

How many lives do we live?

A. Remund

November 1 2019

This is a draft version, please do not cite!

Introduction

How many times did Jeanne Calment cheat death before becoming the oldest person to have ever lived? How exceptional was this feat, compared, for instance, to becoming a centenarian in the 1700s? Usual measures of longevity only give partial answers to this question. As human longevity has increased steadily in most countries during the last century, more and more people reach ages that used to be considered as advanced or even exceptional. For instance, the proportion of person-years lived above 100 years of age has been multiplied by more than 100 in the Netherlands since the 1950s (HMD (2019)). This changing meaning of old age is raising the question of what it takes to become a (very) old person in a given mortality context.

In a nutshell, the dominant model of ageing is the one proposed by Strehler and Mildvan, which postulates that senescence originates from a constant battle between the internal “vitality” of individuals and the external “challenges” that they face (Strehler and Mildvan (1960)). With declining vitality, individuals become more likely to die. From an individual point of view, each time somebody survives one of these random challenges, he or she has “cheated death”. Translated at the population level, this idea can be expressed through the ability to survive to an “expectable” age. In this framework, someone who survives to the life expectancy at birth can be considered to have beaten the odds. Consequently, becoming an “old” person can be measured in terms of outliving a reasonable age given the current mortality context.

This is not exactly the same as a classic life table, which states that, in order to reach a given age t , one needs to survive, year after year, to the next age until reaching t . The probability to reach t is thus simply the value of the survival function l_x at age t , i.e. l_t . An alternative way to represent this process is to think in terms of successively reaching reasonable age milestones. According to this idea, each individual is born with a given number of years of life that he or she can expect to reach. This value is the life expectancy at birth (e_0). An individual who survives to this age receives a “new lease of life”, which is smaller than the initial one if mortality is increasing with age, and corresponds to the value of the remaining life expectancy at age $x = e_0$. By pursuing this reasoning, reaching age t requires thus to repeat this feat successively until reaching age t . This process of successfully beating the odds in order to cheat death a enough times to reach (very) old age can be used as the corner stone of a new measure of (successful) ageing. This measure does not use years of age as its basic unit of time, but instead exploits the concept of successfully “cheating death” from one “life” to the next. Its unit is thus a “life” and it captures longevity by measuring how many “lives” n it takes to reach a given age t . This idea bears resemblances to those of “repeated resuscitations” (Vaupel and Yashin (1987)) and of “lifesaving” (Finkelstein (2005)), albeit without the comparative dimension of contrasting different mortality scenarios.

In this extended abstract I will define and explore the properties of this new proposed measure. I will then apply it to a selection of past and present populations and compare it with other classical measures of mortality such as life expectancy and survival probabilities.

Methods

Definitions

As mentioned above, the cornerstone of the proposed measure is the *Expected Age at Death*, i.e. the age that people can expect to reach, provided that they have survived to a given age x .

$$EAD(x) = x + e(x) \tag{1}$$

At birth, $EAD(x)$ is simply equal to life expectancy, since $EAD(0) = e(0) + 0$. For the individuals who survive to $EAD(0)$, their updated EAD is given by $EAD(EAD(0)) = e(e(0)) + e(0)$. Given an arbitrary target age t , I define n as the number of times that $EAD(x)$ needs to be successively applied to itself until $EAD(x) \geq t$. Algebraically, this is equivalent to saying that $EAD(x)$ is composed with itself n times.

$$n \text{ such that } EAD(x)^{\circ n} \geq t \tag{2}$$

The computation of n can be done with a loop, by successively applying $EAD(x)$ to the value of its previous iteration until reaching $EAD(x) \geq t$. Visually, this method consists in starting from age 0, move vertically up to the curve of $EAD(x)$, bounce to the x-axis until the value of $EAD(0)$, and repeat this until $EAD(x) \geq t$ (Figure 1). This requires interpolating values of $EAD(x)$ for non-integer ages x , for instance using a monotonic spline, since $EAD(x)$ is by definition monotonically increasing because $e(x)$ is positive.

$$i = 1 \quad : \quad EAD(0) = e(0) + 0 < t \tag{3}$$

$$i = 2 \quad : \quad EAD(EAD(0)) = EAD \circ EAD = EAD^{\circ 2} = e(e(0)) + e(0) < t \tag{4}$$

$$\dots \tag{5}$$

$$i = n \quad : \quad EAD(EAD(\dots(EAD(0)))) = EAD^{\circ n} \geq t \tag{6}$$

Figure 1 shows that, in the Netherlands in 2015 (HMD (2019)), it took 6 iterations of $EAD(x)$ (or “lives”) to reach age 100, i.e. $n = 6$. It is possible to consider non-integer values of n if one considers the possibility of fractional iterates (Isaacs (1950)). Empirically, this requires using the inverse function of $EAD(x)$ and looking for the value of x , such that $EAD(x) = t$ exactly. This value can be called t^* , and is the age at which the remaining life expectancy allows to reach exactly t . This value can be used to identify the exact number of iterations ($n^* \in \mathbb{R}$) such that $EAD(n) = t$, by using the inverse function of $EAD(i)$, such that $EAD(x)^{\circ n^*} = t$ (see Appendix for the detailed procedure). This fractional value n^* can also be thought as a continuous function of $tn^* = n(t)$, i.e. the exact number of times someone needs to survive to his or her conditional expected age at death in order to reach age t .

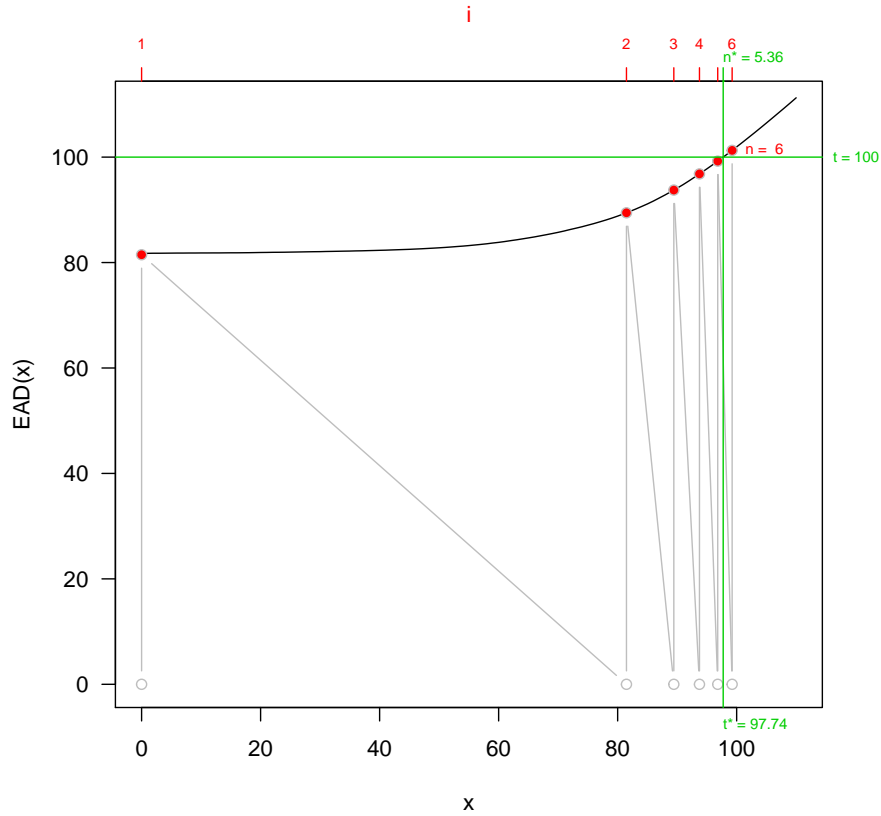


Figure 1: Empirical computation of the number of lives required to become centenarian (Netherlands 2015)

The function $n(t)$ does unfortunately not have a closed expression in terms of t , since it is indirectly defined through $EAD(x)$. It is however possible to approximate it with surprising precision. To understand how, let us define the probability to survive to age t as

$$l(t) = \prod_{i=1}^n p_i \quad (7)$$

where p_i is the probability to survive to the next iteration of $EAD(x)$. Since p_i does not depend on t , this is not sufficient to express n as a function of t . However, we can proceed by analogy to reach a good approximation. Let us imagine that, instead of working with life expectancy as a measure of reasonably attainable age, we used the *Median Age at Death* (MAD) of those alive at age x . By definition, half of the people alive at age x will survive to $MAD(x)$, or in life table terms $\frac{l(MAD(x))}{l(x)} = 0.5$. In this slightly different scenario, surviving to age t means repeatedly surviving to successive values of $MAD(x)$, i.e. n is defined as the number of iterations of $MAD(x)$ until $MAD(x) \geq t$, i.e. n such that $MAD(x)^{on} \geq t$. Since in this case $p(i) = 0.5 \forall i$, the probability to survive to age t is equal to the product of the probability to survive to each iteration of $MAD(x)$, so

$$l(t) = p(i)^n = p^n = 0.5^n \quad (8)$$

Isolating n , we obtain

$$n(t) = \frac{\log(l(t))}{\log(0.5)} \quad (9)$$

Using this equation, it is possible to directly compute the number of (integer) iterations of $MAD(x)$ required to reach age t without having to perform the loop described above. In the case of $EAD(x)$, $p(i)$ is not constant by definition, but it turns out that it follows a relatively regular pattern. In the case of Dutch life tables from 1850 to 2016, $p(1)$ varies between about 0.5 in the 1850s and 0.6 in the 2010s (reflecting the decreasing weight of premature mortality). After a few “lives” though, $p(i)$ seems to stabilize around 0.4 for all the mortality schedules (Figure 2). This value is below 0.5 because each section of the distribution of the ages at death (the life-table dx) above the modal age at death is right-skewed and thus $EAD(x)$ is higher than $MAD(x)$.

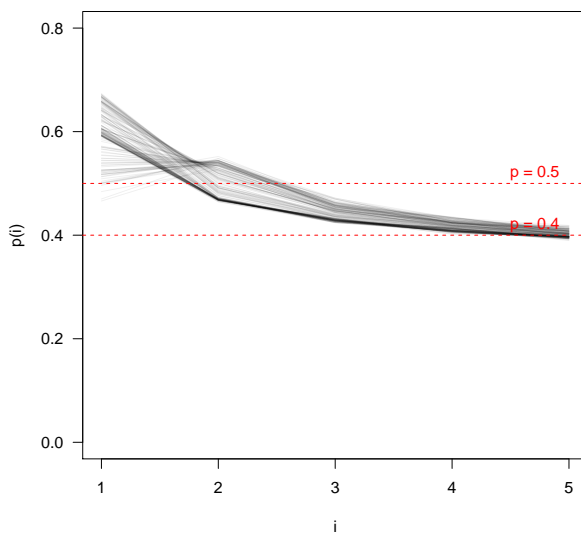


Figure 2: Probability to reach a new life (Netherlands 1850-2016)

Using this information, as well as additional analyses (see Appendix), and proceeding by analogy with the case of the *Median Age at Death*, we can postulate that a good approximation for $n(t)$ is given by the following formula

$$\hat{n}(t) = \frac{\ln(l(t))}{\ln(0.38)} + 1 \cong n(t) \quad (10)$$

where $l(t)$ is the value of the survival function at age t (i.e. the probability to reach age t starting from age 0). The addition of a unit intercept, reflects the fact that, when most people survive to t , $l(t)$ approaches 1 and $\ln(l(t))$ approaches 0, meaning that, on average people need only one life to survive to t . This formula also means that $n(t)$ practically only depends on mortality before age t , reflected in $l(t)$, but not on mortality after age t . This is not intuitive, because $n(t)$ relies on $EAD(x)$, and thus on $e(x)$, which depends on mortality at all ages above age x .

Simulations

In order to test the behavior of $n(t)$ and its approximation $\hat{n}(t)$, I designed a simple simulation using the Gompertz law ($\mu(x) = \alpha \cdot e^{\beta \cdot x}$). Figure 3 shows that, for a given value of $e(0)$, $n(100)$ is smaller when α is

high and β is low (i.e. high baseline and low rate of ageing) than when α is low and β is high (i.e. low baseline and high rate of ageing). For instance, for a life expectancy of 80, $n(100)$ is about 4 when $\alpha = 8e - 5$ and $\beta = 0.08$, but is about 8 when $\alpha = 1e - 5$ and $\beta = 0.12$. In other words, the number of times someone needs to survive its *EAD* is more sensitive to the rate of ageing than the baseline level of mortality. This is not surprising, since $e(x)$ decreases with age, and thus the majority of the “lives” take place at old age, where mortality is more affected by the rate of senescence than the baseline mortality level.

Figure 3 also confirms that the approximation $\hat{n}(t) \cong \frac{\ln(l(t))}{\ln(0.38)} + 1$ is good at all combinations of α and β (the red lines match closely the pattern of the black lines). At very low levels of mortality, when it takes only about 2 “lives” to reach 100 years of age, it slightly underestimates the true value of n^* , but these are extreme cases where the loop approach might be slightly biased. The discrepancy is altogether very small, and the patterns are almost identical. The results of this simulation provide thus additional support to the approximation of $n(t)$ by $\hat{n}(t)$.

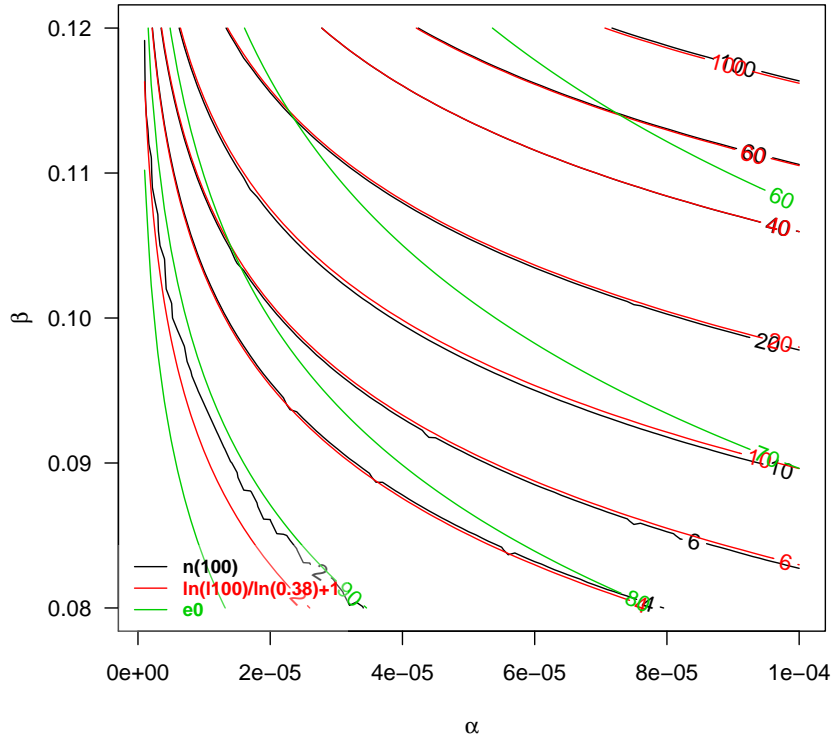


Figure 3: Relation between $n(100)$ and $e(0)$ across simulated Gompertz forces of mortality

Applications

Applications of this measure are manifold. I present here two: the evolution of $n(100)$ over time, and the decomposition of the difference in $n(100)$ by age.

Evolution of $n(100)$ over time

Most countries experienced a steady decrease of $n(100)$ over time. Exceptions to this regularity can be found in certain Eastern European countries, where the number of lives it takes to be a centenarian has increased

sharply between the 1950s and the 2000s, before decreasing again in the last decade (Figure 4). While these trends are sometimes correlated to those of $e(0)$, it is not systematically the case. For instance, during the early 1960s and late 1980s, life expectancy at birth increased sharply while it was getting more and more difficult to become a centenarian. This probably reflects the fact that these sharp gains in life expectancy were due to gains in premature mortality (for instance through a better control of alcohol consumption), while mortality at older ages tended to increase during the same time.

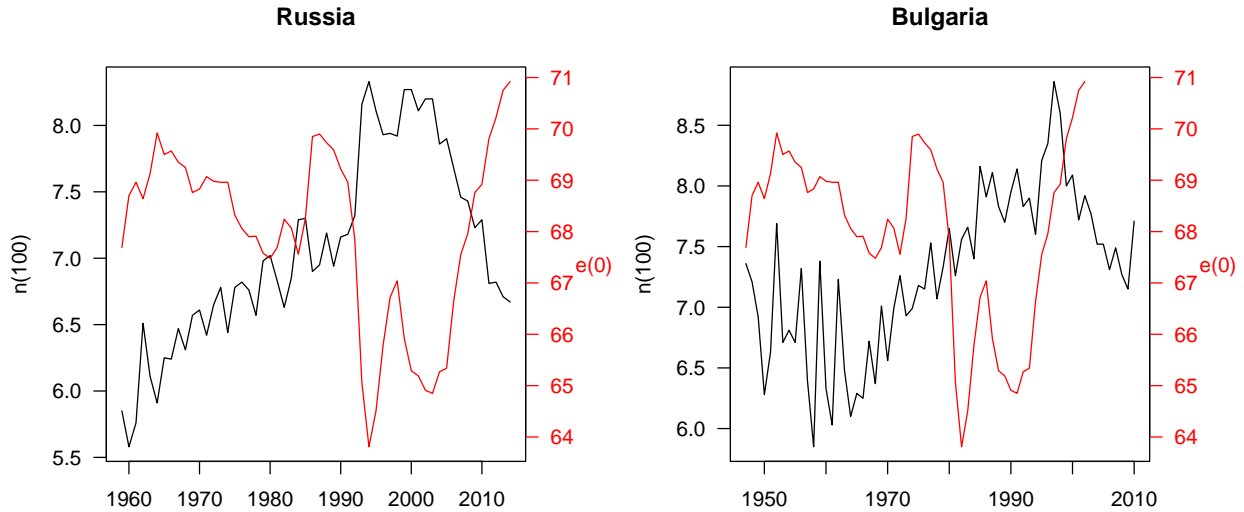


Figure 4: Evolution of $n(100)$ in Russia and Bulgaria since the 1950s

Another interesting pattern is found in some, but not all Western European countries, such as France and Scotland (Figure 5). In these two countries, $n(100)$ gradually increased during the nineteenth century despite a general improvement in $e(0)$. In France, a new born gained about 10 years of life expectancy between 1820 and 1900. In the meantime, he or she had to “cheat death” two more times in order to become a centenarian. This paradox is explained by a simultaneous decrease of premature mortality and increase of old-age mortality. In 1900, it was thus easier for this child to reach adulthood, but harder to become a centenarian. The Scottish case, while less dramatic, shows a similar evolution between the 1850s and 1920s. Other countries who display this paradox are Sweden between the 1750s and the 1850s, and, to a lesser extent, Denmark and Norway between the 1850s and 1900s, Australia between the 1920s and 1940s, Italy between the 1880s and 1940s, and Portugal between the 1940s and 1960s. It seems however absent in countries like Finland, Switzerland, and the Netherlands, which all have data going back at least to the 1880s. This paradox could be the signature of a “failure of success”, which might have increased mortality at old ages by saving frailer children.

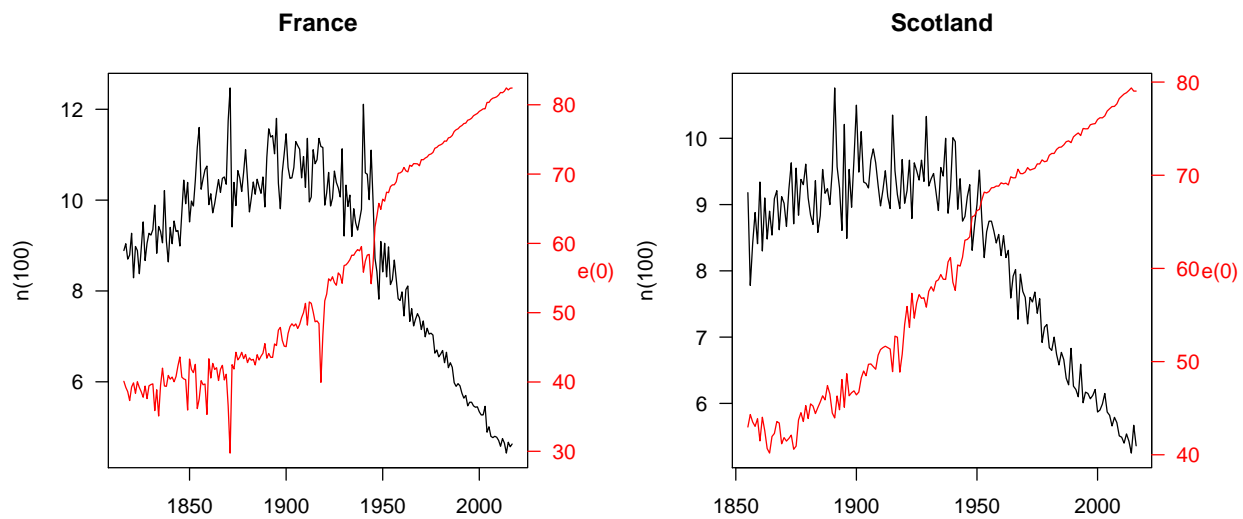


Figure 5: Evolution of $n(100)$ in France and Scotland since the 1800s

Decomposition

The measure $n(t)$ can be manipulated as any other summary measure of mortality. It can for instance be decomposed by age, and even by cause of death, provided that data on age- (and cause-)specific mortality are available. Numerical decomposition techniques (Andreev, Shkolnikov, and Begun (2002)) are perfectly suited for this task. For instance, in 2016 Dutch men had to “cheat death” 6 times (5.99) to become a centenarian, against 5 times (4.97) for women. In other words, today in the Netherlands, a man who reaches 100 years of age has ‘survived’ one more time than a woman of the same age. Figure 6 indicates that most of this additional effort took place after 80 years of age, and almost none of it before age 60

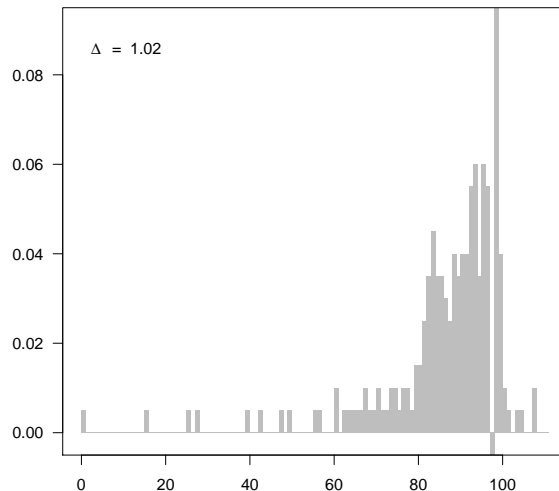


Figure 6: Decomposition by age of the difference in $n(100)$ between Dutch men and women

Conclusion

Humans are reaching old age ever more easily, which questions our approach and measure of ageing. I propose a new measure of mortality that sheds a new light on what it takes to reach old age. This measure is the number of times someone needs to exceed his or her predicted age at death in order to reach an arbitrary target age (for instance 100 years of age). It reflects at the population level, the dominant theories of individual ageing and is relatively easy to compute. Thanks to a surprisingly simple empirical relationship, it also has an alternative interpretation as a function of the probability to survive to a given age.

This measure of “how many lives” people need to become a centenarian provides a new perspective on the mortality in past and present populations. Notably, it suggests that in some European countries, but not all, overall gains in mortality were at some time balanced at older ages by increasing risks of death, leading to a paradoxical situation where it was becoming “easier” to reach adulthood, but “more difficult” to become a centenarian.

Future extensions of this work that I intend to pursue are to test whether the approximation of $\hat{n}(t)$ hold for both sexes, or if the parameters are different, the application of these techniques on cohort data (which necessitates to project the mortality of non-extinct cohorts), and more sophisticated simulations, for instance using a Siller model to observe the behavior of $n(t)$ in the presence of varying infant mortality.

References

- Andreev, Evgeny M, Vladimir M Shkolnikov, and Alexander Begun. 2002. “Algorithm for Decomposition of Differences Between Aggregate Demographic Measures and Its Application to Life Expectancies, Healthy Life Expectancies, Parity-Progression Ratios and Total Fertility Rates.” *Demographic Research* 7: 499–522.
- Finkelstein, MS. 2005. “Lifesaving Explains Mortality Decline with Time.” *Mathematical Biosciences* 196 (2). Elsevier: 187–97.
- HMD. 2019. “Human Mortality Database. University of California, Berkeley (Usa), and Max Planck Institute

for Demographic Research (Germany). Available at Www.mortality.org or Www.humanmortality.de.”

Isaacs, Rufus. 1950. “Iterates of Fractional Order.” *Canadian Journal of Mathematics* 2. Cambridge University Press: 409–16.

Strehler, Bernard L, and Albert S Mildvan. 1960. “General Theory of Mortality and Aging.” *Science* 132 (3418). JSTOR: 14–21.

Vaupel, James W, and Anatoli I Yashin. 1987. “Repeated Resuscitation: How Lifesaving Alters Life Tables.” *Demography* 24 (1). Springer: 123–35.

Appendix

Appendix 1: code for Figure 1

```
# load single-year life table for the Netherlands (both sexes) in 2015
library(HMDHFDplus) # package to access HMD data
load("C:/remund/Data/hmd_user.RData") # personal user and password to HMD
lt <- readHMDweb(CNTRY = "NLD", item = "bltper_1x1", username = us, password = pw)
lt1 <- lt[lt$Year == 2015,]

# age groups
x <- unique(lt$Age)

# function to interpolate EAD(x) for non-integer values of x
eadfun <- splinefun(x = x, y = lt1$ex + x, method = "monoH.FC")

# main loop
nmax <- 20 # maximum value of n
t <- 100 # target age
ead <- rep(NA, nmax) # empty vector to compute successive values of EAD(i)
ead[1] <- lt1$ex[1] # EAD(0) = e(0)
i <- 1 # counter for number of iterations
while(ead[i] < t & i < nmax){
  i <- i + 1
  ead[i] <- eadfun(ead[i-1]) # composition of EAD(x) with itself
}
ead <- na.omit(ead) # remove missing values (if n < nmax)
n <- length(ead) # n (number of iterations until EAD(x) > t)

# plot the loop
plot(x, eadfun(x), type = "l", xlab = "x", ylab = "EAD(x)", las = 1, ylim = range(x), xlim = c(0,max(x)),
points(c(0,ead[-n]), ead, pch = 16, col = 2)
lines(rep(c(0,ead[-n]), each = 2), rep(ead, each = 2) * rep(c(0,1),n), type = "b", col = 8)
axis(3, at = c(0,ead[-n]), labels = 1:n, col = 1, col.ticks = 2, col.axis = 2, cex.axis = 0.7)
mtext(side = 3, text = "i", col = 2, line = 2)
abline(h = t, col = 3)
mtext(side = 4, at = t, text = paste("t =",t), col = 3, las = 1, line = 0.4, cex = 0.7)
text(ead[n-1], ead[n], col = 2, labels = paste("n = ",n), pos = 4, cex = 0.7)

# find and plot non-integer value of n (n*)
tstar <- optim(par = 95, fn = function(x){abs(eadfun(x) - t)}, upper = t, lower = 0, method = "L-BFGS-B")
abline(v = tstar, col = 3)
mtext(text = paste("t* =",round(tstar,2)), side = 1, at = tstar, col = 3, cex = 0.7, adj = 0)
inv.ead.i <- splinefun(x = c(0,ead[-n]), y = 1:n) # inverse function of EAD(i)
axis(3, at = tstar, col = 3, labels = "")
nstar <- round(inv.ead.i(tstar),2) # fractional value of n
mtext(side = 3, at = tstar, text = paste("n* =",nstar), cex = 0.7, line = 0.5, col = 3, adj = 0)
```

Appendix 2: code for Figure 2

```
# load single-year life table for the Netherlands (both sexes) in 2015
library(HMDHFDplus) # package to access HMD data
load("C:/remund/Data/hmd_user.RData") # personal user and password to HMD
lt <- readHMDweb(CNTRY = "NLD", item = "bltper_1x1", username = us, password = pw)
lt1 <- lt[lt$Year == 2015,]

# age groups
x <- unique(lt$Age)

# function to interpolate EAD(x) for non-integer values of x
eadfun <- splinefun(x = x, y = lt1$ex + x, method = "monoH.FC")

# main loop
nmax <- 20 # maximum value of n
t <- 100 # target age
ead <- rep(NA, nmax) # empty vector to compute successive values of EAD(i)
ead[1] <- lt1$ex[1] # EAD(0) = e(0)
i <- 1 # counter for number of iterations
while(ead[i] < t & i < nmax){
  i <- i + 1
  ead[i] <- eadfun(ead[i-1]) # composition of EAD(x) with itself
}
ead <- na.omit(ead) # remove missing values (if n < nmax)
n <- length(ead) # n (number of iterations until EAD(x) > t)

# plot the loop
plot(x, eadfun(x), type = "l", xlab = "x", ylab = "EAD(x)", las = 1, ylim = range(x), xlim = c(0,max(x)),
points(c(0,ead[-n]), ead, pch = 16, col = 2)
lines(rep(c(0,ead[-n]), each = 2), rep(ead, each = 2) * rep(c(0,1),n), type = "b", col = 8)
axis(3, at = c(0,ead[-n]), labels = 1:n, col = 1, col.ticks = 2, col.axis = 2, cex.axis = 0.7)
mtext(side = 3, text = "i", col = 2, line = 2)
abline(h = t, col = 3)
mtext(side = 4, at = t, text = paste("t =",t), col = 3, las = 1, line = 0.4, cex = 0.7)
text(ead[n-1], ead[n], col = 2, labels = paste("n = ",n), pos = 4, cex = 0.7)

# find and plot non-integer value of n (n*)
tstar <- optim(par = 95, fn = function(x){abs(eadfun(x) - t)}, upper = t, lower = 0, method = "L-BFGS-B")
abline(v = tstar, col = 3)
mtext(text = paste("t* =",round(tstar,2)), side = 1, at = tstar, col = 3, cex = 0.7, adj = 0)
inv.ead.i <- splinefun(x = c(0,ead[-n]), y = 1:n) # inverse function of EAD(i)
axis(3, at = tstar, col = 3, labels = "")
nstar <- round(inv.ead.i(tstar),2) # fractional value of n
mtext(side = 3, at = tstar, text = paste("n* =",nstar), cex = 0.7, line = 0.5, col = 3, adj = 0)
```

Appendix 3: code for Figure 3

```
# load single-year life table for the Netherlands (both sexes) in 2015
library(HMDHFDplus) # package to access HMD data
load("C:/remund/Data/hmd_user.RData") # personal user and password to HMD
lt <- readHMDweb(CNTRY = "NLD", item = "bltper_1x1", username = us, password = pw)
lt1 <- lt[lt$Year == 2015,]

# age groups
x <- unique(lt$Age)

# function to interpolate EAD(x) for non-integer values of x
eadfun <- splinefun(x = x, y = lt1$ex + x, method = "monoH.FC")

# main loop
nmax <- 20 # maximum value of n
t <- 100 # target age
ead <- rep(NA, nmax) # empty vector to compute successive values of EAD(i)
ead[1] <- lt1$ex[1] # EAD(0) = e(0)
i <- 1 # counter for number of iterations
while(ead[i] < t & i < nmax){
  i <- i + 1
  ead[i] <- eadfun(ead[i-1]) # composition of EAD(x) with itself
}
ead <- na.omit(ead) # remove missing values (if n < nmax)
n <- length(ead) # n (number of iterations until EAD(x) > t)

# plot the loop
plot(x, eadfun(x), type = "l", xlab = "x", ylab = "EAD(x)", las = 1, ylim = range(x), xlim = c(0,max(x)),
points(c(0,ead[-n]), ead, pch = 16, col = 2)
lines(rep(c(0,ead[-n]), each = 2), rep(ead, each = 2) * rep(c(0,1),n), type = "b", col = 8)
axis(3, at = c(0,ead[-n]), labels = 1:n, col = 1, col.ticks = 2, col.axis = 2, cex.axis = 0.7)
mtext(side = 3, text = "i", col = 2, line = 2)
abline(h = t, col = 3)
mtext(side = 4, at = t, text = paste("t =",t), col = 3, las = 1, line = 0.4, cex = 0.7)
text(ead[n-1], ead[n], col = 2, labels = paste("n = ",n), pos = 4, cex = 0.7)

# find and plot non-integer value of n (n*)
tstar <- optim(par = 95, fn = function(x){abs(eadfun(x) - t)}, upper = t, lower = 0, method = "L-BFGS-B")
abline(v = tstar, col = 3)
mtext(text = paste("t* =",round(tstar,2)), side = 1, at = tstar, col = 3, cex = 0.7, adj = 0)
inv.ead.i <- splinefun(x = c(0,ead[-n]), y = 1:n) # inverse function of EAD(i)
axis(3, at = tstar, col = 3, labels = "")
nstar <- round(inv.ead.i(tstar),2) # fractional value of n
mtext(side = 3, at = tstar, text = paste("n* =",nstar), cex = 0.7, line = 0.5, col = 3, adj = 0)
```

Appendix 4: Additional analyses using the Human Mortality Database

I computed $n(t)$ for different values of t between 85 and 110 years of age, using almost all available life tables from the HMD (yielding 104,324 observations), and fitted a linear regression between $\ln(n(t))$ and $\ln(l(t))$. The results show that the best fit is obtained by fixing the intercept at 1, yielding $p = 0.38$. In this case, R^2 reaches 0.999. These results will receive more attention in the final version of the paper.

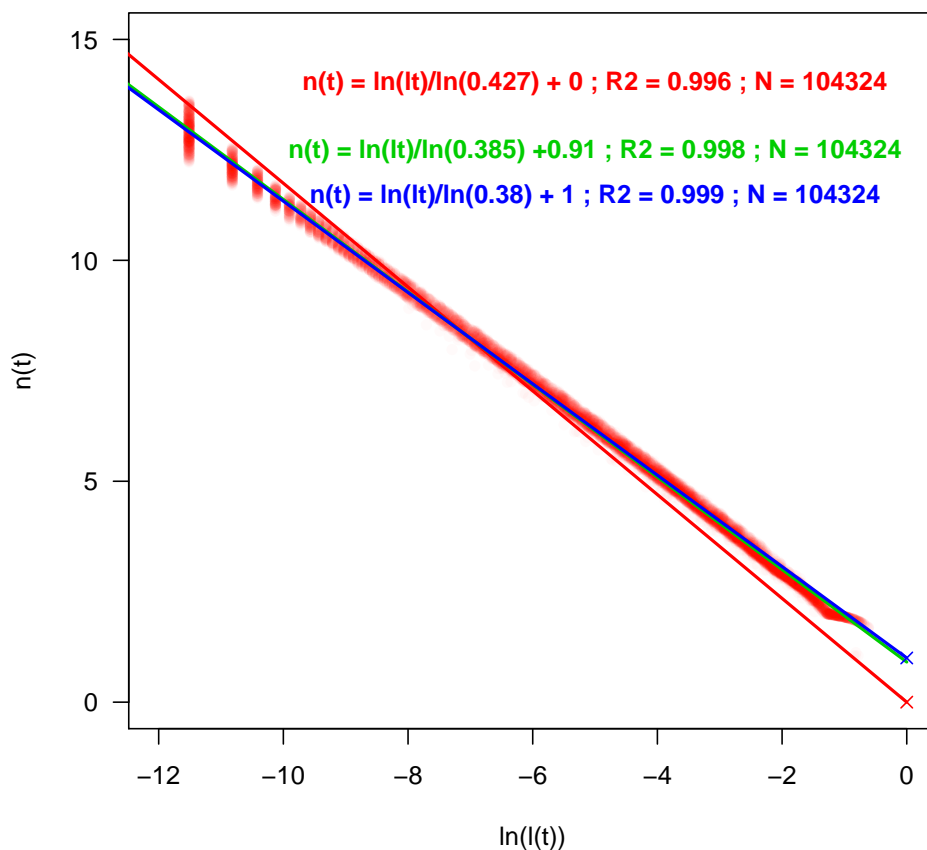


Figure 7: Empirical relationship between $n(t)$ and $l(t)$ in a large set of real populations