## A. FORMAL RELATION

It can be shown that the outsurvival statistic relates to the joint probability density function of two lifespan distributions, which gives the probability of realizations of two lifespans and thus is related to the overlap of the two distributions. Assume two populations of individuals, with ages at death *x* and *y*, respectively. Assume the two populations are independent, meaning that the length of life *x* does not depend on the length of life *y* and vice versa. This implies that the joint probability density function, $d_{1,2}(x,y)$, equals the product of the marginal densities so that $d_{1,2}(x,y) = d_1(x)\, d_2(y)$. We are interested in calculating the probability ($\varphi$) of individuals in the first population outliving those in the second population. This implies that $0 \leq y < x$ so:

$$\begin{aligned}\varphi &= \int_0^\infty \int_0^x d_{1,2}(x,y)dydx = \int_0^\infty d_1(x)\int_0^x d_2(y)dydx \\ &= \int_0^\infty d_1(x)D_2(x)dx \\ &= \int_0^\infty d_1(x)[1-l_2(x)]dx = 1-\int_0^\infty d_1(x)l_2(x)dx \\ &= \textstyle\int_0^\infty d_2(x)\ell_1(x)dx.\end{aligned} \tag{A1}$$

Following the same approach, we can find the complement of $\varphi$, labeled $\varphi'$, which is the probability of individuals in the second population to outlive those in the first:

$$\begin{aligned}\varphi' &= \int_0^\infty \int_0^y d_{1,2}(x,y)dxdy \\ &= \int_0^\infty d_2(y)\int_0^y d_1(x)dxdy \\ &= \int_0^\infty d_2(y)D_1(y)dy \\ &= \int_0^\infty d_2(y)[1-l_1(y)]\,dy \\ &= 1-\int_0^\infty d_2(y)l_1(y)\,dy.\end{aligned} \tag{A2}$$

From Equations (B1) and (B2) it can be shown that $\varphi + \varphi' = 1$. Thus, $\varphi$ is also equal to:

$$\begin{aligned}\varphi &= 1-\varphi' \\ &= 1-\left[1-\int_0^\omega d_2(x)l_1(x)\,dx\right] \\ &= \int_0^\omega d_2(x)l_1(x)\,dx.\end{aligned} \tag{A3}$$

## B. SIMULATIONS AND DISCREATE APPROXIMATION

We simulated age at death distributions, using the Gompertz model, using various scale (M) and shape (β) parameters [1]. The distributions were first found using an age width (*n*) of 0.0001, after



Bergeron-Boucher M-P, *et al. BMJ Open* 2022; 12:e059964. doi: 10.1136/bmjopen-2021-059964

which the data were aggregated within 1-year and 5-years age-groups. The probability that individuals in both population died within the same age-group, $\sum_{x=0}^{\omega} {}_nd_x^1\, {}_nd_x^2$, was then redistributed between $\varphi$ and $\varphi'$ based on two assumptions: equal (equation B1) and proportional redistributions (equation B2). The results are presented in Table B1.

$$\varphi \approx \sum_{x=0}^{\omega} {}_nd_x^1\ {}_nD_{x-n}^2 + \frac{\sum_{x=0}^{\omega} {}_nd_x^1\, {}_nd_x^2}{2} \quad \text{(B1)}$$

$$\varphi \approx \sum_{x=0}^{\omega} {}_nd_x^1\ {}_nD_{x-n}^2 + \sum_{x=0}^{\omega} {}_nd_x^1\, {}_nd_x^2 \frac{{}_nd_x^1\ {}_nD_{x-n}^2}{{}_nd_x^1\ {}_nD_{x-n}^2 + {}_nd_x^2\ {}_nD_{x-n}^1} \quad \text{(B2)}$$

The simulations show that equally redistributing $\sum_{x=0}^{\omega} {}_nd_x^1\, {}_nd_x^2$ between the two other statistics provide very similar results to the continuous data (*n=0.0001*), especially for the 1-year age-group. More differences are found when aggregating by 5-years age-groups, but the difference in $\varphi$ between the different age-width remains less than 1 percentage point, when equally redistributing $\sum_{x=0}^{\omega} {}_nd_x^1\, {}_nd_x^2$.

*Table B1. Assumptions to redistribute $\sum_{x=0}^{\omega} {}_nd_x^1\, {}_nd_x^2$ for different mortality scenarios.*

| | $\varphi$ | $\varphi'$ | $\sum_{x=0}^{\omega} {}_nd_x^1\, {}_nd_x^2$ | Eq. B1 | Eq. B2 |
|---|---|---|---|---|---|
| Gompertz: $M_A = 61, M_B = 65,\ \beta_A = 0.12,\ \beta_B = 0.14$ | | | | | |
| Continuous | **36.3** | 63.7 | 0.0 | - | - |
| 1-year | 34.8 | 62.2 | 3.0 | **36.3** | 35.9 |
| 5-years | 28.2 | 55.8 | 15.0 | **36.7** | 34.3 |
| Gompertz: $M_A = 61, M_B = 70,\ \beta_A = 0.10,\ \beta_B = 0.14$ | | | | | |
| Continuous | **23.6** | 76.4 | 0.0 | - | - |
| 1-year | 22.5 | 75.2 | 2.3 | **23.6** | 23.0 |
| 5-years | 18.5 | 70.0 | 11.3 | **24.2** | 20.9 |
| Gompertz: $M_A = 68, M_B = 70,\ \beta_A = 0.13,\ \beta_B = 0.14$ | | | | | |
| Continuous | **42.8** | 57.2 | 0.0 | - | - |
| 1-year | 41.2 | 55.5 | 3.3 | **42.8** | 42.6 |
| 5-years | 34.9 | 48.8 | 16.3 | **43.0** | 41.7 |
| Gompertz: $M_A = 69, M_B = 70,\ \beta_A = 0.10,\ \beta_B = 0.12$ | | | | | |
| Continuous | **46.1** | 53.9 | 0.0 | - | - |
| 1-year | 44.7 | 52.6 | 2.7 | **46.1** | 46.0 |
| 5-years | 39.4 | 47.2 | 13.4 | **46.1** | 45.5 |

To further test the model and the redistribution of $\sum_{x=0}^{\omega} {}_nd_x^1\, {}_nd_x^2$, we simulated 100,000 individual lifespans from an exponential distribution with piece-wise constant rates [2]. We performed this procedure for every population and by sex using as an input empirical death rates retrieved from the HMD [3]. Then we randomly paired males and females and calculated the proportions of males outliving the paired female. Table 3 compares the discrete approach introduced in the main document (eq. B1) and the continuous approach based on simulations. Both approaches provided very similar results.

*Table B2. Proportions of males outliving females based on a discrete and continuous approach (simulations).*

| | % males outliving females | | % females outliving males | |
|---|---|---|---|---|
| | Continuous | Discrete | Continuous | Discrete |
| Denmark | | | | |
| 1850 | 46.1 | 46.4 | 53.9 | 53.6 |
| 1900 | 45.8 | 45.8 | 54.2 | 54.2 |
| 1950 | 46.2 | 46.3 | 53.8 | 53.7 |
| 2016 | 40.7 | 40.6 | 59.3 | 59.4 |
| France | | | | |
| 1850 | 48.6 | 48.5 | 51.4 | 51.5 |
| 1915 | 25.6 | 25.5 | 74.4 | 74.5 |
| 1950 | 39.8 | 39.8 | 60.2 | 60.2 |
| 2016 | 35.9 | 36.0 | 64.1 | 64.0 |
| Japan | | | | |
| 1950 | 44.1 | 44.2 | 55.9 | 55.8 |
| 2016 | 33.8 | 33.7 | 66.2 | 66.3 |
| Russia | | | | |
| 1960 | 35.6 | 35.5 | 64.4 | 64.5 |
| 2014 | 30.2 | 30.0 | 69.8 | 70.0 |

## C. DATA

*Table C1. Countries/ regions and years with available data in the HMD*

| Country\region | Years | Country\region | Years |
|---|---|---|---|
| Australia | 1921-2018 | Japan | 1947-2019 |
| Austria | 1947-2019 | Latvia | 1959-2019 |
| Belarus | 1959-2018 | Lithuania | 1959-2019 |
| Belgium | 1841-2018 | Luxembourg | 1960-2019 |
| Bulgaria | 1947-2017 | Netherlands | 1850-2019 |
| Canada | 1921-2018 | New Zealand | 1948-2013 |
| Chile | 1992-2017 | Norway | 1846-2020 |
| Croatia | 2001-2019 | Poland | 1958-2019 |
| Czechia | 1950-2019 | Portugal | 1940-2018 |
| Denmark | 1835-2020 | Republic of Korea | 2003-2018 |
| Estonia | 1959-2019 | Russia | 1959-2014 |
| Finland | 1878-2019 | Slovakia | 1950-2017 |
| France | 1816-2018 | Slovenia | 1983-2017 |
| Germany-East | 1956-2017 | Spain | 1908-2018 |
| Germany-West | 1956-2017 | Sweden | 1751-2019 |
| Greece | 1981-2017 | Switzerland | 1876-2018 |
| Hong Kong | 1986-2017 | Taiwan | 1970-2019 |
| Hungary | 1950-2017 | UK – England and Wales | 1841-2018 |
| Iceland | 1838-2018 | UK- Scotland | 1855-2018 |
| Ireland | 1950-2017 | UK- Northern Ireland | 1922-2018 |
| Israel | 1983-2016 | USA | 1933-2019 |
| Italy | 1872-2018 | Ukraine | 1959-2013 |



Bergeron-Boucher M-P, *et al. BMJ Open* 2022; 12:e059964. doi: 10.1136/bmjopen-2021-059964

## D. EXTRA ANALYSIS: COHORTS, WWP AND MALES' STANDARD DEVIATION

Similar relations as those for period data were also found for cohorts (Figure D1). In the HMD, life table for cohorts were only available for 11 countries: Denmark, England and Wales, Finland, France, Iceland, Italy, Netherlands, Norway, Scotland, Sweden and Switzerland. For cohorts with complete mortality history, the proportions of males outliving females varied between 35% and 49%. Only small changes in $\varphi$ were observed for cohorts born prior to 1870-1890, with $\varphi$ varying around 46.5%. For the cohorts born afterwards, $\varphi$ decreases, reaching a mean of 38.4% for the cohort born in 1925, with values varying between 35.3% (Finland) and 40.4% (Scotland).

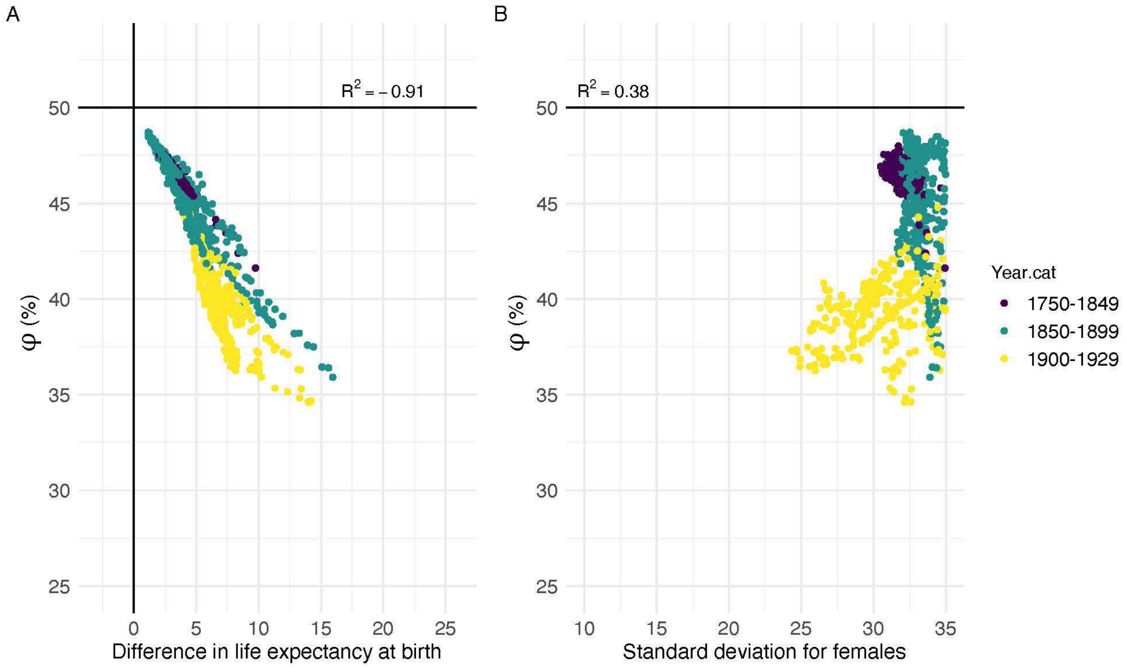


**Figure D1** Relation between $\varphi$ and (a) the sex differences in life expectancy and (b) the standard deviation for females for HMD cohort data.
Source: HMD [3] and authors' own calculations.

Figure D2 also shows the relation between $\varphi$ and (A) the sex differences in life expectancy at birth and (B) the standard deviation for females for countries in the WWP from 1950-55 to 2015-19. The relation between $\varphi$ and the two measures is similar to that shown in the main text using the HMD data.

Figure D3 shows the relation between (A) the relation between $\varphi$ at the standard deviation of the lifespan distribution from birth for males and (B) the same relation, but conditional to survival to age 50. The relation between $\varphi$ and the SD is similar whether we used the SD for females (as in the main text) or for males (Figure D3).

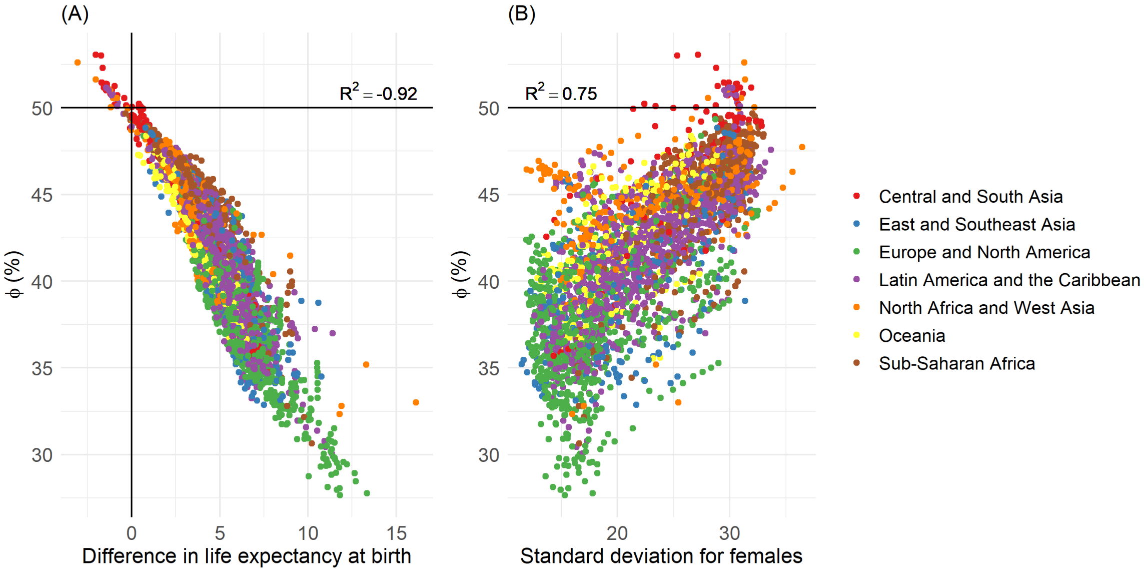


**Figure D2** Relation between $\varphi$ and (a) the sex differences in life expectancy and (b) the standard deviation for females for WPP data.
Source: WPP [4] and authors' own calculations.

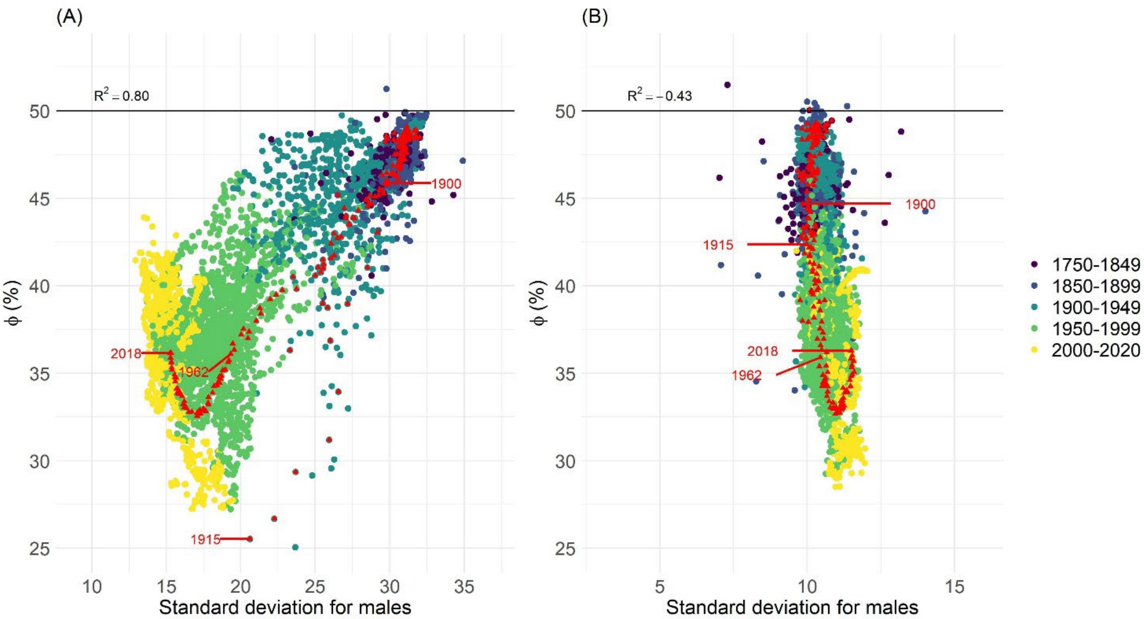


**Figure D3** Relation between (a) $\varphi$ and the standard deviation for males form birth and (b) $\varphi$ and the standard deviation for males form age 50 for HMD period data.
Source: HMD [3] and authors' own calculations.



Bergeron-Boucher M-P, *et al. BMJ Open* 2022; 12:e059964. doi: 10.1136/bmjopen-2021-059964

## E. OTHER MEASURES OF OVERLAP

Figure E1 shows the relation between $\varphi$ and the stratification index used by Shi et al. [5]. Both indicators are strongly correlated with a correlation coefficient of 0.98. Figure E2 shows a similar relation between $\varphi$ and the Kullback-Leibler divergence, with a correlation coefficient of -0.93.

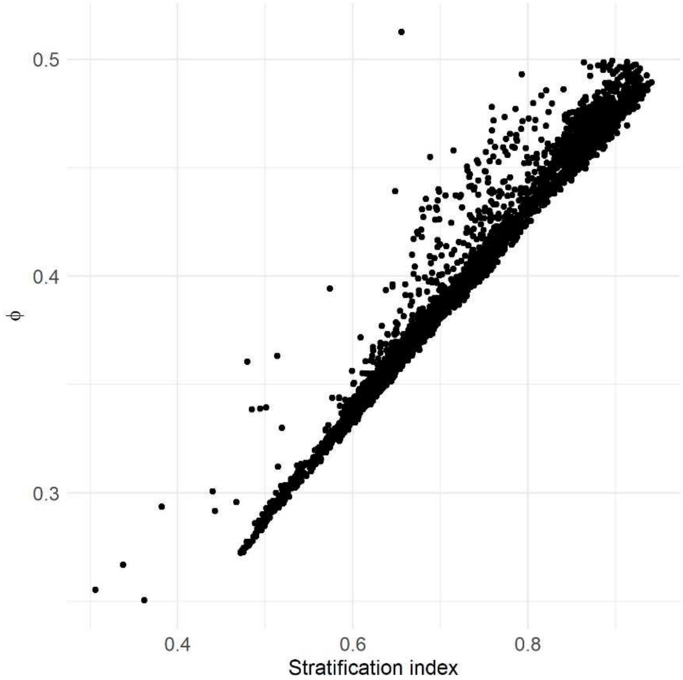


**Figure E1**. Relation between $\varphi$ and the stratification index for HMD period data. Source: HMD [3] and authors' own calculations.

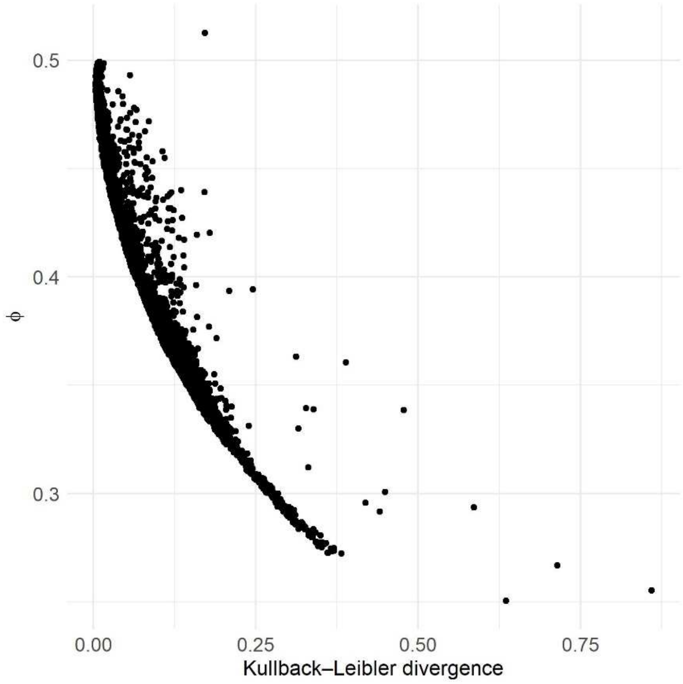


**Figure E2**. Relation between $\varphi$ and the Kullback-Leibler divergence for HMD period data. Source: HMD [3] and authors' own calculations.

## REFERENCES

1. Missov TI, Lenart A, Nemeth L, Canudas-Romo V, Vaupel JW: **The Gompertz force of mortality in terms of the modal age at death**. *Demographic Research* 2015, **32**:1031-1048.
2. Willekens F: **Continuous-time microsimulation in longitudinal analysis**. In: *New frontiers in microsimulation modelling.* edn. Edited by Harding A, Zaidi A. London: Routledge; 2009: 353-376.
3. HMD: **Human Mortality Database**. In.: University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany), Available at www.mortality.org (Accessed on July 15) . 2021.
4. United Nations: **World Population Prospects 2019, Online Edition. Rev. 1.** In.: Department of Economic and Social Affairs, Population Division, Available at https://population.un.org/wpp/Download/Standard/Mortality/ (Accessed on May 21); 2019.
5. Shi J, Aburto JM, Martikainen P, Tarkiainen L, van Raalte AA: **A distributional approach to measuring lifespan stratification**. *Population Studies* 2022, **(Accepted)**.