Telomeres and the Statistical Mechanics of Mortality

Charles Dermer

August 23, 2023

Abstract

Shortening of telomeres is implicated in aging and, ultimately, of cell senescence and cell death. A model is proposed wherein the statistical distribution in the times at which the telomeres of an ensemble of cells are destroyed and cell death occurs gives the model mortality rate for the organism. Assuming that the number of natal human telomeres is $N_T = 50N_{50}$ ($N_{50} \approx 1$), and that the rate of telomere loss is such that, on average, the cell reaches zero telomeres after the Biblical age $t_B = 70t_{70}$ years ($1 \leq t_{70} \leq 8/7$), the model reduces essentially to a 1-parameter model for the probability p of telomere loss per step.

Population data from the Population Division of the United Nations is used for a handful of nations in epoch 2010-2015 to calculate male and female mortality rates for different 5-year age cohorts. The telomere loss model better fits the current mortality rates with $t_B \cong 80$ years. The model is critiqued.

Intro

The Psalmist, probably David, recorded this sentiment some 3000 years ago:

> The days of our years *are* threescore and ten;
> and if by reason of strength *they* be fourscore,
> yet *is* their strength labour and sorrow;
> for it is soon cut off,
> and we fly away.
>
> (Psalms 90:10, King James Version)

The ancients recognized a transition to old age at 70 to 80 years, followed by the inevitability of death. Let's call this transition the Biblical age, and denote it by $t_B = 70\ t_{70}$ years, with $1 \leq t_{70} \leq 8/7$, in accordance with the Psalmist.

Guided by this ancient quantitative wisdom, an underlying mechanism is proposed that leads to an abrupt mortality rate increase to decay and death of the human organism at about the Biblical age. It is based on a classical statistical model applied to telomere shortening. Results of the model in its simplest form are examined.

Telomeres and the Hayflick Limit

Telomeres were discovered by the geneticist Hermann J. Müller in the UK in 1938. These endcaps of chromosomes contain no essential genetic material for reproduction, but function to protect the DNA that contains the genetic code. As cells age, the telomeres shorten. A human telomere consists of many kilobase pairs at birth, but may contain only ≈1500 in old age. A human cell's telomere loses ≅ 25 to 100 base pairs per year.

The Hayflick limit was named after anatomist Leonard Hayflick for work done at the Wistar Institute in Philadelphia in 1961. He showed that after a certain number N_T of cell divisions, a cell will cease dividing and enter a senescence phase. Hayflick found that the number of cell divisions for human fetal cells was $N_T = 50(\pm 10)$. This limit applies to somatic cells, not germ cells, cancer cells or prokaryotes.

This suggests that as somatic cells divide, collections of sequences of length ~50-100 base pairs are lost with each cell division. Each of these subdivisions is called a telomere here, recognizing that the entire endcap is the entity usually referred to as a telomere.

Telomere Loss Model

Telomere shortening is modeled as a one-dimensional random walk problem. The probability of losing a telomere with each step is denoted by $p \equiv 1 - q$, with $0 < p \leq 1$. Starting with $N_T \equiv 50N_{50}$ telomeres, the cell will, on average, lose all of its telomeres after N_T/p steps, during which time the system is hypothesized to reach the Biblical age $t_B = 70\ t_{70}$ years, with $1 \leq t_{70} \leq 8/7$. Thus the number of steps/yr for a cell to reach 0 telomeres is, on average, $(N_T/p) / (70t_{70}) = (5/7p)\ (N_{50}/t_{70})$ steps/year.

A Gaussian already provides an excellent approximation to the exact combinatorial results for a statistical distribution when the number of steps exceeds ≈10, let alone ≥ 50 N_{50}/p. The probability of surviving to time t(yr) in this model therefore is given by the area of the Gaussian that has not yet swept through the zero telomere endpoint.

Thus the survival probability to dimensionless step time τ is simply

$$P_S(\tau) = (2\pi pq\tau)^{-1/2} \int_0^{N_T} dx \exp\left[-\frac{(x-x_1)^2}{2pq\tau}\right], \qquad (1)$$

where $x_1 \equiv N_T - p\tau$, and x is a continuous variable replacing the discrete integer variable representing the number of steps to the right, or number of telomeres destroyed. Eq. (1) divides into two regimes: $\tau < N_T/p$, representing the young, early-time regime, and $\tau > N_T/p$, representing the old, late-time regime (one is tempted to say *l'ancien regime*). As constructed, the transition time between the two regimes is the Biblical age t_B.

Eq.(1) is equal to

$$P_S(\tau) = \frac{1}{2}(1 + \operatorname{sgn} y_1\ \operatorname{erf}|y_1|)\ , \quad y_1 \equiv \frac{N_T - p\tau}{\sqrt{2pq\tau}}, \quad (2)$$

where $\operatorname{sgn} y \equiv y/|y|$ is the sign function and $\operatorname{erf} y \equiv \frac{2}{\sqrt{\pi}} \int_0^y dx\ e^{-x^2}$ is the error function.

The loss rate per step, $\nu_S(\tau) = -\frac{\left[\frac{\partial P_S(\tau)}{\partial \tau}\right]}{P_S(\tau)}$, becomes

$$\nu_S(\tau) = \frac{(N_T + p\tau)\exp(-y_1^2)}{\tau\sqrt{2\pi pq\tau}\,(1+\operatorname{sgn} y_1\ \operatorname{erf}|y_1|)}\ , \quad (3)$$

upon substituting eq. (2).

The mortality rate ν_B at the Biblical time $\tau_B = N_T/p$ is $\nu_B = \sqrt{\frac{2}{\pi N_T(1-p)}}\ p$ step^{-1}, and the late-time asymptote is

$$\nu_S\left(\tau > \frac{N_T}{p}\right) \cong \frac{p}{2q}\ [\,1 - (\frac{p\tau}{N_T})^{-2}]/[1 - \left(\frac{1}{2y_1^2}\right)]\ . \quad (4)$$

The 0th, 1st, and 2nd-order terms are clearly seen in this expression.

Recalling the relationship between step time τ and physical time t(yr) = $(7pt_{70}/5N_{50}/)\tau$, the mortality rate in physical units is ν_M (t[yr]) [yr^{-1}] = $\nu_M(\tau(t))$ [step^{-1}]×(5/7p) (N_{50}/t_{70}) [steps/yr]. Thus the mortality rate at the Biblical time is $\nu_B \cong \frac{0.114}{t_{70}}\sqrt{\frac{N_{50}}{2(1-p)}}\ \ yr^{-1}$. The very late-time (0th-order) asymptote is simply $\nu_M(t \gg t_B) = \frac{5}{7}\frac{N_{50}}{[2(1-p)]\,t_{70}}\ \ yr^{-1}$.

Fig. 1 shows a calculation of the mortality rate in physical units for the standard model $N_{50} = t_{70} = 2p = 1$. The 0th, 1st, and 2nd–order asymptotes are shown by the short-dashed line at (5/7) yr^{-1}, the dotted, and the dashed curves, respectively. The solid curve gives the analytic result, eq.(3). The dot-dashed curve additionally includes an assumed underlying constant mortality rate ν_{min} = 0.008 yr^{-1}. The value of ν_B at the Biblical age of 70 years is pointed out.

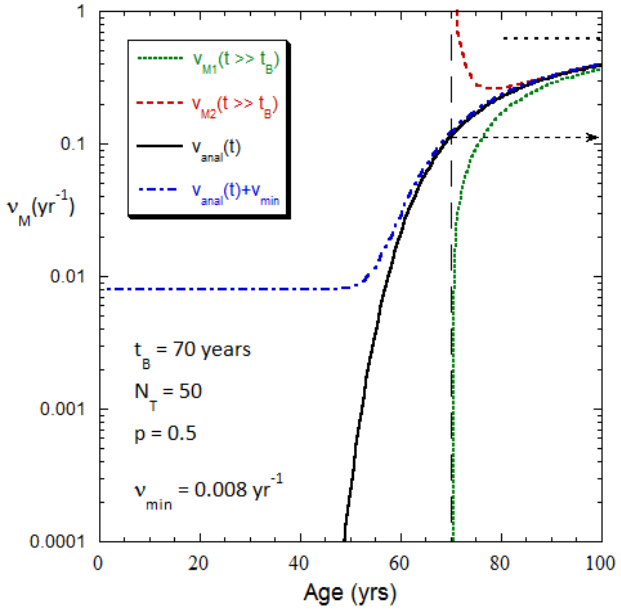


Figure 1. Telomere model for standard parameters, including asymptotes, and analytic results including an ad hoc additional mortality rate ν_{min} = 0.008 yr^{-1}.

Parameter surveys are shown in Fig. 2 for p = 0.1, 0.3,0.5, 0.7 and 0.9 (Fig. 2a), and N_T = 30, 40, 50, 60 and 80 (Fig. 2b), calculated for the Biblical age of 70 (dotted lines). The pivot point is insensitive to the studied range of parameters. The

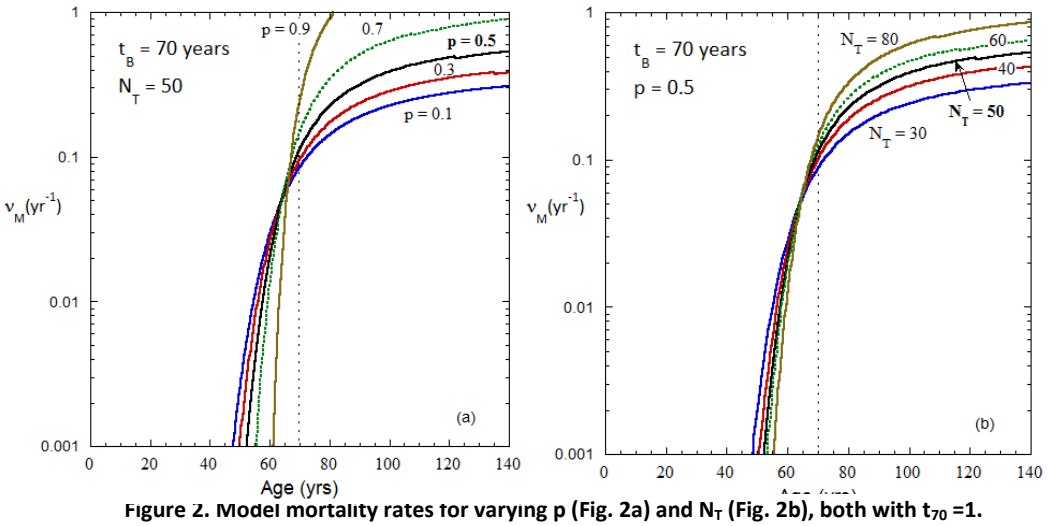


Figure 2. Model mortality rates for varying p (Fig. 2a) and N_T (Fig. 2b), both with t_{70} =1.

more rapid rise of the mortality rate with increasing p (Fig. 2a) reflects a narrower Gaussian sweeping though the zero telomere endpoint; indeed, as p approaches unity, its width vanishes. For p and t_B held constant, the mortality rate steepens with increasing N_T because the relative temporal width of the Gaussian describing telomere survival at a given time broadens with fewer telomeres. Specifically, the ratio of the square root of the dispersion σ to the Biblical time is $\frac{1}{5\sqrt{N_{50}}}\ \frac{\sqrt{2(1-p)}}{2p}$ in this model.

Country Population Data

Year-by-year population data from the United Nations, Department of Economic and Social Affairs, Population Division are found online at populationpyramid.net for many nations of the world over many years as far back as 1950. The population data give the population of male and female 5-year age cohorts beginning with the 0-4 year age cohorts, though the eldest cohort is 100+ years of age. Total population of the country for each year is also given.

We accept these data at face value, except for one adjustment described shortly. First, note that each cohort ages (of course) after 5 years into a 5-year older age cohort. Hence the mortality rate is simply $\nu_M(t)[yr^{-1}] = -\ln\left[\frac{N_{i+1}(t+5\ yr)}{N_i(t)}\right]$ / 5 years, where $N_i(t)$ is the number of women or men in age cohort i at time t(yr) (the biological gender index is suppressed). Arbitrarily choosing the 2010-2015 CE reference epoch (a time of relative world political stability and no global pandemic) for several countries gives the results shown in Figs. 3a, c, d, e and f for the USA, Iceland, Hungary, Japan, and China, respectively, Fig. 3b shows the results for the USA in the 2000-2005 epoch.

The logic for the choice of these countries is subjective. The USA is chosen for its intrinsic interest, and a second epoch is taken for comparison. The others are advanced nations with probably accurate records (though other nations could have excellent records, and first-world nations' population data may well be corrupted). Iceland, though with low statistics, should be able to track each of its citizens precisely. The same goes for a totalitarian society which can surveil its subjects with impunity. The high statistics and astonishing data from China may bear this conjecture out. Hungary and Japan were chosen for their low immigration rates, and as countries suffering from fertility collapse.

The extraordinary exponentially increasing power law above ~60 years of age is apparent in every nation chosen. For younger people, the mortality rate strongly varies from country to country, sometimes going negative when using raw population numbers (the open symbols). To crudely correct for unaccounted migration (the most likely reason for negative mortality), the population data are normalized to the total population of the country at that epoch (closed symbols). This is consistent with the supposition that most migration is undertaken by young adults, so the raw population data are most affected in the younger age cohorts. The raw and normalized data indeed give similar results above about 60 years of age. The rates inferred from the normalized data are also generally better behaved than the rates from the raw data at younger ages.

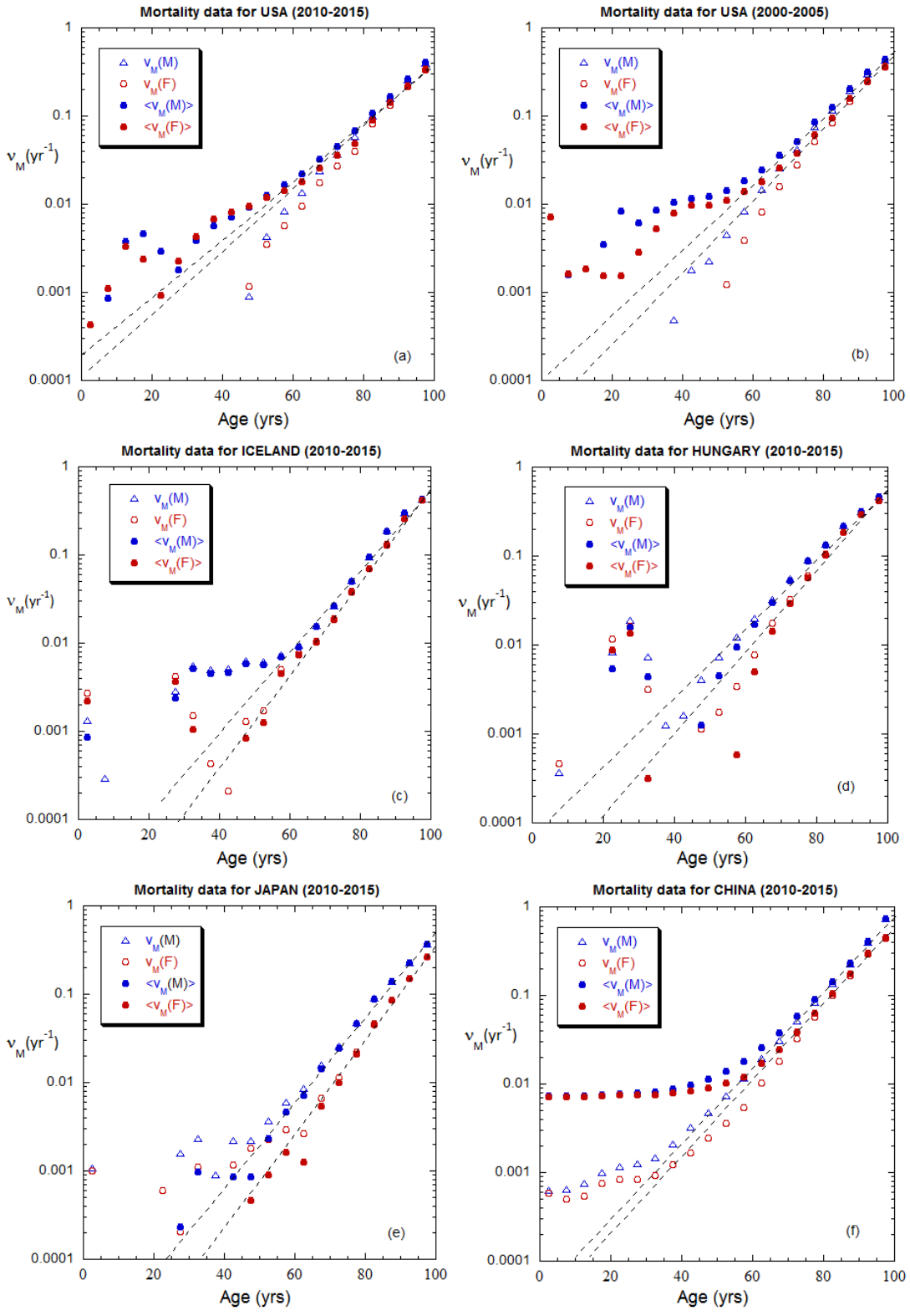


Figure 3. Male and female mortality rates derived from UN population (open symbols) and normalized population (closed symbols) data. The rates are calculated for the different countries at the epochs given in the labels. Also shown are best fit power laws by eye to the old-age male and female mortality rates for each country.

More remarkable still is the similarity in the magnitudes and slopes of the mortality rates for the elderly, irrespective of nation, suggesting a common (but not identical, as witness the comparatively low Japanese mortality) mechanism that governs death.

In all countries, the male mortality rate exceeds the female mortality rate within that country, not only for the elderly, but also generally for younger men and women. The lower mortality rate and consequent longer average lifespan for women compared to men is well known, but these data quantify the difference. A slight systematic affecting the rates of the most elderly is addressed in Appendix A.

Many questions arise concerning the Fig. 3 mortality rate data for persons younger than ≈60 years of age. For example, are the changes in rates between the two US epochs statistically significant?; why were the raw population rates for Iceland oftentimes negative for the young—was there significant migration into Iceland during this epoch?; what is the reason for the odd mortality rate spike in the 2010-2015 Hungarian data (for both males and females) between ≈25 and 30 years of age? What is in the Japanese lifestyle/diet, environment or genetics that is conducive to long life?; Why do the raw and normalized Chinese rates differ by a constant scale factor up to the age of about 40 (inconsistent with an explanation from migration)? Most important, what is the nature of the mortality of the young?

These questions are alas beyond the scope of this paper, and we resume the study of old-age mortality.

Telomere Model vs. Mortality Data

Comparing the mortality rate data with the model curves in Fig. 2, we see that the $t_{70} = 1$ model fails, but that a much improved fit is obtained by moving the Biblical age to 80 years. This is shown in Fig. 4, where mortality rates of the elderly are taken from the visual fits to the data shown in Fig. 3 for the countries and epochs considered. The mortality rates are plotted at 70, 80, and 100 years of age (see also App. B). The mortality data are compared with different values of the parameter p in Fig. 4a, with $N_T = 50$, and different values of the parameter N_T in Fig. 4b, with $p = ½$.

Quite obviously, even for the Biblical age of 80 years, the fit is not perfect, particularly for the rates of the most elderly plotted at 97.5 years of age (about the average age of the 90-94 year-old age cohort in 2010 that ages into the 95-99 year-old age cohort in 2015). Correcting for a systematic effect identified at the oldest age point (App. A) does not improve the fit; in fact it does slightly the opposite.

On the other hand, the model shown in Fig. 4 with a Biblical age of 80 years reproduces quite well the range of mortality data between 70 to 80 years. In this age range, the mortality rate data for the different countries differ by only a factor of ≈4 (≈2 without the Japanese data), and the model falls right on the midpoint of this range for the standard parameters $p = ½$, $N_T = 50$.

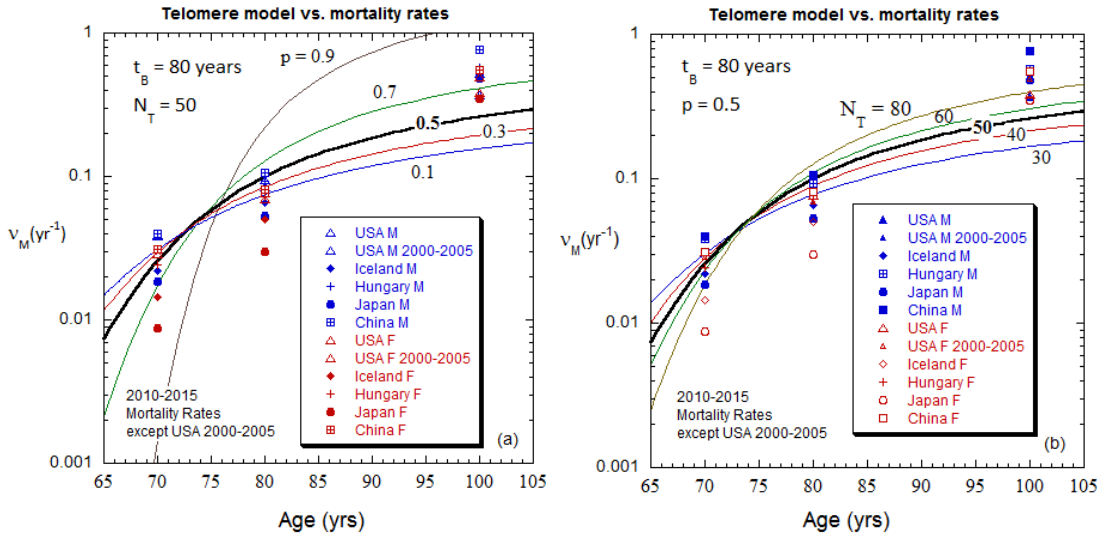


Figure 4. Model predictions for t_B = 80 years of age compared with male and female mortality rates for select countries. (a) Model dependence on p, with N_T = 50. (b) Model dependence on N_T, with p = 0.5.

Fig. 5 shows a comparison of the telomere loss model to the mortality rates derived from the US and Chinese population data for 2010-2015. A constant fixed mortality rate ν_{min} for all age groups is added to the mortality model. Other than ν_{min}, the model uses the same parameters for both countries.

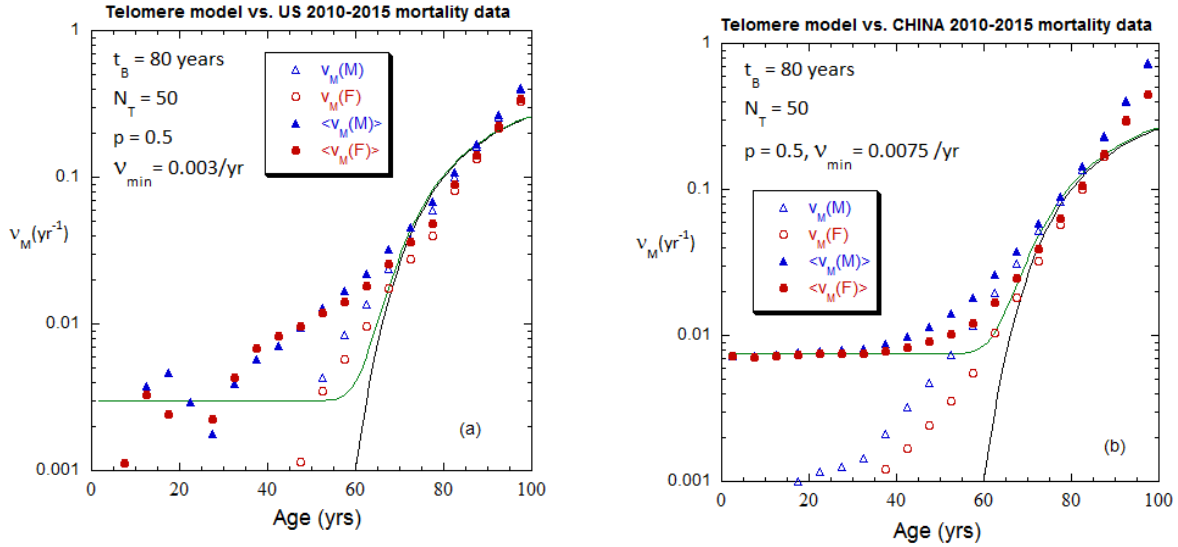


Figure 5. Comparison of telomere loss model with data from the US (Fig. 5a), and China (Fig. 5b).

Discussion

One may conclude from Figs. 4 and 5 that the model fails, as it neither reproduces the manifest exponential behavior of old-age mortality—the model rates are curved rather than linear on a log-linear plot of mortality rate with age---nor does it fit all the data, particularly for the most elderly. Or one can take a glass half-full approach and say that that the model is so easily falsifiable is its strength—death should not be so simply explained.

A more complex model might better fit, for example, by invoking a temporally evolving p with cell age. Or one could reverse-engineer the mortality data to try and find a process that reproduces the old-age exponential behavior. Absent a realistic physico-chemical mechanism, however, any such explanation would be unsatisfying. Pounding the data, one could suppose that there is additional old-age mortality, e.g., iatrogenic disease or euthanasia, not included in the model.

A conceptual hurdle of this model remains unresolved. The basic assumption is that the survival probability of the human organism mirrors the survival probability of an ensemble of cells as they age, their telomeres shorten, and they reach senescence and death. Central to this issue is the production of $\sim 2^{N_T}$ cells over the course of a human lifetime leading to the ~10-100 trillion human cells currently in a human body. This issue deserves further investigation, as do:

1. A mortality-rate catalog for many countries at many epochs (no doubt already performed by professional demographers);
2. Statistical and potential systematic errors in the rates derived from the population data.
3. Empirical fits to mortality data to calculate life expectancy for persons of various ages, assuming 2010-2015 mortality rates for that country (App. B).
4. Modeling temporally-evolving population profiles, with source and mortality rates also varying with time.
5. A numerical combinatorial model to assess the accuracy of the Gaussian approximation.

Only a full numerical study will put this last issue to rest, while also providing a more reliable and flexible (though far more computer-intensive) tool to treat population evolution (pt. 4).

Outro

Cell division and the problem of aging and death in the human organism is infinitely (cardinality **>>** $\aleph_0$) more complex than the toy model presented here. Yet in spite of, or perhaps because of its simplicity, an unambiguous and straightforward prediction can be made that is falsifiable, possibly even with current data. Given that it is well known that telomeres shorten with age, the prediction is that the sub-units referred to here as telomeres are distinct aggregations or agglomerations of base pairs that break off rather uniformly as the cell divides. The composition of these groupings involves some fraction ~$1/N_T$ of the number of natal base pairs in order to satisfy the Hayflick limit,

In contrast, if it were found or has been found that telomeres gradually erode with age absent cell division, or that telomere base-pair length is highly irregular in the course of cell division, the basis of this model would be undermined and the model, at least in its current form, should have to be abandoned.

The Psalmist affords us this final thought:

> …we spend our years as a tale *that is told*. (Psalms 90:9, KJV)

Presented August 23, 2023 with respect and admiration to Professor Péter Mészáros on the occasion commemorating his 80th birthday (as I approach my 70th). [v. 4,1/9/2024]

References

Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., *Essential Cell Biology* (3rd ed.), ©2010 (Garland Science, Taylor & Francis Group, New York)

PopulationPyramid.net: Population Pyramids of the World from 1950 to 2100

Reif, F., *Fundamentals of statistical and thermal physics*, ©1965 (McGraw-Hill: San Francisco)

Appendix A

An obvious correction to the mortality rates derived for those aging from the 95-99 yo age cohort into the 100+ yo age is the need to subtract the number of people living past 105 years of age from the 100+ yo age cohort. Examining the 2000, 2005, 2010, and 2010 epoch data for the USA, one finds that

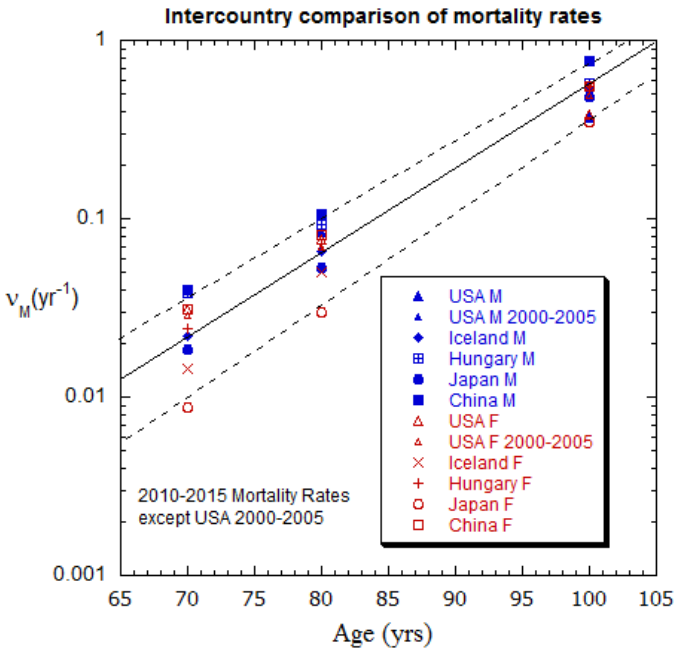
1. the fractions of the US population in the 90-95, 95-99, and 100+ yo age cohorts are in the ranges 0.37-0.46%, 0.09-0.14%, and 0.014-0.02%, respectively. In the US in 2010 specifically, 1 in every 6500 persons was older than 100 years of age.
2. The male/female ratio varies from ~30-40% for the 90-95 yo age cohort, decreasing to ~20-35% and 16-20% for the 95-99 and 100+ yo age cohorts, respectively.

Thus the probability of surviving from the 90-95 yo age cohort to the 96-99 yo age cohort is ~25-40%, and decreases by a factor of ~ 7 for survival from the 95-99 to the 100+ age cohorts. Given this trend, it is doubtful that any more than 10%, and probably a far smaller fraction of those who reach 100 years of age survive to the age of 105. Making a 10% reduction to the 100+ yo population has the effect of increasing the mortality rate by ~0.02 yr^{-1}, or a correction of no more than 5% upward for the 97.5 yo data points in Figs. 3 and 5.

Appendix B

Fig. 6 shows the mortality rates (see also Fig. 4) for men and women at 70, 80 and 100 years of age, taken from the visual fits to the data shown in Fig. 3 for the countries and epochs considered here.

As can be seen, there is considerable overlap between male and female mortality, though male mortality exceeds female mortality within each country considered. Assuming that these countries are representative of the mortality rates for the elderly, the straight-line visual fits spanning the range of these rates invite the following calculations to derive the average lifespan and



different countries, as noted, from fits to data in Fig. 3. Visual exponential fits spanning the range of data are shown by the solid and dashed lines.

expected remaining lifespan for members of a hypothetical population who have reached age t(yr). The premise of the calculation is that the mortality law does not significantly change over the lifespan of the individual. For a given country, this can be checked from available population data at different epochs.

As Fig. 6 shows, contemporary old-age male and female mortality are well described by an empirical mortality rate law of the form $\nu_M(t)\ [\,yr^{-1}] = \nu_{min} + \nu_o \exp(k_1 t)$, where $k_1 \equiv \ln(\nu_{t_2}/\nu_{t_1})\ /(\,t_2 - t_1)$, ν_{t_1} and ν_{t_2} are the mortality rates at t_1 and t_2 years of age taken from the Fig. 6 fits, and $\nu_0 = \nu_1 \exp(-k_1 t_1)$. The term ν_{min} is purely empirical. A power-law form for the mortality of the young would be more accurate, but adds another parameter while not being very relevant for old-age mortality.

The steady-state form of the master population equation in the absence of sources or sinks is $\frac{\partial N(t)}{\partial t} = -\nu_M(t)N(t)$ for $t \geq 0$, which is trivially solved to give the fraction

$$\frac{N(t)}{N(t_0)} = \exp\left[-\nu_{min}t - \left(\frac{\nu_0}{k_1}\right)\left(e^{k_1 t} - 1\right)\right]$$

of the original number $N(t_0)$ that survive to age t. $N(t_0) = 1$ gives the survival probability for a single model individual. The 50% probability to survive to age $t_{1/2}$ if one has already reached age t is obtained by numerically solving $N(t_{1/2})/N(t) = ½$.*

The average lifespan and the expected 50% survival time of individuals born and living in conditions governed by the range of mortality rates shown in Fig. 6 are calculated and shown in Fig. 7, with $\nu_{min} = 0.001$ yr^{-1} and $\nu_{min} = 0.01$ yr^{-1} for the "background" mortality rates. The expected lifespan at birth ranges from 76 to 87 years of age, with a median of ~80 years for the low background, $\nu_{min} = 0.001$ yr^{-1}. Life expectancy at birth is ~25% less when $\nu_{min} = 0.01$ yr^{-1}.

*A related calculation compares areas of the survival probability.

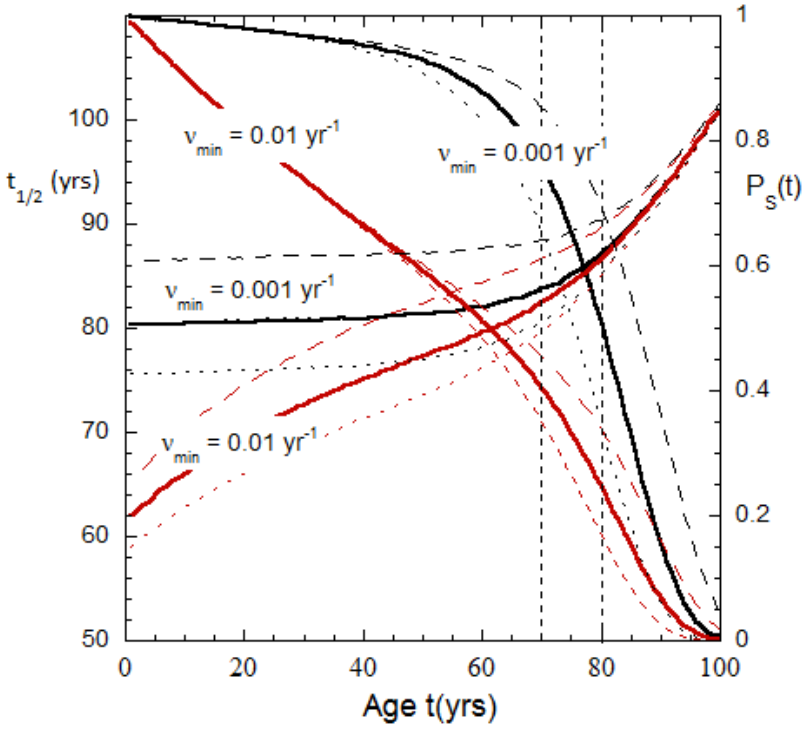


Figure 7. Probability $P_S(t)$ of survival to age t(yr) (downward curves, right axis) and the age $t_{1/2}(t)$ to which 50% of persons who have already reached age t(yr) survive (upward curves, left axis). The empirical fits in Fig. 6 for the range of mortality rates for the elderly give the range of values of $P_S(t)$ and $t_{1/2}(t)$, as shown. The background mortality rates are set equal to $\nu_{min} = 0.001$ yr^{-1} and $\nu_{min} = 0.01$ yr^{-1} as noted. The Biblical ages of 70 and 80 years are indicated by the vertical short-dashed lines.