Telomere length analysis from minimally-invasively collected samples: methods development and meta-analysis of the validity of different sampling techniques

Abbreviated Title: Telomere measurement from minimally invasive samples

Keywords: senescence, dried blood spots, saliva, buccal cells, aging

Peter H. Rej¹, Madison H. Bondy², Jue Lin³, Aric A. Prather⁴, Brandon A. Kohrt⁵, Carol M. Worthman², Dan T.A. Eisenberg^{1,6}

¹Department of Anthropology, University of Washington, Seattle, WA, 98105

²Department of Anthropology, Emory University, Atlanta, GA, 30322

³Department of Biochemistry and Biophysics, University of California, San Francisco, CA, 94143

⁴Department of Psychiatry, University of California, San Francisco, CA, 94143

⁵Department of Psychiatry and Behavioral Services, George Washington University, Washington DC, 20037

⁶Center for Studies in Demography and Ecology, Seattle, WA, 98195

September 24, 2019 Draft

Corresponding Author: Peter H. Rej, Ph.D., Department of Anthropology, University of Washington, 314 Denny Hall, Box 353100, Seattle, WA 98195-3100, USA Tel: (206) 543-5240, Email: peterrej@uw.edu

Abstract

Objectives: Telomeres are the protective caps of chromosomes. They shorten with cell replication, age, and possibly environmental stimuli (e.g., infection and stress). Short telomere length (TL) predicts subsequent worse health. Although venous whole blood (VWB) is most commonly used for TL measurement, other, more minimally-invasive, sampling techniques are becoming increasingly common due to their field-friendliness, allowing for feasible measurement in low-resource contexts. We conducted validation work for measuring TL in dried blood spots (DBS) and incorporated our results into a meta-analysis evaluating minimally-invasive sampling techniques to measure TL.

Methods: We isolated DNA extracts from DBS using a modified extraction protocol and tested how they endured different shipping conditions and long-term cryostorage. We then included our in-house DBS TL validation statistics (correlation values with VWB TL and age) in a series of meta-analyses of results from 24 other studies that published similar associations for values between TL measured in DBS, saliva, and buccal cells (BC).

Results: Our modified DBS extraction technique produced DNA yields that were roughly twice as large as previously recorded. Partially extracted DBS DNA was stable for seven days at room temperature, and still provided reliable TL measurements, as determined by external validation statistics. In our meta-analysis, DBS TL had the highest external validity (i.e., strongest overall association with both VWB TL and age), followed by saliva, and then BC.

Conclusions: Measuring TL using DNA isolated from DBS is the best option for minimally-invasively collected samples, which can be used to expand research to diverse settings and populations.

Introduction

Telomeres are the protective caps of chromosomes. Vertebrate telomeres are comprised of several kilobases of six base-pair repeated DNA sequence (TTAGGG), which function to protect the chromosomal ends from nucleotide loss due to incomplete end replication (Wright, Tesmer, Huffman, Levene, & Shay, 1997). The loss of telomeric DNA occurs with every cell division because there is not enough DNA substrate for the cell's replication machinery to bind at the end of the chromosome (Lingner, Cooper, & Cech, 1995). Thus, while telomere length (TL) is maintained by the reverse transcriptase, telomerase (Greider & Blackburn, 1987), telomeres still shorten over the lifespan, and TL is considered an important biomarker of aging. Furthermore, a cell whose TL reaches a critically short length can no longer replicate itself and enters into senescence (Armanios & Blackburn, 2012; Cong, Wright, & Shay, 2002; Palm & de Lange, 2008). Accordingly, short blood TL predicts an array of adverse health outcomes ranging from common killers, such as cardiovascular disease (CVD), to a general increase in all-cause mortality (Cawthon, Smith, O'Brien, Sivatchenko, & Kerber, 2003; Rode, Nordestgaard, & Bojesen, 2015; Serrano & Andrés, 2004; Wang, Zhan, Pedersen, Fang, & Hägg, 2018). In addition to cell division and age, environmental factors may also modify TL over the lifespan; most notably experiences of psychosocial stress (Epel & Prather, 2018; Quinlan et al., 2014). This plasticity makes telomeres of particular interest to social scientists.

Currently, our understanding of the relationship among telomeres, health, and exposures, including infections and social stress, are predominantly based on

populations from high-income countries (HIC). While there are exceptions (e.g., Rej, Tennyson, Lee, & Eisenberg, 2019; Tennyson et al., 2018; Zalli et al., 2014), a recent systematic review of stress and TL studies did not include any conducted in low- and middle-income countries (LMIC) (Mathur et al., 2016). Because the majority of the world's population lives in LMIC, and these settings differ from HIC with regard to exposure to infectious diseases and social stressors, especially traumatic stressors such as war and disasters, improved methods need development in order to conduct human TL research in these countries.

One way to increase the number of TL studies in LMIC is to improve field friendliness of sample collection. Venous whole blood (VWB) is the standard source of DNA for assessment of human TL. However, VWB collection in low-resource settings has inherent risks and challenges. First, venipuncture is an invasive procedure that requires a trained phlebotomist. Second, once collected, VWB samples must be frozen immediately or processed promptly to maintain DNA integrity. These infrastructural and logistical requirements are difficult to satisfy when working in settings without extensive medical and laboratory facilities. The few TL studies that have been conducted in LMIC have been mostly limited to hospitals, universities, or other high-resource settings (e.g., Auld et al., 2016; Chatterjee et al., 2015; Eisenberg, Lee, Rej, Hayes, & Kuzawa, 2019; Farrukh, Baig, Hussain, Shahid, & Khan, 2019; Farrukh, Baig, Hussain, & Lucky, 2019; Jiménez, Pereira-Morales, Adan, & Forero, 2019; Malan, Hemmings, Kidd, Martin, & Seedat, 2011; Oliveira et al., 2018; Pathai et al., 2013; Rej et al., 2019; Tennyson et al., 2018; Von Känel, Malan, Hamer, & Malan, 2015; but see Zahran et al., 2015).

Motivated by these demands and the overall expense of VWB collection in the low-resource settings, many researchers now assess TL using biological samples from minimally-invasive sources. These include saliva, buccal swabs, and dried blood spots (DBS) collected via capillary finger prick. Because researchers can collect each with minimal participant discomfort and without specialized training, all three serve as attractive alternative sources of DNA to VWB. Nevertheless, there are additional costs and benefits to each. While bulky and expensive, whole saliva collected using Oragene's DNA kit (DNA Genotek, Ottawa, Ontario, Canada) yields samples that are putatively stable for five years at room temperature (RT), and indefinitely at -20°C. For conventional DNA collection, buccal swab samples need to be processed within 48 hours of collection and should be maintained at 4°C at all times. Alternatively, they can be transferred to a cryovial and stored at -80°C (Woo & Lu, 2019). Like with VWB, if not processed soon after collection, it is difficult to obtain high-quality DNA from buccal cells (BC) (Nedel, André, Oliveira, Tarquinio, & Demarco, 2009), which will likely result in inaccurate TL measurement. DBS, on the other hand, have started to show promise for TL research in low-resource settings as demonstrated by work with rural populations in northern India (Zahran et al., 2015). DBS collection cards take up considerably less space than VWB collection tubes or saliva collection kits and most partially stabilize DNA, which makes DBS more robust than whole blood (McDade, Williams, & Snodgrass, 2007). This durability provides easier sample collection, transportation, and storage. That notwithstanding, investigators still report problems isolating DNA of adequate quality and quantity for use in the TL assay from blood collected on Whatman FTA and 903 protein saver cards (e.g., Goldman et al., 2017).

While DBS, saliva collection kits, and buccal swabs are all well-validated means of collecting DNA in less invasive ways than via VWB, TL measured using these minimally-invasive samples should not be assumed to be adequate substitutes for VWB TL measures without validation. There are technical/analytical as well as biological differences among these sample types. Technically, performance of extraction techniques differs for each modality, which affects DNA quality/quantity and subsequently influences downstream TL measures (Cunningham et al., 2013; Dagnall et al., 2017; Denham, Marques, & Charchar, 2014; Raschenberger et al., 2016). Biologically, TL varies across cell types due to different replicative histories and environmental exposures of different cell lineages. DBS, saliva, and BC samples are each composed of different cell types. One small study of bone marrow transplant patients suggests that BC composition is highly variable, primarily consisting of epithelial cells (37-95%, median 79%; Thiede, Prange-Krex, Freiberg-Richter, Bornhauser, & Ehninger, 2000). To the extent that factors affecting telomeres in immune cells differ from those acting on other tissues (e.g., by infections), BC may provide a poor proxy of leukocyte TL. Saliva appears to be even more heterogeneous than BC (5-84% epithelial cells, median 26%; Thiede et al., 2000). This heterogeneity could be influenced by confounding variables (e.g., someone with an oral lesion or infection might have more leukocytes in their saliva; Aps, Van Den Maagdenberg, Delanghe, & Martens, 2002; Theda et al., 2018). While, in theory, DBS samples should have the same cell composition as VWB, experimental analyses show significantly larger leukocytes (e.g., granulocytes) and greater leukocyte-subtype sampling error in DBS compared to VWB, as well as reduced precision with subtype representation in

repeated complete blood count measures (Bond & Richards-Kortum, 2015; Yang et al., 2001). Different leukocytes possess different mean TLs (Lin et al., 2016). Thus, even if an infection shortens epithelial cell and leukocyte telomeres equally, the varying means would add noise to TL measurements.

In this study, we conducted analyses to better validate DBS TL measures that could be adapted for use in low-resource settings. To do this, we assessed TL in two sets of matched DBS and VWB specimens. The first came from a sample of 19 individuals from the University of Washington (UW), Seattle, and the second set consisted of 15 individuals from Emory University. These two samples allowed us to test how different conditions, likely to be encountered in low-resource settings and over the course of sample handling, affect TL measurement accuracy. We report our DBS extraction technique and optimized qPCR conditions to serve as a reference for future work. We also suggest best practices for TL measurement derived from our reading of the literature and experience.

To complement our in-house validation work, we conducted a systematic literature review, from which we extracted published correlation values between DBS, salivary, and BC TL with both VWB TL and age. Correlations of TL values yielded from these modalities with VWB TL and age are used as external validation measures of TL measurements (Bateson, Eisenberg, & Nettle, 2018; Eisenberg, 2016; Stout et al., 2017). We subsequently incorporated the published correlation values, along with those from our new DBS validation work, into a meta-analysis.

Methods

UW Matched Specimens

In 2013, we recruited participants from the UW campus to donate matched biological specimens. We sampled matched DBS and VWB specimens from 19 individuals aged 20-50 years. Capillary finger-prick, using a BD Microtainer Lancet (Becton Dickinson, Franklin Lakes, NJ, USA) was used to draw blood that was preserved on Whatman 903 protein saver cards (GE Healthcare, Chicago, IL, USA). The first drop of blood was wiped away and discarded (to minimize skin or other tissue contaminating the sample). We spotted successive, non-overlapping drops of blood onto the card. DBS cards were then air-dried at RT overnight and then placed in Ziploc bags with desiccant packs. After collection, the VWB samples were stored immediately at -20°C and remained frozen until extraction with the Gentra Puregene Blood Kit protocol (Qiagen, Venlo, The Netherlands).

The VWB DNA used in this study, along with the Whatman 903 cards, were kept at -20°C for approximately five and a half years, undergoing between zero and three freeze-thaws in that time window. In 2019, we re-thawed and re-extracted the DBS specimens using a modified version of Qiagen's QIAamp DNA investigator kit extraction protocol (modifications explained in more detail below).

Emory Matched Specimens

In 2017, we collected VWB from 15 volunteers, filling one 9 mL sodium heparin tube for each volunteer. From this, we immediately pipetted 50 µL spots onto Whatman 903 cards and then centrifuged the tube of remaining blood at RT (18 to 20°C) using a

horizontal rotor (swing-out head) for a minimum of 20 minutes at 1,500 to 1,800 RCF (Relative Centrifugal Force). After removing the plasma layer by pipette, the buffy coat was pipetted (slowly to avoid RBC contamination) into 1.5 mL tubes. Some of the buffy coat was spotted (50 μ L spots) onto Whatman FTA classic cards. DBS were stored in plastic containers with desiccant packs after drying at RT for at least three hours.

The buffy coats and their accompanying FTA cards were shipped overnight to the University of California, San Francisco (UCSF) on dry ice and at RT, respectively. The 903 cards underwent an additional assessment of how storage and handling conditions influence DNA quantity and TL measurement. Specifically, we applied three conditions designed to reflect the options for sample processing and transport from Nepal (Table 1). Condition A assumed that exportation of DBS is possible, but that freezing is not. Condition B assumed that exporting DBS and freezing were possible and that dried ice was available for shipping. Finally, condition C assumed that shipping DBS cards was not possible, but that a local lab had a freezer and the required equipment for partial processing. DBS were kept in sealed plastic containers with desiccant packs. When at “room temperature,” they were in an air-conditioned laboratory.

Table 1. Experimental conditions applied to Emory DBS samples collected on Whatman 903 protein saver cards.

Condition	Description
A	7 days at room temperature (RT) before being shipped at RT.
B	Frozen immediately at -30°C and kept frozen while shipped.
C	Frozen immediately at -30°C, kept frozen while shipped, then partially extracted using the QIAamp DNA Investigator kit (up through addition of proteinase K), kept for 7 days at RT, then completed DNA extraction.

Dried Blood Spot DNA Extraction Protocol

DNA was extracted from DBS as described below. We used a modified version of Qiagen's protocol for isolation of total DNA from FTA and Guthrie cards (See page 17 of the June 2012 version of the QIAamp DNA Investigator Handbook—downloaded March 7, 2019 from <https://www.qiagen.com/us/applications/human-identity-and-forensics/human-identification-applications/assault-sample-analysis/differential-wash-workflow/downstream-sample-preparation/qiaamp-dna-investigator-kit/#resources>). First, we modified step 1 by punching six 3 mm (~1/8 in) holes instead of three. Second, we added 20 µl of Buffer ATE in step 17 for the UW samples and 50 µl for the Emory samples. Third, at step 18, we incubated at 56°C for three minutes instead of at RT. Finally, for the UW samples only, we repeated steps 17 and 18 twice, which yielded three 20 µl elutions. The UW samples were quantified by spectrophotometry using an Epoch Microplate Reader (BioTek, Winsooki, VT), and the Emory samples by Quant-iT™ PicoGreen™ dsDNA Assay Kit (Thermo-Fisher, Waltham, MA). For each, samples were purified using the same lot of the DNA extraction kit.

While developing this modified extraction protocol, we tested multiple elution volumes before we decided on 20 µl because of its good total DNA yield at a high concentration (Table S1). We also attempted extraction of the UW DBS samples using the Gentra Puregene Kit, QIAamp Mini Kit, QIAamp Micro Kit, and a Chelex bead protocol modified from Ceffa et al. (2012), all of which failed to yield sufficient DNA, before settling on the DNA investigator kit. When collecting the Emory samples, we also filled Vacutainer CPT™ tubes (cell preparation tubes with sodium heparin), from which we centrifuged peripheral blood mononuclear cells (PBMCs). For most samples, neither

the whole PBMC extracts nor those extracted from spotted FTA cards consistently yielded sufficient DNA to run the TL assay.

Telomere Length Measurement

For both the UW and Emory specimens, we used qPCR-based approaches to measure TL.

UW Samples. These samples were analyzed at UW using a modified version of the monochrome multiplex (MMQPCR) version of the assay (Cawthon, 2009). We assayed relative TL on a CFX 384 real-time PCR detection system (Bio-Rad, Hercules, CA, USA) using the same cycling profile reported by Eisenberg et al. (2012). We incorporated the primer pairs reported by Tackney et al. (2014) along with PowerUp SYBR Green master mix (Applied Biosystems, Waltham, MA). Our final reaction volume was 5 μ l, which contained 1.0 μ l DNA, 2.5 μ l master mix, 1.1 μ l primers, and 0.4 μ l water. Detailed instructions are provided in our online supplement along with master mix, primer, and water storage/stock concentration information, which is listed in Table S2.

We used high-quality/quantity (260/280 ratio = 1.79; conc. = 498.23 ng/ μ l) DNA extracted in-house from whole blood to create a six-point standard reference curve, used to determine the quantity of targeted templates for each sample. A three-fold serial dilution was used. Standard curve concentrations ranged from 5 ng/ μ l to 0.02 ng/ μ l. From a single-use frozen stock, we diluted standard DNA to working concentrations immediately prior to assay. We then diluted the same high-quality DNA to 6.25 ng/ μ l for use as a positive control.

We assayed all UW specimens twice on separate PCR plates. All DNA standards, samples, and negative controls were included in triplicate on each assayed plate. Average standard curve R^2 was 0.99 (SD < 0.01) for the telomere (T) amplicon and 0.99 (SD < 0.01) for that of the single copy gene (S—albumin for UW, human β -globin for Emory). Plate assay efficiency was 103.2% (SD = 1.3%) on average for T, and 94.05% (SD < 0.95%) for S.

Emory samples. Relative TL for the Emory samples was measured at UCSF using a modified version of Cawthon's (2002) singleplex qPCR procedure. We employed the same thermocycling profile, primer pairs, and PCR reagent concentrations as Stout et al. (2017), with one exception: we used 200 μ M of each dNTP (see Supplemental Material for detailed protocol). For our standard curve, we used reference DNA derived from pooled human genomic DNA purchased from Roche (Cat# 11691112001, Basel, Switzerland) to create a six-point, ~three-fold curve (ranging from 2.36 ng/ μ l to 0.01 ng/ μ l). The same reference DNA was used for all PCR runs. We also included eight positive controls in each run. For the samples, each was run in triplicate across two plates (one plate for T and one for S). Plate assay efficiencies were 88.8% (SD = 1.3%) for T and 90.1% (SD = 1.0%) for S. Average standard curve R^2 was 0.99 (SD < 0.01) for both the T and S curves.

Relative Telomere Length Data Analysis

When using either the singleplex or the multiplex version of the qPCR relative assay, the unit of analysis is T/S ratio. For both the UW and Emory samples, we calculated T/S ratio by dividing the estimated starting quantity (SQ) of T by the

estimated SQ of S. SQ is determined by where the amplified sample falls on the standard curve.

UW. We averaged raw T/S ratio across sample triplicates, which we included in our analyses. Dixon's Q test was used to identify any replicate outliers, which led to the removal of three observations prior to analysis. Since the coefficient of variation (CV) is not valid to assess reproducibility (Eisenberg, 2016; Verhulst et al., 2016, 2015), we calculated the intraclass correlation coefficient (ICC) of the average of triplicate values run on separate plates (ICC1 = 0.86, 95% CI = 0.74, 0.92; ICC1k = 0.92, 95% CI = 0.85, 0.96).

Emory. The T/S ratio of each positive control was divided by the average T/S ratio for the same DNA from ten runs to get a normalizing factor, which we used to correct the batch differences and obtain the final T/S ratio included in our analyses. ICC of averaged triplicate values were: ICC1 = 0.97, 95% CI = 0.95, 0.99; ICC1k = 0.99, 95% CI = 0.98, 0.99. Each sample was run twice. If T/S ratios varied by more than 7%, the sample was run in triplicate for a third time and the two closest values were reported. We ran 12 samples three times out of 45 total samples.

Literature Search and Inclusion Criteria

To synthesize the literature on minimally-invasive sampling techniques applied to TL analysis, we used the following criteria to search both the PubMed and Scopus databases: telomere AND (Saliva OR Dried Blood Spot OR Buccal) AND Human on March 7, 2019. We screened the abstracts of the resulting 101 studies and excluded any that did not actually measure TL, included an animal model, or were not written in

English. We then assessed the full-text of the remaining 73 articles for eligibility. Our eligibility criteria included the following: TL measured in either DBS, saliva, or BC; at least one external validation statistic reported with blood TL (from whole blood or a derivate) and/or continuously reported age data; and a test statistic that is convertible to a correlation coefficient. Our eligibility scan led to a candidate set of 23 articles. We also included one of our unpublished correlation values. A more detailed summary of our protocol is summarized in a PRISMA flow diagram (Figure S1).

Statistical Analysis

For our two sets of matched DBS and VWB specimens, we calculated Pearson's correlation coefficients between the two TL measurements (UW DBS TL vs. UW VWB TL AND Emory DBS TL vs. Emory buffy coat TL) and between UW DBS TL and age. We then included these correlation values, along with those identified in our literature search, in a series of six meta-analyses in order to measure the effect sizes across all studies reporting measures of external validity: 1) DBS TL vs. blood TL; 2) Saliva TL vs. blood TL; 3) BC TL vs. blood TL; 4) DBS TL vs. age; 5) Saliva TL vs. age; 6) BC TL vs. age. Before inclusion, we standardized all reported associations to Fisher z-transformed correlation coefficients (as in Pepper et al. 2018). We ran our analysis and created all subsequent forest plots using the "metafor" package in R (R Development Core Team, 2019; Wolfgang, 2010). One-way ANOVAs were used to determine if DBS and VWB TL were significantly different.

Results

In-House Matched Specimen Correlations

Our modified DBS extraction protocol used for the 19 UW subjects produced average DNA yields of 530.34 ± 127.88 ng, which is up to ten times more than published values (Goldman et al., 2017; Stout et al., 2017). The average yields for the Emory samples were: 369 ± 140 ng, 469 ± 198 ng, and 354 ± 96 for conditions A, B, and C, respectively. For the UW, the average 260/280 ratio was 1.97 across all elutions. TL measures for the UW matched specimens were moderately correlated ($r = 0.56$, 95% CI = 0.15, 0.81, Figure 1A). The correlation between DBS TL measures and age was not significant; however, it was trending in the expected negative direction ($r = -0.28$, 95% CI = -0.65, 0.20, Figure S2). This association was also slightly stronger than the correlation between the matched VWB specimens and age ($r = -0.25$, 95% CI = -0.63, 0.23). Both sets of samples manifested a significant difference in mean measured TL between DBS and blood, with DBS being significantly longer (UW: $F = 7.949$, $p = 0.019$; Emory DBS vs. frozen buffy coat samples: $F = 6.915$, $p = 0.003$). For the Emory samples, measured TL was significantly different ($F = 6.915$, $p = 0.003$) between DBS handled under different conditions (A, B, and C). DBS handled under condition C exhibited the strongest correlation with buffy coat TL in matched specimens (Condition A: $r = 0.84$, 95% CI = 0.58, 0.95; Condition B: $r = 0.93$, 95% CI = 0.80, 0.98; Condition C: $r = 0.95$, 95% CI = 0.86, 0.98; Figure 1B). Using standard ANOVA testing, we found that TL measured from no one condition was a significantly better predictor of buffy coat TL. Ages were not available for the Emory participants.

Figure 1. Correlations between telomere length (TL) measured from dried blood spots and A) whole blood for the University of Washington (UW) matched specimens and B) Buffy coats for the Emory University matched specimens.

A) UW matched specimens ($r = 0.56$).

B) Emory matched specimens (Condition A: $r = 0.84$; Condition B: $r = 0.93$; Condition C: $r = 0.95$).

Description of Dataset for Meta-Analyses

Including our new DBS data from the UW and Emory samples, our final dataset consisted of 25 studies. Of those, five reported external validation statistics for DBS, eleven for saliva, and ten for BCs. Stout et al. (2017) measured TL in VWB, DBS, and saliva, while our UW study and Goldman et al. (2018) report correlations with VWB and age. Of the eleven studies that measured salivary TL, only two did not use Oragene collection kits (Fillman, Shimizu-Furusawa, Ng, Parajuli, & Watanabe, 2016; Lahnert, 2005). Table S3 reports the associations and provides details of the papers included. All

but one of these studies used a qPCR-based method to measure TL (Lahnert et al., 2005).

Overall Associations

Figure 2 includes the correlations for the reported associations between blood TL and TL measured in DBS, saliva, and BC. DBS produced the strongest overall correlation with blood ($r = 0.82$, 95% CI = 0.73, 0.88), then saliva ($r = 0.61$, 95% CI = 0.46, 0.72), then BC ($r = 0.40$, 95% CI = 0.36, 0.43). None of the 95% confidence intervals overlapped.

Figure 2. Correlations estimates and 95% confidence intervals for separate meta-analytic models between blood TL and TL measured from A) dried blood spots, B) saliva, and C) buccal cells.

A) Dried Blood Spot Telomere Length vs. Venus Whole Blood Telomere Length

B) Saliva Telomere Length vs. Venous Whole Blood Telomere Length

C) Buccal Cell Telomere Length vs. Venous Whole Blood Telomere Length

We present the correlations of the reported overall associations between age and A) DBS TL, B) salivary TL, and C) BC TL in Figure 3. The overall association with age is the strongest for DBS ($r = -0.39$, 95% CI = -0.51 , -0.26), followed by saliva ($r = -0.18$, 95% CI = -0.23 , -0.13), then BC ($r = -0.14$, 95% CI = -0.20 , -0.07). As before, there is no overlap in the 95% confidence intervals for the DBS vs. age comparison and the other two comparisons. However, there is overlap between saliva and BC, with the mean estimate for each falling within the other's interval.

Figure 3. Correlations estimates and 95% confidence intervals for separate meta-analytic models between chronological age and TL measured from A) dried blood spots, B) saliva, and C) buccal cells.

A) Dried Blood Spot Telomere Length vs. Age

B) Saliva Telomere Length vs. Age

C) Buccal Cell Telomere Length vs. Age

We ran two additional saliva TL vs. age analyses. First, after excluding the two studies that did not use Oragene kits for extraction, we saw a slight increase in overall effect strength, $r = -0.201$ (-0.258, -0.144). In the second, we only excluded the single

study that measured TL using TRF, which led to a slight decrease in effect strength, $r = -0.175$ ($-0.225, -0.124$).

Discussion

In this study, we presented our modified DBS DNA extraction protocol for use with TL measurement. This procedure provided substantially larger yields than those previously reported: 530.34 ng (UW samples) and 397.33 ng (Emory samples) vs. 59.1 ng (Stout et al., 2017) and 110.4 ng (Goldman et al., 2017). However, considering our A280/A260 ratios, and the lack of an RNase A treatment step during the extraction protocol (also not employed by either Stout or Goldman et al.), yield estimates may be slightly inflated. Using this protocol, we extracted two sets of DBS samples for which we had matched VWB specimens. After calculating external validation statistics for both sets (correlation values with VWB TL and age), we examined how different storage/shipping conditions affected TL measurements, and reported the optimal environments for those facing strict sample export constraints. We also incorporated these validation measures into our meta-analysis, along with results from 23 other studies identified by systematic literature review and one value from our own unpublished work. This meta-analysis established that DBS TL had the strongest overall association with both VWB and age, followed by saliva, then BC. For both comparisons, there was no 95% CI overlap between DBS and saliva.

Shipping and storage conditions are both important considerations for accurate TL measurement. To simulate the international transport of DBS from countries with

varying requirements and restrictions that could result in delays during shipment, pre-exportation extraction, or varying access to frozen shipment facilities, we exposed the DBS from one of our sets of matched specimens to three different conditions: 1) stored for seven days at RT and then shipped at RT; 2) stored at -30°C and shipped in dry ice; and 3) frozen immediately at -30°C and shipped in dry ice then partially extracted and stored for seven days at RT before extraction was completed. All three conditions produced similar external validation statistics, which indicates that researchers could use the partial extraction protocol when faced with regulations that prohibit the transport of unprocessed blood spots out of a host country but where in-country facilities are not available for final analysis. As for our other set of matched specimens, in addition to being housed for six years at -80°C, the VWB samples were extracted using a different technique (Gentra Puregene) than the Emory samples. There are multiple lines of evidence to suggest that differing extraction techniques influence TL measurements (Cunningham et al., 2013; Dagnall et al., 2017; Denham et al., 2014; Raschenberger et al., 2016).

Our meta-analysis suggested that TL dynamics vary across different sample types collected using minimally invasive techniques. We expected this result due to the different cellular compositions of DBS (which should be similar to VWB), saliva, and BC. The latter two consist of roughly 79% and 26% epithelial cells, respectively (Thiede et al., 2000). Although white blood cells likely contribute most of the DNA content in saliva, DBS TL was significantly more strongly correlated with both blood TL and age. This result indicates that DBS TL most strongly approximates TL measured from VWB, suggesting that DBS may be useful in research aimed at predictions/inferences about

TL and the immune system. Conversely, caution is warranted when interpreting TL results derived from saliva and BC. Although these orally derived samples might be informative about the telomere biology of the epithelium, it is also likely that these samples yield noisier TL measures. For in addition to tissue variability, these samples come from an orifice that is dynamically interacting with and influenced by many environmental and behavioral factors such as eating, drinking, and oral health (e.g., microbial colonization, lesions). All of these elements could directly change the cell types from which DNA is derived and, thus, influence the purity of DNA extracts.

Recommendations for Additional Best Practices in Measuring Telomeres

Several publications have reported different TL measurement techniques (Cawthon, 2002, 2009; Kimura et al., 2010; Mender & Shay, 2016; O'Sullivan et al., 2005), their relative costs and benefits (e.g., Lai, Wright, & Shay, 2018; Montpetit et al., 2014), and some of the best practices for analysis and reporting (Lin, Smith, Esteves, & Drury, 2019). Nonetheless, several measurement issues persist in the telomere biology literature. We outline some of these below, focusing on issues of statistical power, measurement error, internal validity, and external validity.

Power. Previous analysts have noted that many published TL investigations are underpowered (e.g., Pepper et al., 2018). As in any study, we recommend that researchers consult observed effect sizes from past studies (especially from meta-analyses) and conduct a power analysis during study design to determine if their sample is large enough to have reasonable statistical power. For example, taking the meta-analysis derived estimated association of TL with stresses and adversities of $r = -$

0.15 (Pepper et al., 2018), a sample size of 359 would be needed to detect a significant effect ($p < 0.05$) of this size as different from zero with 80% power. In this same meta-analysis, 62% of the 138 studies fell below this sample size threshold. A publication bias-corrected estimate (only including studies with sample sizes over 100) in this same study yielded a stress/adversity-TL association of $r = -0.09$ ($p = 0.04$); $r = -0.03$ ($p = \text{NS}$ for their entire sample). Detecting an effect of -0.09 would require a sample size of 1,007. We note that these effect sizes assume that investigators are merely trying to detect a main effect. Human biologists are often interested in variation of effects with age, sex, and SES—all of which would require larger samples to discern. These effect sizes are largely based on data derived using qPCR techniques, and therefore assume typical levels of measurement error observed in the literature. As discussed in more detail below, measurement error, which is a considerable problem in TL studies, could further reduce statistical power.

Internal validity. Investigators can expect to encounter substantial measurement issues during TL assay development. When analyzing qPCR output data, researchers should confirm that all samples are amplifying within the bounds of their standard curves (aka calibration curves; Bustin et al., 2009). For the MMQPCR assay, in particular, T should be amplifying several cycles before S, and there should be two observable melt curves (note: in some cases, the T peak may be absent but still appear in gel electrophoresis).

As discussed earlier, measurement error is a substantial problem in the telomere biology literature. Exacerbating this situation is the recent realization that the conventionally used CV measure of error is invalid for TL (Eisenberg, 2016; Nettle,

Seeker, Nussey, Froy, & Bateson, 2018; Verhulst et al., 2016, 2015). Researchers should instead report the ICC of their repeated measures (average T/S ratio across multiple plates), which allows for a more accurate gauge of both the measurement error and statistical power.

External validity. Even with strong internal repeatability of a putative TL quantification, TL might not be measured well—i.e., it may have low external validity. Many publications assume that, because a TL assay has been externally validated in another lab with another cohort of samples, and because they have good internal validity, their results must also be externally valid. After exhaustively troubleshooting the qPCR telomere length assay, studying the telomere literature, and discussing telomere length measurement with colleagues, we believe this is an assumption that should be more carefully questioned. In our experience, it is easy to generate qPCR TL results that appear valid at first, but upon closer examination, have a high degree of error and/or potential bias. For example, we previously showed dramatic differences in TL across European countries (Eisenberg, Salpea, Kuzawa, Hayes, & Humphries, 2011), which have not been observed in subsequent studies and we now believe represent collection/analytical artifacts.

External validity of TL measures can be established in multiple ways. The best method to assess TL is by measuring the same samples using multiple techniques (e.g., comparing qPCR to telomere restriction fragment (TRF) analysis or telomere analysis by fluorescence in situ hybridization (Tel-FISH)). If this is not possible, known robust predictors of TL provide some limited evidence of external validity. These include predictors such as age, familial relatedness, paternal age, and specific genetic

polymorphisms. However, using such predictors can be problematic if one is studying populations where these factors themselves might vary biologically. For example, infections likely accelerate TL shortening (Eisenberg, Borja, Hayes, & Kuzawa, 2017). Thus, in an environment with more randomly fluctuating rates of infection, we would expect to observe an attenuation of the effects from other predictors. The correlation between baseline and follow up TL measures also can provide a measure of external validity (Bateson et al., 2018), but again this may be problematic if substantial changes in TL reflect not only environmental factors but also the contexts in which those factors occur.

Finally, because DBS makes TL research more feasible in low-resource settings, it is important to consider ethical issues in these contexts (Gautam, Pedersen, Wahid, & Kohrt, 2019). Developing collaborations with researchers in the setting where DBS are being collected can not only improve the quality of science but can also help to give back by building local capacity and knowledge. Awareness of international ethical guidelines and local policies and regulations are crucial to respect both participants and local institutions. Efforts should be undertaken to inform local ethical review boards about the process and benefits of DBS.

Conclusions

In conclusion, we recommend DBS to researchers interested in using minimally invasively sampling for TL measurement in settings with limited medical and laboratory resources. Of the minimally invasive sample types included in our meta-analysis, DBS measured TL was that closest approximates to TL from whole blood measured. Saliva

samples are amongst the easiest to collect and store, however, researchers using them for TL measurements must acknowledge differences in cell composition, as well as the complexities of the oral cavity and how they moderate the predictions that can be made. We advise against using BCs for telomere measurement unless hypotheses are specifically related to epithelial cell TL. While not invasively collected, BCs exhibited the weakest overall correlation values with our measures of external validity, and they also have demanding storage requirements, similar to blood. In addition to these suggestions, we provided researchers with a variant of the QIAamp DNA Investigator Kit extraction protocol for DBS that produced yields twice as large as any previously reported method. Based on our results, a potential next step in the modification/development of this extraction protocol is the treatment of the extract with RNase A.

We also present data that begin to address best practices for maintaining DNA when shipping and storing DBS. We note that the set of storage conditions that we tested experimentally are just a starting point and certainly do not encompass all conditions researchers are likely to experience. Lastly, based on our experiences, we strongly suggest that any TL work with minimally invasively collected samples be done only after labs have first integrated the discussed extraction protocol AND have validated the qPCR/MMQPCR method in house. These methods will create new opportunities for research in settings with high rates of infectious disease and among populations affected by war, disasters, and other humanitarian emergencies.

Acknowledgements

We thank Josh Snodgrass and Geeta Eick for organizing this special issue and the conference session that begat it. We also thank Kathy Ton for her efforts during method development and expansion, and Amanda Guyton and Ellie Brindle for help with UW matched specimens collections and study design. Jue Lin is a co-founder of Telomere Diagnostics Inc. and serves on its scientific advisory board. The other authors have no conflict of interest.

Grant Sponsorship

NSF Grant BCS 151911

References

- Aps, J. K. M., Van Den Maagdenberg, K., Delanghe, J. R., & Martens, L. C. (2002). Flow cytometry as a new method to quantify the cellular content of human saliva and its relation to gingivitis. *Clinica Chimica Acta*, 321(1–2), 35–41.
[https://doi.org/10.1016/S0009-8981\(02\)00062-1](https://doi.org/10.1016/S0009-8981(02)00062-1)
- Armanios, M., & Blackburn, E. H. (2012). The telomere syndromes. *Nature Reviews Genetics*, 13(10), 693–704. <https://doi.org/10.1038/nrg3246>
- Auld, E., Lin, J., Chang, E., Byanyima, P., Ayakaka, I., Musisi, E., ... Huang, L. (2016). HIV infection is associated with shortened telomere length in ugandans with suspected tuberculosis. *PLoS ONE*, 11(9).
<https://doi.org/10.1371/journal.pone.0163153>
- Bateson, M., Eisenberg, D. T. A., & Nettle, D. (2018). Controlling for baseline telomere length biases estimates of the effect of smoking on leukocyte telomere attrition in longitudinal studies. *Zenodo*, (July), 1–22. <https://doi.org/10.5281/zenodo.1009086>
- Bond, M. M., & Richards-Kortum, R. R. (2015). Drop-to-drop variation in the cellular components of fingerprick blood: Implications for point-of-care diagnostic development. *American Journal of Clinical Pathology*, 144(6), 885–894.
<https://doi.org/10.1309/AJCP1L7DKMPCHPEH>
- Bustin, S. A., Hellemans, J., Pfaffl, M. W., Wittwer, C. T., Garson, J. A., Huggett, J., ... Nolan, T. (2009). The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry*, 55(4), 611–622.
<https://doi.org/10.1373/clinchem.2008.112797>

- Cawthon, R. M. (2002). Telomere measurement by quantitative PCR. *Nucleic Acids Research*, 30(10), e47. <https://doi.org/10.1093/nar/30.10.e47>
- Cawthon, R. M. (2009). Telomere length measurement by a novel monochrome multiplex quantitative PCR method. *Nucleic Acids Research*, 37(3), 1–7. <https://doi.org/10.1093/nar/gkn1027>
- Cawthon, R. M., Smith, K. R., O'Brien, E., Sivatchenko, A., & Kerber, R. A. (2003). Association between telomere length in blood and mortality in people aged 60 years or older. *Lancet*, 361(9355), 393–395. [https://doi.org/10.1016/S0140-6736\(03\)12384-7](https://doi.org/10.1016/S0140-6736(03)12384-7)
- Ceffa, S., Erba, F., Brambilla, D., Luhanga, R., Armel Pépé, M., Paturzo, G., ... Marazzi, M. C. (2012). Evaluation of different deoxyribonucleic acid (DNA) extraction methods using dried blood spot for early infant diagnosis of HIV1 in Sub-Saharan Africa. *African Journal of Microbiology Research*, 6(27). <https://doi.org/10.5897/ajmr11.1469>
- Chatterjee, D., Bhattacharjee, P., Sau, T. J., Das, J. K., Sarma, N., Bandyopadhyay, A. K., ... Giri, A. K. (2015). Arsenic exposure through drinking water leads to senescence and alteration of telomere length in humans: A case-control study in West Bengal, India. *Molecular Carcinogenesis*, 54(9), 800–809. <https://doi.org/10.1002/mc.22150>
- Cong, Y., Wright, W. E., & Shay, J. W. (2002). Human Telomerase and Its Regulation Human Telomerase and Its Regulation. *Microbiology and Molecular Biology Reviews*, 66(3), 407. <https://doi.org/10.1128/MMBR.66.3.407>
- Cunningham, J. M., Johnson, R. A., Litzelman, K., Skinner, H. G., Seo, S., Engelman,

- C. D., ... Boardman, L. A. (2013). Telomere length varies by DNA extraction method: Implications for epidemiologic research. *Cancer Epidemiology Biomarkers and Prevention*, 22(11), 2047–2054. <https://doi.org/10.1158/1055-9965.EPI-13-0409>
- Dagnall, C. L., Hicks, B., Teshome, K., Hutchinson, A. A., Gadalla, S. M., Khincha, P. P., ... Savage, S. A. (2017). Effect of pre-analytic variables on the reproducibility of qPCR relative telomere length measurement. *PLoS ONE*, 12(9). <https://doi.org/10.1371/journal.pone.0184098>
- Denham, J., Marques, F. Z., & Charchar, F. J. (2014). Leukocyte telomere length variation due to DNA extraction method. *BMC.Res.Notes*, 7, 877-.
- Eisenberg, D. T. A., Hayes, M. G., & Kuzawa, C. W. (2012). Delayed paternal age of reproduction in humans is associated with longer telomeres across two generations of descendants. *Proceedings of the National Academy of Sciences*, 109(26), 10251–10256. <https://doi.org/10.1073/pnas.1202092109>
- Eisenberg, D. T. A., Salpea, K. D., Kuzawa, C. W., Hayes, M. G., & Humphries, S. E. (2011). Substantial variation in qPCR measured mean blood telomere lengths in young men from eleven European countries. *American Journal of Human Biology*, 23(2), 228–231. <https://doi.org/10.1002/ajhb.21126>
- Eisenberg, D. T.A. (2016). Telomere length measurement validity: the coefficient of variation is invalid and cannot be used to compare quantitative polymerase chain reaction and Southern blot telomere length measurement techniques. *International Journal of Epidemiology*, (Cv), dyw191. <https://doi.org/10.1093/ije/dyw191>
- Eisenberg, D. T. A., Borja, J. B., Hayes, M. G., & Kuzawa, C. W. (2017). Early life

- infection, but not breastfeeding, predicts adult blood telomere lengths in the Philippines. *American Journal of Human Biology*, 29(4).
<https://doi.org/10.1002/ajhb.22962>
- Eisenberg, D. T. A., Lee, N. R., Rej, P. H., Hayes, M. G., & Kuzawa, C. W. (2019). Older paternal ages and grandpaternal ages at conception predict longer telomeres in human descendants. *Proceedings of the Royal Society B: Biological Sciences*, 286(1903). <https://doi.org/10.1098/rspb.2019.0800>
- Epel, E. S., & Prather, A. A. (2018). Stress, Telomeres, and Psychopathology: Toward a Deeper Understanding of a Triad of Early Aging. *Annual Review of Clinical Psychology*, 14, 371–397. <https://doi.org/10.1146/annurev-clinpsy-032816-045054>
- Farrukh, S., Baig, S., Hussain, R., Shahid, A., & Khan, S. T. (2019). Telomere reprogramming during fetal life in low socioeconomic mothers. *Egyptian Journal of Medical Human Genetics*, 20(1), 9. Retrieved from 10.1186/s43042-019-0007-4
- Farrukh, Sadia, Baig, S., Hussain, R., & Lucky, M. H. (2019). Association of Cord Blood Telomere Biology with Mother's Education. *International Journal of Biochemistry Research & Review*, 24(3), 1–9. <https://doi.org/10.9734/ijbcr/2018/46192>
- Fillman, T., Shimizu-Furusawa, H., Ng, C. F. S., Parajuli, R. P., & Watanabe, C. (2016). Association of cadmium and arsenic exposure with salivary telomere length in adolescents in Terai, Nepal. *Environmental Research*, 149, 8–14.
<https://doi.org/10.1016/j.envres.2016.04.037>
- Gautam, K., Pedersen, G. A., Wahid, S. S., & Kohrt, B. A. (2019). Ethical considerations for conducting cross-cultural biological psychiatry and prevention research on depression among adolescents in low- and middle-income countries.

<https://doi.org/10.1016/bs.dnb.2019.05.001>

- Goldman, E. A., Eick, G. N., Compton, D., Kowal, P., Snodgrass, J. J., Eisenberg, D. T. A., & Sterner, K. N. (2017). Evaluating minimally invasive sample collection methods for telomere length measurement. *American Journal of Human Biology*, (September 2016), e23062. <https://doi.org/10.1002/ajhb.23062>
- Greider, C. W., & Blackburn, E. H. (1987). The telomere terminal transferase of tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell*, 51(6), 887–898. [https://doi.org/10.1016/0092-8674\(87\)90576-9](https://doi.org/10.1016/0092-8674(87)90576-9)
- Jiménez, K. M., Pereira-Morales, A. J., Adan, A., & Forero, D. A. (2019). Telomere length and childhood trauma in Colombians with depressive symptoms. *Revista Brasileira de Psiquiatria (Sao Paulo, Brazil)*: 1999, 41(3), 194–198. <https://doi.org/10.1590/1516-4446-2018-0027>
- Kimura, M., Stone, R. C., Hunt, S. C., Skurnick, J., Lu, X., Cao, X., ... Aviv, A. (2010). Measurement of telomere length by the Southern blot analysis of terminal restriction fragment lengths. *Nature Protocols*, 5(9), 1596–1607. <https://doi.org/10.1038/nprot.2010.124>
- Lahnert, P. (2005). An improved method for determining telomere length and its use in assessing age in blood and saliva. *Gerontology*, 51(5), 352–356. <https://doi.org/10.1159/000086374>
- Lai, T. P., Wright, W. E., & Shay, J. W. (2018). Comparison of telomere length measurement methods. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1741). <https://doi.org/10.1098/rstb.2016.0451>
- Lin, J., Cheon, J., Brown, R., Coccia, M., Puterman, E., Aschbacher, K., ... Blackburn,

- E. H. (2016). Systematic and Cell Type-Specific Telomere Length Changes in Subsets of Lymphocytes. *Journal of Immunology Research*, 2016.
<https://doi.org/10.1155/2016/5371050>
- Lin, J., Smith, D. L., Esteves, K., & Drury, S. (2019). Telomere length measurement by qPCR – Summary of critical factors and recommendations for assay design. *Psychoneuroendocrinology*, 99, 271–278.
<https://doi.org/10.1016/j.psyneuen.2018.10.005>
- Lingner, J., Cooper, J. P., & Cech, T. R. (1995). Telomerase and DNA end replication: No longer a lagging strand problem? *Science*, 269(5230), 1533–1534.
<https://doi.org/10.1126/science.7545310>
- Malan, S., Hemmings, S., Kidd, M., Martin, L., & Seedat, S. (2011). Investigation of telomere length and psychological stress in rape victims. *Depression and Anxiety*, 28(12), 1081–1085. <https://doi.org/10.1002/da.20903>
- Mathur, M. B., Epel, E., Kind, S., Desai, M., Parks, C. G., Sandler, D. P., & Khazeni, N. (2016). Perceived stress and telomere length: A systematic review, meta-analysis, and methodologic considerations for advancing the field. *Brain, Behavior, and Immunity*, 54, 158–169. <https://doi.org/10.1016/j.bbi.2016.02.002>
- McDade, T. W., Williams, S., & Snodgrass, J. J. (2007). What a drop can do: Dried blood spots as a minimally invasive method for integrating biomarkers into population-based research. *Demography*, 44(4), 899–925.
<https://doi.org/10.1353/dem.2007.0038>
- Mender, I., & Shay, J. (2016). Telomere Restriction Fragment (TRF) Analysis. *Bio-Protocol*, 5(22). <https://doi.org/10.21769/bioprotoc.1658>

- Montpetit, A. J., Alhareeri, A. A., Montpetit, M., Starkweather, A. R., Elmore, L. W., Filler, K., ... Jackson-Cook, C. K. (2014). Telomere Length: A review of methods for measurement. *Nursing Research*, 63(4), 289–299.
<https://doi.org/10.1097/nnr.0000000000000037>
- Nedel, F., André, D. de A., Oliveira, I. O. de, Tarquinio, S. B. C., & Demarco, F. F. (2009). Buccal cells submitted to three different storage conditions before DNA extraction. *Journal of Applied Oral Science*, 17(2), 113–115.
<https://doi.org/10.1590/s1678-77572009000200008>
- Nettle, D., Seeker, L., Nussey, D., Froy, H., & Bateson, M. (2019). Consequences of measurement error in qPCR telomere data: A simulation study. *PLoS ONE*, 14(5).
<https://doi.org/10.1371/journal.pone.0216118>
- O’Sullivan, J. N., Finley, J. C., Risques, R., Shen, W.-T., Gollahon, K. A., & Rabinovitch, P. S. (2005). Quantitative Fluorescence In Situ Hybridization (QFISH) of Telomere Lengths in Tissue and Cells. *Current Protocols in Cytometry*.
<https://doi.org/10.1002/0471142956.cy1206s33>
- Oliveira, B. S., Zunzunegui, M. V., Quinlan, J., Batistuzzo De Medeiros, S. R., Thomasini, R. L., & Guerra, R. O. (2018). Lifecourse Adversity and Telomere Length in Older Women from Northeast Brazil. *Rejuvenation Research*, 21(4), 294–303. <https://doi.org/10.1089/rej.2017.1937>
- Palm, W., & de Lange, T. (2008). How Shelterin Protects Mammalian Telomeres. *Annual Review of Genetics*, 42(1), 301–334.
<https://doi.org/10.1146/annurev.genet.41.110306.130350>
- Pathai, S., Lawn, S. D., Gilbert, C. E., McGuinness, D., McGlynn, L., Weiss, H. A., ...

- Shiels, P. G. (2013). Accelerated biological ageing in HIV-infected individuals in South Africa: A case-control study. *Aids*, 27(15), 2375–2384.
<https://doi.org/10.1097/QAD.0b013e328363bf7f>
- Pepper, G. V., Bateson, M., & Nettle, D. (2018). Telomeres as integrative markers of exposure to stress and adversity: A systematic review and meta-analysis. *Royal Society Open Science*, 5(8). <https://doi.org/10.1098/rsos.180744>
- Quinlan, J., Tu, M. T., Langlois, E. V, Kapoor, M., Ziegler, D., Fahmi, H., & Zunzunegui, M. V. (2014). Protocol for a systematic review of the association between chronic stress during the life course and telomere length. *Syst Rev*, 3, 40.
<https://doi.org/10.1186/2046-4053-3-40>
- Raschenberger, J., Lamina, C., Haun, M., Kollerits, B., Coassin, S., Boes, E., ... Kronenberg, F. (2016). Influence of DNA extraction methods on relative telomere length measurements and its impact on epidemiological studies. *Scientific Reports*, 6. <https://doi.org/10.1038/srep25398>
- Rej, P. H., Tennyson, R. L., Lee, N. R., & Eisenberg, D. T. A. (2019). Years of caregiving for chronically ill and disabled family members is not associated with telomere length in the Philippines. *Psychoneuroendocrinology*, 103(October 2018), 188–194. <https://doi.org/10.1016/J.PSYNEUEN.2019.01.019>
- Rode, L., Nordestgaard, B. G., & Bojesen, S. E. (2015). Peripheral blood leukocyte telomere length and mortality among 64 637 individuals from the general population. *Journal of the National Cancer Institute*, 107(6).
<https://doi.org/10.1093/jnci/djv074>
- Serrano, A. L., & Andrés, V. (2004). Telomeres and Cardiovascular Disease: Does Size

Matter? *Circulation Research*, 94(5), 575–584.

<https://doi.org/10.1161/01.RES.0000122141.18795.9C>

Stout, S. A., Lin, J., Hernandez, N., Davis, E. P., Blackburn, E., Carroll, J. E., & Glynn, L. M. (2017). Validation of minimally-invasive sample collection methods for measurement of telomere length. *Frontiers in Aging Neuroscience*, 9(DEC).

<https://doi.org/10.3389/fnagi.2017.00397>

Tackney, J., Cawthon, R. M., Coxworth, J. E., & Hawkes, K. (2014). Blood cell telomere lengths and shortening rates of chimpanzee and human females. *American Journal of Human Biology*, 26(4), 452–460. <https://doi.org/10.1002/ajhb.22538>

Team, R. D. C. (2019). R: A language and environment for statistical computing.

Retrieved from <http://www.gbif.org/resource/81287>

Tennyson, R. L., Gettler, L. T., Kuzawa, C. W., Hayes, M. G., Agustin, S. S., & Eisenberg, D. T. A. (2018). Lifetime socioeconomic status and early life microbial environments predict adult blood telomere length in the Philippines. *American Journal of Human Biology*, (September 2017), 1–10.

<https://doi.org/10.1002/ajhb.23145>

Theda, C., Hwang, S. H., Czajko, A., Loke, Y. J., Leong, P., & Craig, J. M. (2018). Quantitation of the cellular content of saliva and buccal swab samples. *Scientific Reports*, 8(1). <https://doi.org/10.1038/s41598-018-25311-0>

Thiede, C., Prange-Krex, G., Freiberg-Richter, J., Bornhauser, M., & Ehninger, G. (2000). Buccal swabs but not mouthwash samples can be used to obtain pretransplant DNA fingerprints from recipients of allogeneic bone marrow transplants. *Bone Marrow Transplantation*, 25(5), 575–577.

<https://doi.org/10.1038/sj.bmt.1702170>

Verhulst, S., Susser, E., Factor-Litvak, P. R., Simons, M., Benetos, A., Steenstrup, T., ... Aviv, A. (2016). Response to: Reliability and validity of telomere length measurements. *International Journal of Epidemiology*, 45(4), 1298–1301.

<https://doi.org/10.1093/ije/dyw194>

Verhulst, S., Susser, E., Factor-Litvak, P. R., Simons, M. J. P., Benetos, A., Steenstrup, T., ... Aviv, A. (2015). Commentary: The reliability of telomere length measurements. *International Journal of Epidemiology*, 44(5), 1683–1686.

<https://doi.org/10.1093/ije/dyv166>

Von Känel, R., Malan, N. T., Hamer, M., & Malan, L. (2015). Comparison of telomere length in black and white teachers from south africa: The sympathetic activity and ambulatory blood pressure in africans study. *Psychosomatic Medicine*, 77(1), 26–32. <https://doi.org/10.1097/PSY.0000000000000123>

Wang, Q., Zhan, Y., Pedersen, N. L., Fang, F., & Hägg, S. (2018). Telomere Length and All-Cause Mortality: A Meta-analysis. *Ageing Research Reviews*, 48, 11–20.

<https://doi.org/10.1016/j.arr.2018.09.002>

Wolfgang, V. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1–48.

Woo, J. S., & Lu, D. Y. (2019). Procurement, transportation, and storage of saliva, buccal swab, and oral wash specimens. *Methods in Molecular Biology*, 1897, 99–105. https://doi.org/10.1007/978-1-4939-8935-5_10

Wright, W. E., Tesmer, V. M., Huffman, K. E., Levene, S. D., & Shay, J. W. (1997). Normal human chromosomes have long G-rich telomeric overhangs at one end.

Genes and Development, 11(21), 2801–2809.

<https://doi.org/10.1101/gad.11.21.2801>

Yang, Z. W., Yang, S. H., Chen, L., Qu, J., Zhu, J., & Tang, Z. (2001). Comparison of blood counts in venous, fingertip and arterial blood and their measurement variation. *Clinical and Laboratory Haematology*, 23(3), 155–159.

<https://doi.org/10.1046/j.1365-2257.2001.00388.x>

Zahran, S., Snodgrass, J. G., Maranon, D. G., Upadhyay, C., Granger, D. a, & Bailey, S. M. (2015). Stress and telomere shortening among central Indian conservation refugees. *Proceedings of the National Academy of Sciences of the United States of America*, 112(9), E928-36. <https://doi.org/10.1073/pnas.1411902112>

Zalli, A., Carvalho, L. a, Lin, J., Hamer, M., Erusalimsky, J. D., Blackburn, E. H., & Steptoe, A. (2014). Shorter telomeres with high telomerase activity are associated with raised allostatic load and impoverished psychosocial resources. *Proceedings of the National Academy of Sciences of the United States of America*, 111(12), 4519–4524. <https://doi.org/10.1073/pnas.1322145111>