerma, T., Bhutta, Z.A., Bendavid, E.: Women and children living in areas of armed conflict in Africa: a geospatial analysis of mortality and orphanhood. The Lancet Global Health **7**(12), 1622–1631 (2019). doi:10.1016/s2214-109x(19)30407-3
10. Wang, H., *et al.*: Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950–2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019. The Lancet **396**(10258), 1160–1203 (2020). doi:10.1016/S0140-6736(20)30977-6
11. United Nations: World Population Prospects: The 2022 Revision. United Nations, New York (2022). https://esa.un.org/unpd/wpp/
12. Helleringer, S., Pison, G., Masquelier, B., Kanté, A., Douillot, L., Duthé, G., Sokhna, C., Delaunay, V.: Improving the quality of adult mortality data collected in demographic surveys: Validation study of a new siblings' survival questionnaire in Niakhar, Senegal. PLoS Medicine **11**(5), 1001652 (2014). doi:10.1371/journal.pmed.1001652
13. Adjiwanou, V., et al.: Measuring excess mortality during the COVID-19 pandemic in low- and lower-middle income countries: the need for mobile phone surveys. Center for Open Science (2020). doi:10.31235/osf.io/4bu3q. https://doi.org/10.31235/osf.io/4bu3q
14. Rutenberg, N., Sullivan, J.: Direct and indirect estimates of maternal mortality from the sisterhood method. Proceedings of the Demographic and Health Surveys World Conference, August 5-7 1991, Washington, D.C., Vol. III. IRD/MACRO International Inc., Columbia , Maryland. (1991)
15. Graham, W., Brass, W., Snow, R.: Estimating maternal mortality : the sisterhood method. Studies in Family Planning **20(3)**, 125–135 (1989). doi:10.2307/1966567
16. Timæus, I., Zaba, B., Ali, M.: Estimation of adult mortality from data on adult siblings. In: Zaba, B., Blacker, J. (eds.) Brass Tacks: Essays in Medical Demography, pp. 43–66. Athlone Press, London (2001)
17. Zlotnik, H., Hill, K.: The use of hypothetical cohorts in estimating demographic parameters under conditions of changing fertility and mortality. Demography **18(1)**, 103–122 (1981). doi:10.2307/2061052
18. Timæus, I.M.: An assessment of methods for estimating adult mortality from two sets of data on maternal orphanhood. Demography **23 3**, 435–450 (1986). doi:10.2307/2061440
19. Keyfitz, N., Caswell, H.: The demographic theory of kinship. In: Keyfitz, N., Caswell, H. (eds.) Applied Mathematical Demography (3rd Ed.), pp. 370–398. Springer, New York (2005). Chap. 15
20. Hill, K., Trussell, J.: Further developments in indirect mortality estimation. Population Studies **31(2)**, 313–334 (1977). doi:10.2307/2173920
21. Masquelier, B.: Sibship sizes and family sizes in survey data used to estimate mortality. Population-E **69**(2), 249–268 (2014). doi:10.3917/popu.1402.0249
22. Garenne, M., Friedberg, F.: Accuracy of indirect estimates of maternal mortality: A simulation model. Studies in Family Planning **28**(2), 132–142 (1997). doi:10.2307/2138115
23. Saifuddin, A., Li, Q., Scrafford, C., Pullum, T.: An Assessment of DHS Maternal Mortality Data and Estimates. DHS Methodological Reports No. 13 (2014)
24. Obermeyer, Z., Rajaratnam, J., Park, C., Gakidou, E., Hogan, M., Lopez, A., Murray, C.: Measuring adult mortality using sibling survival: A new analytical method and new results for 44 countries, 1974-2006. PLoS Medicine **7**(4), 1000260 (2010). doi:10.1371/journal.pmed.1000260
25. Trussell, J., Rodriguez, G.: A note on the sisterhood estimator of maternal mortality. Studies in Family Planning **21 (6)**, 344–346 (1990). doi:10.2307/1966923
26. Masquelier, B.: Adult mortality from sibling survival data: A reappraisal of selection biases. Demography **50**(1), 207–228 (2013). doi:10.1007/s13524-012-0149-1
27. Gakidou, E., King, G.: Death by survey: estimating adult mortality without selection bias from sibling survival data. Demography **43**(3), 569–585 (2006). doi:10.1353/dem.2006.0024

28. Feehan, D.M., Borges, G.M.: Estimating adult death rates from sibling histories: A network approach. Demography **58**(4), 1525–1546 (2021). doi:10.1215/00703370-9368990
29. Timæus, I.: Estimation of adult mortality from sibling histories. In: Moultrie, T., Dorrington, R., Hill, A., Hill, K., Timæus, I., Zaba, B. (eds.) Tools for Demographic Estimation. International Union for the Scientific Study of Population, Paris (2013)
30. Zimbabwe National Statistics Agency and I. C. F.: Zimbabwe Demographic and Health Survey 2015: Final Report. Zimbabwe National Statistics Agency (ZIMSTAT) and ICF International, Rockville, Maryland, USA (2016). http://dhsprogram.com/pubs/pdf/FR322/FR322.pdf
31. Coale, A., Demeny, P., Vaughan, B.: Regional Model Life Tables and Stable Populations (2nd Ed.). Academic Press, New York (1983)
32. Gregson, S., Mugurungi, O., Eaton, J., Takaruza, A., Rhead, R., Maswera, R., Mutsvangwa, J., Mayini, J., Skovdal, M., Schaefer, R., Hallett, T., Sherr, L., Munyati, S., Mason, P., Campbell, C., Garnett, G.P., Nyamukapa, C.A.: Documenting and explaining the hiv decline in east zimbabwe: the manicaland general population cohort. BMJ Open **7**(10) (2017). doi:10.1136/bmjopen-2017-015898
33. Wachter, K., Blackwell, D., Hammel, E.: Testing the validity of kinship microsimulation. Journal of Mathematical and Computer Modeling **26**, 89–104 (1997). doi:10.1016/s0895-7177(97)00172-6
34. Zagheni, E.: The impact of the HIV/AIDS epidemic on kinship resources for orphans in Zimbabwe. Population and Development Review **37**(4), 761–783 (2011). doi:10.1111/j.1728-4457.2011.00456.x
35. Verdery, A.M., Smith-Greenaway, E., Margolis, R., Daw, J.: Tracking the reach of COVID-19 kin loss with a bereavement multiplier applied to the United States. Proceedings of the National Academy of Sciences **117**(30), 17695–17701 (2020). doi:10.1073/pnas.2007476117
36. Alburez-Gutierrez, D., Mason, C., Zagheni, E.: The "sandwich generation" revisited: Global demographic drivers of care time demands. Population and Development Review **47**(4), 997–1023 (2021). doi:10.1111/padr.12436
37. Brass, W.: On the scale of mortality. In: Brass, W. (ed.) Biological Aspects of Demography, pp. 69–110. Taylor et Francis, London (1971)
38. Brass, W.: Perspectives in population prediction: illustrated by the statistics of England and Wales. Journal of the Royal Statistical Society **137**, 532–583 (1974). doi:10.2307/2344713
39. Booth, H.: Transforming Gompertz's function for fertility analysis: the development of a standard for the relational Gompertz function. Population Studies **38(3)**, 495–506 (1984). doi:10.2307/2174137
40. Masquelier, B., Kanyangarara, M., Pison, G., Kanté, A., Ndiaye, C., Douillot, L., Duthé, G., Sokhna, C., Delaunay, V., Helleringer, S.: Errors in reported ages and dates in surveys of adult mortality: A record linkage study in Niakhar (Senegal). Population Studies, 1–19 (2021). doi:10.1080/00324728.2020.1854332
41. Reniers, G., Masquelier, B., Gerland, P.: Adult mortality in Africa. In: Rogers, R., Crimmins, E. (eds.) International Handbook of Adult Mortality. Springer, New York (2011)
42. Masquelier, B., Eaton, J.W., Gerland, P., Pelletier, F., Mutai, K.: Age patterns and sex ratios of adult mortality in countries with high HIV prevalence. AIDS **31 Suppl 1**, 77–85 (2017). doi:10.1097/QAD.0000000000001332
43. Vaupel, J.W.: Biodemography of human ageing. Nature **464**(7288), 536–542 (2010). doi:10.1038/nature08984
44. Stanton, C., Noureddine, A., Hill, K.: An assessment of DHS maternal mortality indicators. Studies in Family Planning **31**(2), 111–123 (2000). doi:10.1111/j.1728-4465.2000.00111.x
45. Checchi, F., Roberts, L.: Documenting mortality in crises: What keeps us from doing better? PLoS Medicine **5**(7), 146 (2008). doi:10.1371/journal.pmed.0050146
46. Rajaratnam, J., Linda, T., Lopez, A., Murray, C.: Measuring under-five mortality: Validation of new low-cost methods. PLoS Medicine **7**(4) (2010). doi:10.1371/journal.pmed.1000253
47. Hill, K., Brady, E., Zimmerman, L., Montana, L., Silva, R., Amouzou, A.: Monitoring change in child mortality through household surveys. PLOS ONE **10**(11), 0137713 (2015). doi:10.1371/journal.pone.0137713

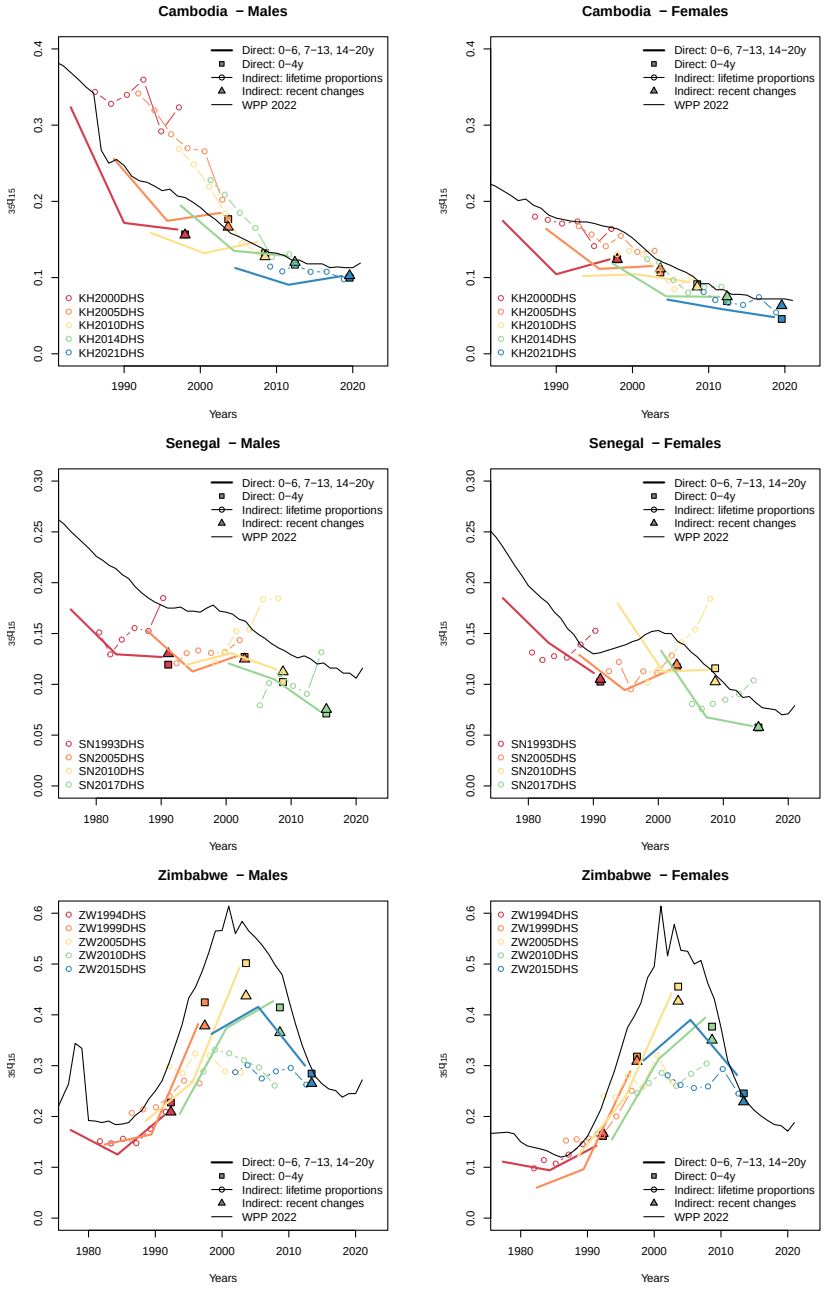

Figure 1 Estimates of the probability of dying between ages 15 and 50 ($_{35}q_{15}$) from SSH data using different methods in Cambodia, Senegal and Zimbabwe.
Note: one set of direct estimates is calculated for periods of 0-6, 7-13, and 14-20 completed years before the survey (solid lines). Another set is calculated for the period 0-4 completed years before the survey (squares) to compare with indirect estimates derived from recent changes in proportions (triangles).

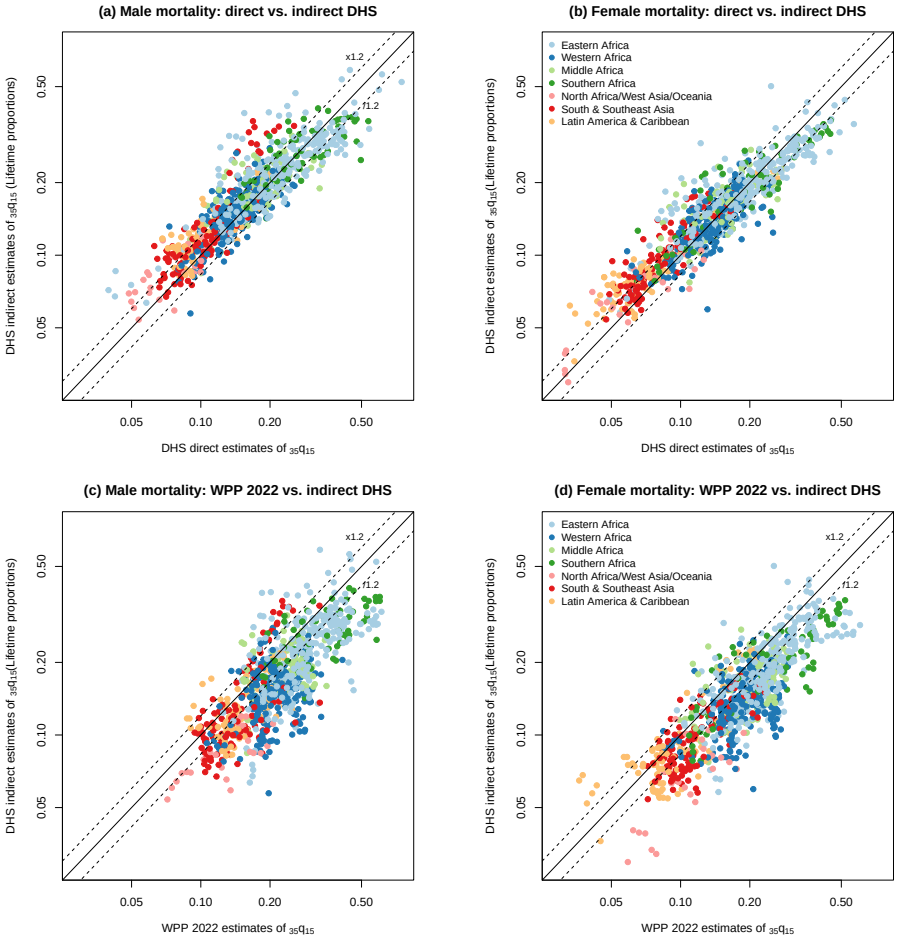

Figure 2 Indirect estimates of the probability of dying between ages 15 and 50 ($_{35}q_{15}$) obtained from adult lifetime proportions (0-15 years before data collection), compared to direct estimates from full SSH and the 2022 World Population Prospects

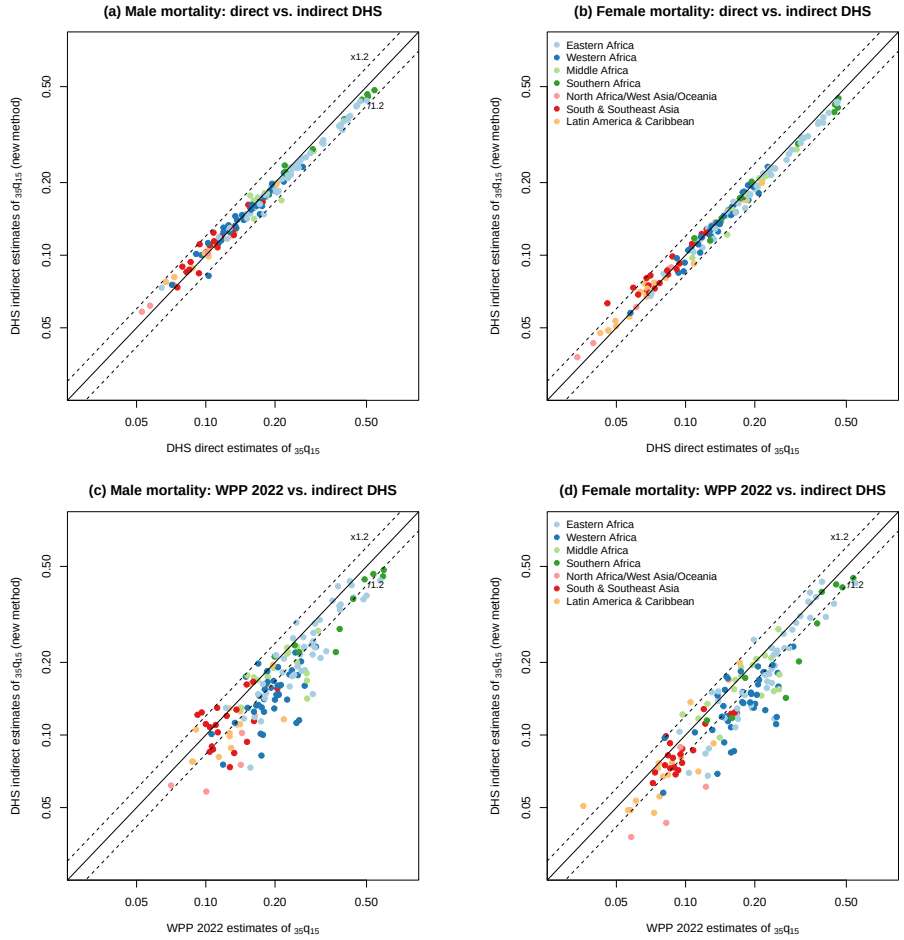

Figure 3 Indirect estimates of the probability of dying between ages 15 and 50 $_{35}q_{15}$ probability obtained from changes in proportions surviving (0-4 years before data collection), compared to direct estimates from full SSH and the 2022 World Population Prospects

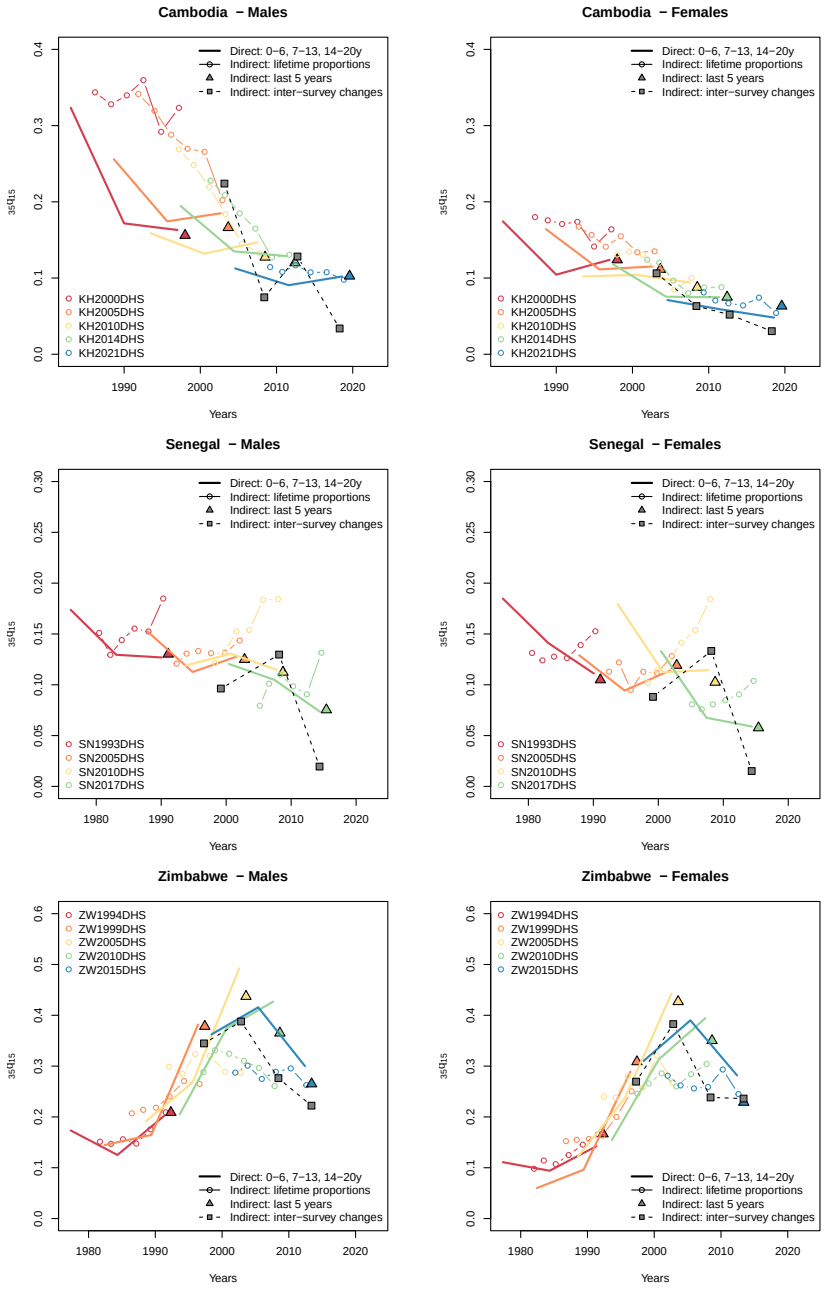

Figure 4 Estimates of the probability of dying between ages 15 and 50 ($_{35}q_{15}$) based on the full SSH, on lifetime survival of adult siblings, on the survival to survey of cohorts of siblings alive 5 years earlier, and on cohort changes in sibling survival in between successive surveys

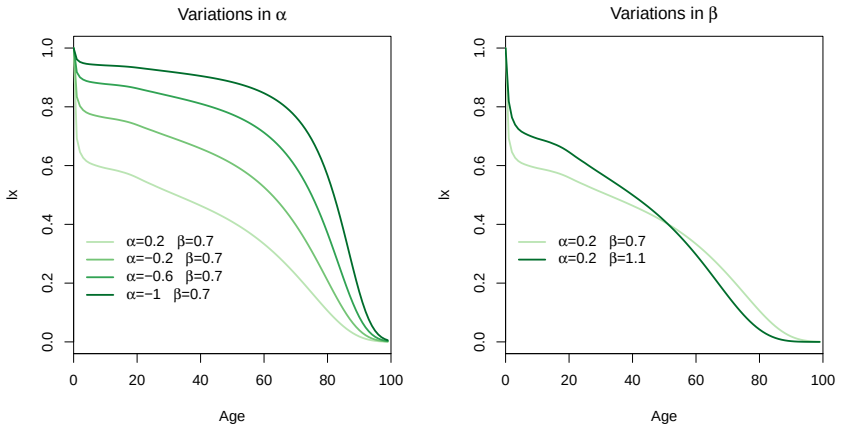

Figure S1 Age patterns of mortality introduced in the simulations, Brass General Standard [37]

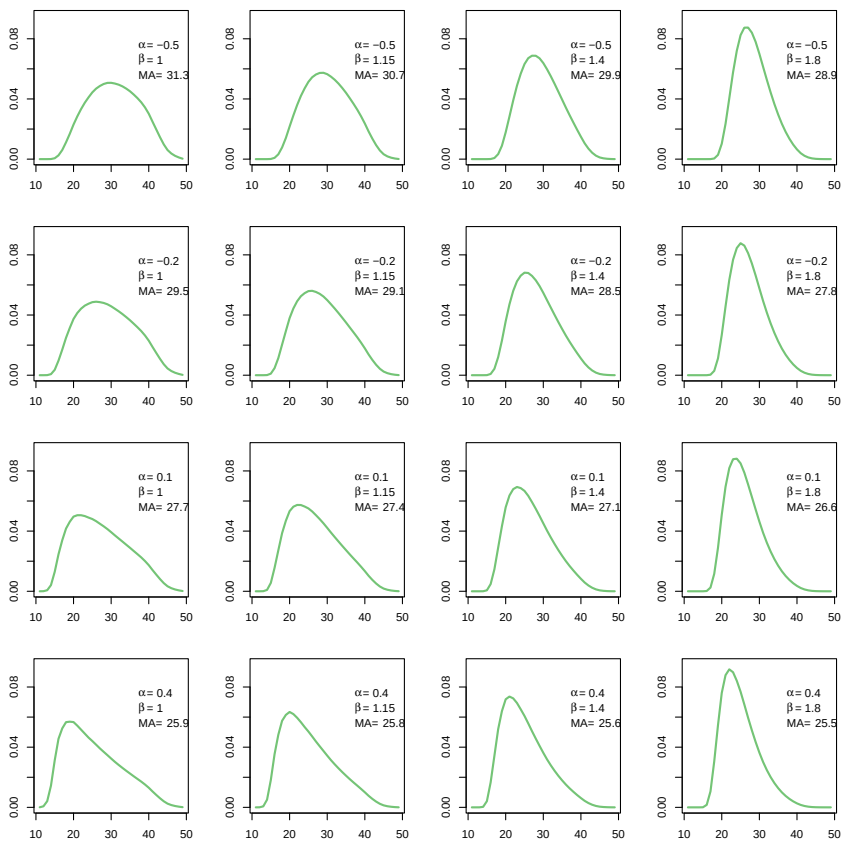

Figure S2 Age patterns of fertility introduced in the simulations, Booth standard [39] - MA is the mean age of the fertility schedule

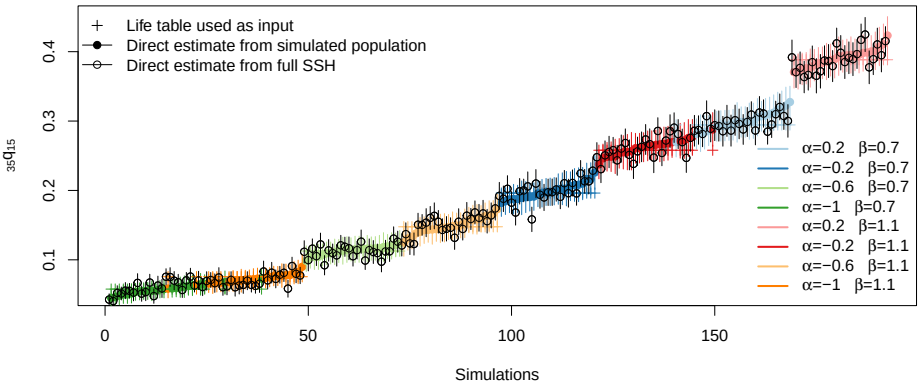

Figure S3 Probability of dying between ages 15 and 50 ($_{35}q_{15}$) introduced in the simulations, re-estimated from deaths in the overall populations and estimated directly from full SSH, for the different mortality age patterns

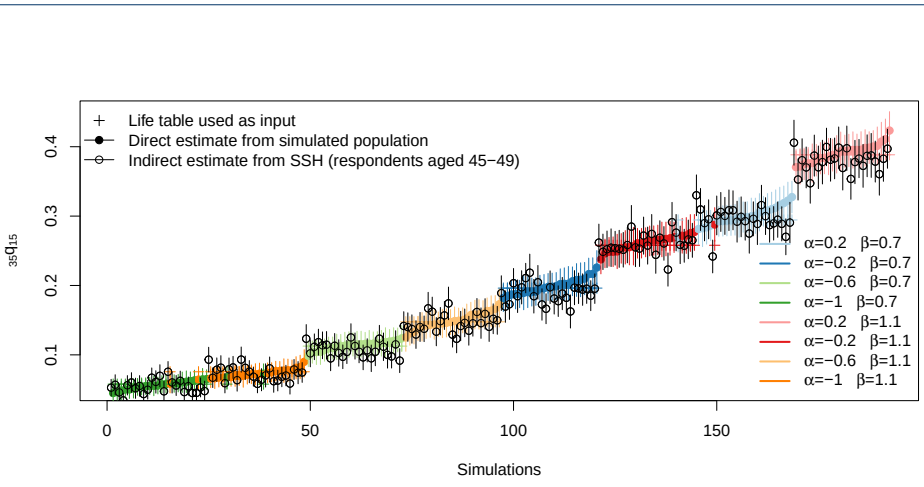

Figure S4 Probability of dying between ages 15 and 50 ($_{35}q_{15}$) introduced in the simulations, re-estimated from deaths in the overall populations, and estimated indirectly from adult lifetime proportions, for the different mortality age patterns