Supporting Information of:

Telomere length behaves as biomarker of somatic redundancy rather than biological age

Jelle J. Boonekamp, Mirre J.P. Simons, Lia Hemerik, Simon Verhulst

Contents:

SI–I  Meta-analysis  p. 2
   I.A  Search and selection of studies  p. 2
   I.B  Data extraction and effect size calculations  p. 2
   I.C  Meta-analysis  p. 2
       Table S1  p. 3
   I.D  Meta-analysis results  p. 4

SI–II  Simulation study  p. 5
   II.A  General simulation procedures  p. 5
   II.B  Model optimization  p. 5
   II.C  Model comparison  p. 5

SI–III  Simulation model 1: biological age  p. 7
   III.A  Weibull  p. 7
   III.B  Gompertz  p. 7

SI–IV  Simulation model 2: somatic redundancy  p. 9

References  p. 11
SI–I Meta-analysis procedures

A. Search and selection of studies:
We searched papers with (i) ISI Web of Knowledge and Google Scholar using combinations of multiple keywords: human, telomeres, telomere length, age, ag(e)ing, survival, mortality, and (ii) by checking references of relevant papers. In addition, (iii) we checked all the papers that cited (Cawthon et al., 2003), the first paper showing an association of human telomere length with mortality. The last search was carried out on 2-Feb-2012.

From the retrieved papers we selected studies that contained human leucocyte telomere length (TL) measurements combined with a follow-up period in which mortality was recorded. Further inclusion criteria were: (i) the study used “healthy” subjects, i.e. studies in which subjects were not selected for carrying a particular disease or other health problem. Causes of death were unfortunately available in only a few cases, thus we could not take into account whether these were aging-related or not. We note however that since this increases measurement error, this makes our test more conservative, i.e. decreases type-I error probability. (ii) Whether the necessary data could either be extracted from the paper, or received after contacting the authors, which was the case for each otherwise eligible study. See Table S1 below for an overview of the studies and study-specific details on data extraction.

B. Data extraction and effect size calculations:
From each study we extracted: the natural logarithm (ln) of the hazard ratio and its 95% confidence interval associated with TL, the mean age of the study population at TL sampling, the length of the follow up period, and the TL assay method (qPCR, Southern Blot, or flow-FISH). Studies differed in the number of covariates included in the survival analysis, possibly rendering the TL estimates across studies to be incomparable. Therefore, we used only the simplest survival models reported, in which besides TL only age was taken into account.

Studies varied in whether they used TL as continuous variable or instead compared TL quantiles, which in principle renders the hazard ratio estimates to be incomparable, because the units of analysis differ (Kavvoura & Liberopoulos, 2007). We therefore determined for each study the unit of analysis and converted the HR’s accordingly (see Table S1 for details). For example, 1.23 in table S1 denotes that the HR was based on 1.23 kbp TL difference (if this was not reported in the paper we estimated it based on the reported mean TL and standard deviation, assuming a normal distribution). All analyses and figures were based on these converted HR values, but we note that this conversion had only minor effects on the results.

C. Meta-analysis:
We performed meta-analyses using the Metafor package (Viechtbauer, 2010) in R (R development core team, 2011) using a random-effects model fitted with restricted maximum likelihood. Sampling variances were calculated from the confidence intervals, and we used 1/s.e.^2 as weighting factor in the meta-analysis (Hedges & Olkin, 1985). Heterogeneity was evaluated using Q tests. With respect to testing whether the association of TL and mortality diminished with age we used the natural logarithm of age rather than age, because when the ln HR declines with age it can be expected that it will asymptotically approach zero, and this is better captured by ln age when compared to age.
### Table S1. Studies used in the meta-analysis. Sample sizes are the total numbers of individuals sampled. Ln HR (C.I.) denotes the natural logarithm of hazard ratio of TL with the 95% confidence interval in brackets. The letters a-c denote method of HR extraction: a=HR directly from paper, b=HR from author, c=HR calculated by us (see below for details). TL assay denotes whether TL was determined using quantitative PCR (q-pcr), Southern blot (s-blot), or flow-cytometry (flow-FISH). Mean age denotes the mean age in years at blood draw of the sampled subjects. Follow-up is the number of years after blood draw during which survival was recorded. Unit of analysis denotes the difference in TL in kbp that was used as unit of analysis in the study’s survival analysis (as determined by us). Corrected Ln HR (C.I.) denotes the study ln HR corrected for the unit of analysis, i.e. the study HR divided by the unit of analysis factor.

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample size</th>
<th>ln HR (C.I.)</th>
<th>TL assay</th>
<th>Mean age</th>
<th>Follow-up</th>
<th>Unit of analysis (kbp)</th>
<th>Corrected ln HR (C.I.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Bakaysa et al., 2007)</td>
<td>350</td>
<td>0.531 (0.182:0.956) a</td>
<td>s-blot</td>
<td>78.8</td>
<td>6.9</td>
<td>1.23</td>
<td>0.430 (0.148:0.774)</td>
</tr>
<tr>
<td>(Bischoff et al., 2006)</td>
<td>42</td>
<td>-0.101 (-0.375:0.189) c</td>
<td>s-blot</td>
<td>101.0</td>
<td>6.0</td>
<td>1.00</td>
<td>-0.101 (-0.375:0.189)</td>
</tr>
<tr>
<td>(Cawthon et al., 2003)</td>
<td>143</td>
<td>0.621 (0.199:1.040) a</td>
<td>q-pcr</td>
<td>71.9</td>
<td>15.0</td>
<td>1.55</td>
<td>0.402 (0.129:0.676)</td>
</tr>
<tr>
<td>(Epel et al., 2009)</td>
<td>235</td>
<td>0.329 (-0.067:0.725) b</td>
<td>q-pcr</td>
<td>73.7</td>
<td>12.0</td>
<td>1.55</td>
<td>0.213 (-0.044:0.471)</td>
</tr>
<tr>
<td>(Fitzpatrick et al., 2011)</td>
<td>1,136</td>
<td>0.278 (0.095:0.451) a</td>
<td>s-blot</td>
<td>73.9</td>
<td>8.1</td>
<td>1</td>
<td>0.278 (0.095:0.451)</td>
</tr>
<tr>
<td>(Harris et al., 2006)</td>
<td>190</td>
<td>0.092 (-0.147:0.330) b</td>
<td>q-pcr</td>
<td>79.0</td>
<td>5.0</td>
<td>1</td>
<td>0.092 (-0.147:0.330)</td>
</tr>
<tr>
<td>(Honig et al., 2006)</td>
<td>132</td>
<td>-0.223 (-1.204:0.588) a</td>
<td>q-pcr</td>
<td>81.4</td>
<td>NA</td>
<td>2.08</td>
<td>-0.107 (-0.578:0.282)</td>
</tr>
<tr>
<td>(Houben et al., 2011)</td>
<td>203</td>
<td>-0.215 (-0.673:0.248) a</td>
<td>q-pcr</td>
<td>78.5</td>
<td>7.0</td>
<td>0.9</td>
<td>-0.239 (-0.747:0.276)</td>
</tr>
<tr>
<td>(Kimura et al., 2008)</td>
<td>548</td>
<td>0.211 (-0.030:0.446) a</td>
<td>s-blot</td>
<td>78.8</td>
<td>7.3</td>
<td>1</td>
<td>0.211 (-0.030:0.446)</td>
</tr>
<tr>
<td>(Martin-Ruiz et al., 2005)</td>
<td>598</td>
<td>0.000 (-0.166:0.236) a</td>
<td>q-pcr</td>
<td>89.8</td>
<td>13.0</td>
<td>2</td>
<td>0.000 (-0.083:0.118)</td>
</tr>
<tr>
<td>(Martín-Ruiz et al., 2011)</td>
<td>751</td>
<td>0.365 (-0.198:0.936) b</td>
<td>q-pcr</td>
<td>85.0</td>
<td>1.5</td>
<td>1.45</td>
<td>0.252 (-0.137:0.646)</td>
</tr>
<tr>
<td>(Njajou et al., 2009)</td>
<td>2,721</td>
<td>0.000 (-0.105:0.095) a</td>
<td>q-pcr</td>
<td>73.6</td>
<td>10.0</td>
<td>1</td>
<td>0.000 (-0.105:0.095)</td>
</tr>
<tr>
<td>(Strandberg et al., 2011)</td>
<td>622</td>
<td>0.000 (-0.186:0.174) b</td>
<td>s-blot</td>
<td>75.6</td>
<td>7.0</td>
<td>1</td>
<td>0.000 (-0.186:0.174)</td>
</tr>
<tr>
<td>(Willeit et al., 2010)</td>
<td>787</td>
<td>0.467 (0.010:0.763) a</td>
<td>q-pcr</td>
<td>62.6</td>
<td>10.0</td>
<td>1.07</td>
<td>0.435 (0.009:0.626)</td>
</tr>
<tr>
<td>(Woo et al., 2008)</td>
<td>2,006</td>
<td>0.166 (-0.493:0.878) a</td>
<td>q-pcr</td>
<td>72.4</td>
<td>4.0</td>
<td>4.9</td>
<td>0.034 (-0.099:0.176)</td>
</tr>
<tr>
<td>(Zekry et al., 2012)</td>
<td>444</td>
<td>0.058 (-0.261:0.365) a</td>
<td>flow-FISH</td>
<td>85.3</td>
<td>5.0</td>
<td>2.08</td>
<td>0.028 (-0.125:0.175)</td>
</tr>
</tbody>
</table>

1 The cohort of Danish twins analyzed by Bischoff et al. was studied again later by Kimura et al. However, the paper by Bischoff et al included a separate analysis for centenarians, which was not replicated and we used only this study effect size in our meta-analysis. Survival and TL of centenarians were taken from figure 1 in the paper of Bischoff et al. HR was determined using a Cox-proportional hazard model without censoring of the data (all individuals died).

2 The total sample of individuals analyzed by Honig et al. was selected for the prevalence of Alzheimer’s disease, which led us to only include the HR of the control group in our meta-analysis.
D. Meta-analysis results:
We tested for publication bias using a funnel plot in combination with a rank test (Viechtbauer, 2010), and no publication bias was detected (Fig. S1 below; Kendall's tau = 0.150; P = 0.450). There was significant heterogeneity among effect sizes (Q = 33.1; P = 0.005). Residual heterogeneity was substantially reduced when adding subject sampling age to the model, but remained significant (–18%, Q = 27.2; P = 0.018), suggesting that in addition to subject sampling age, differences in study methods and, or, population differences may affect the association of TL and mortality. We tested for such study differences, i.e. TL assay method and study follow-up period, but these were not significant as main effect (TL assay method P = 0.502, follow-up period P = 0.767), or interacting with age (P = 0.678 and P = 0.148 respectively).

Figure S1. Funnel plot of the studies in the meta-analysis on the association of TL and mortality
SI–II Simulation study procedures

A. General simulation procedures:
With the simulation models of biological age and somatic redundancy (described in SI–III and –IV respectively), we simulated survival times per individual per study, using the number of individuals, mean subject sampling age, and follow-up period as in the studies used in the meta-analysis. In the simulation, we generated individual survival data from one age to the next by using the age and TL specific mortality probability (determined by either one of the model equations 1, 2, or 3 described in SI–III and –IV) and a random value drawn from the uniform distribution $U(0,1)$. Each study was simulated 50 times and we calculated the HR of TL using Cox’s proportional hazards with right censoring (Kleinbaum & Klein, 2005) per simulation cycle, and subsequently we averaged these HR’s over the 50 simulations. Thus, we obtained a simulated data set for each parameter combination for each of the models. We then optimized the parameters to maximize the resemblance between the simulated data and the meta-regression line of the real data, and subsequently compared which of the models generated data that best matched the observed pattern.

B. Model optimization:
To enable a quantitative comparison with the meta-regression results we optimized the model parameters for the simulated HR values to yield the closest possible fit to the meta-regression line of the observed studies. This was achieved by minimizing the sum of the weighted squared differences between the simulated study HR’s and the meta-regression line fitted through the observed HR’s. The weight factor that we applied to these squared differences was the same weight factor as used in the meta-analysis of the corresponding empirical studies, i.e. $1/s.e^2$. To find the optimal parameter values we started with a wide range of parameter combinations and applied bisectioning to find the optimal parameter values.

Theoretically, a good fit of the simulated data to the meta-regression line of the observed studies could be based on lifespan distributions in the simulated data that strongly deviate from the empirically observed lifespan, which would render the model uninformative. We avoided this problem by additionally fitting the simulated lifespan distributions to the observed lifespan distribution obtained from the Dutch bureau of statistics (Dutch bureau of statistics, 1996-2009) and omitted all model parameter combinations that yielded a fit of $r < 0.90$. We were limited to this selection of models, because a quantitative approach, i.e. directly optimizing the simulation model to the observed lifespan, requires the lifespan data of the studies that we used in our meta-analysis, and these are unavailable. Since all studies were done in recent years, and in Western countries, we consider it is safe to assume that these distributions are sufficiently similar when compared to our selection criterion. We calculated $r$ as follows

$$r = 1 - \frac{SSe}{SSot}$$

where $SSe$ is the sum of the squared differences between the observed and simulated probability density lifespan distributions, and $SSot$ is the sum of squared differences between the observed probability density lifespan distribution and its mean. For the calculations of $r$ we used matched age ranges of the simulated- and the observed lifespan distributions, and thus observed age at deaths of age $< 63$ and simulated age at deaths of age $> 98$ were ignored.

C. Model comparison:
To formally compare the fit of the simulation models to the observed pattern we performed
additional meta-regression analyses of the observed hazard ratios, pooled with the hazard ratios generated by one of the simulation models with the optimized parameters. Pooling data and then fitting one meta-regression is informative, because when the simulated data fit the observed data less well this results in a poorer total fit. As measure of goodness of fit we used Akaike’s “An Information Criterion” (AIC) (Akaike, 1974), calculated on the basis of the maximum log-likelihood (Metafor package in R). Following general convention, we considered models to fit equally well if their AIC’s differed by less than two (Burnham & Anderson, 2002).
SI–III Simulation model 1: biological age

A. Weibull:
We here describe how in our model TL determines biological age and how this was implemented in the Weibull distribution. At the start of each simulation, for each study, TL was generated from a normal distribution with the mean and SD approximating the mean TL and SD of the actual studies (TL mean=6.6 kbp, SD=1.0 kbp). TL shortening was included of 40 base pairs per year, which approximates the measured TL shortening rate in some longitudinal studies, e.g. (Aviv et al., 2009; Chen et al., 2011; Ehrlenbach et al., 2009; Houben et al., 2011). We stress however that the exact value has no effect on the outcome of the simulations, because the entire distribution shifted to shorter TL with increasing age, but the exact same range and relative differences between means were maintained. As measure of TL we used the age-specific deviation from the mean TL as follows:

\[ \delta T_t = T_t - \bar{T}_t \]

where \( T_t \) is TL (kbp) at age \( t \), and \( \bar{T}_t \) is the population mean TL at age \( t \). Subsequently we defined biological age as

\[ t' = t - b_1 \delta T_t \]

where \( t \) is age in years and \( b_1 > 0 \) is the parameter indicating how many years the age is adjusted per \( \delta T_t \). This would generate negative biological ages early in life but not in our simulations in which the lowest age is 63 years. At young age (after birth) we assume the effect of TL on biological age to increase non-linearly with age, levelling-off at medium to older ages, but note that we cannot test this because data on ages < 63 are unavailable.

We based equation (1) below on the Weibull distribution, which has been shown to describe the distribution of human life span well (Weibull, 1951), but for comparison repeated the analysis using the Gompertz distribution (see below). We assumed the hazard rate \( h(t) \) to increase with biological age \( t' \) as follows:

\[
(1) \quad h(t) = \lambda p(\lambda t')^{p-1}
\]

where \( \lambda \) and \( p \) are the Weibull scale and shape parameters respectively. In this model the effect of TL on mortality diminishes with age, because for \( p > 1 \) mortality increases as a power function of age while the modulating effect of TL on mortality does not. This results in TL becoming relatively less important for survival, because the mortality risk of other factors increases with age, suggesting qualitative agreement between the biological age model and the observed pattern. The parameter range that we tested for this model was \( [\lambda (10 \cdot 10^{-3}, 17 \cdot 10^{-3}); p (1, 10); b_1 (1, 10)] \). The optimal parameter values were: \( \lambda = 14.29 \cdot 10^{-3}; p = 4.0; b_1 = 3.0 \), resulting in a fit to the meta-regression line with AIC = -46.8 (calculated as described in SI–II.D); see Fig S2 below.

B. Gompertz:
Alternatively we based our model of biological age on the Gompertz function, because some discrepancy between these functions exists when fitting to old ages (Juckett, 1993). We used the same definition of biological age as previously described, and in the Gompertz model the hazard rate \( h(t) \) increases with biological age \( t' \) as follows:
where $R$ is the initial mortality rate and $a$ is the age dependent mortality. The parameter range that we tested for this model was $[R(1\cdot10^{-4}, 1\cdot10^{-3}); a(0.01, 0.2); b(1-15)]$. The optimal parameter values were: $R = 5\cdot10^{-4}; a = 0.045; b = 10.1$, resulting in a fit to the meta-regression line with AIC = -45.3 (calculated as described in SI–II.D); see Fig. S2 below).

Fig. S2 HR of TL according to the biological age model based on the Weibull and Gompertz distributions. The data points are the study HR values and the solid line is the meta-regression line as in Figure 1 of the main paper.
SI–IV Simulation model 2: somatic redundancy

The initial number of redundancy elements and the rate at which these fail characterizes a redundancy system. We considered TL as index of the number of redundancy elements, and we assumed the redundancy element failure rate to be constant, i.e. independent of age. This results in that the cumulative survivorship of a single redundancy element with failure rate $k$ decreases with age exponentially ($S(t)=e^{-kt}$) and the hazard function of an organism with multiple redundancy elements is therefore given by:

\[ h(t) = \frac{nke^{-kt}(1-e^{-kt})^{n-1}}{1-(1-e^{-kt})^n} \]

where

\[ n \equiv a + b_2\delta T \]

where $t$ is age in years, $k$ is the constant (age-independent) failure rate of $n$ redundancy elements, and $c$ is a scaling factor (Gavrilov & Gavrilova, 2001). As measure of TL we used the deviation from the population mean TL (kbp) at sampling age ($\delta T \equiv T - \bar{T}$). We set $a$ to 500, meaning that the average redundancy (at mean TL) at the start of our simulation was 500, and $b_2 > 0$ determines the redundancy per unit $\delta T$. In this model the effect of TL on mortality diminishes with age because the variation in the number of redundancy elements between individuals diminishes with age, because individuals with a high level of redundancy also lose more elements per unit of time, compared to individuals with a low redundancy level. We optimized the parameters of equation (3) using the same procedure as used for the previous model (see SI II.C). The parameter range that we tested was $[k \ (0.18, 0.25); \ c \ (0.25, 0.35); \ b_2 \ (15, 110)]$. The optimal parameter values were: $k = 0.235; \ c = 0.328; \ b_2 = 90$, resulting in a fit to the meta-regression line with AIC = -50.8 (see SI–II.D for details on calculations of the AIC).

The fit of the redundancy model in Fig. 1 is a meta-regression fit using the simulated data, which explains why the line goes below zero at ages > 92, instead of approaching zero asymptotically. For the exact outcome of the model see Fig. S3 below.
Figure S3. Exact ln HR of TL as calculated by the redundancy model (solid line). The data points are the study HR values as in Figure 1. The line asymptotically approaches zero because the rate at which total redundancy diminishes asymptotically approaches the redundancy element failure rate. This results in that at very old age all individuals face the same mortality risk, because mortality risk is then equal to redundancy element failure rate.
References


