Title

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

Authors

Rej, Peter H
Bondy, Madison H
Lin, Jue
et al.

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

Peter H. Rej[1], **Madison H. Bondy**[2], **Jue Lin**[3], **Aric A. Prather**[4], **Brandon A. Kohrt**[5], **Carol M. Worthman**[2], **Dan T. A. Eisenberg**[1,6]

[1]Department of Anthropology, University of Washington, Seattle, Washington

[2]Department of Anthropology, Emory University, Atlanta, Georgia

[3]Department of Biochemistry and Biophysics, University of California, San Francisco, California

[4]Department of Psychiatry, University of California, San Francisco, California

[5]Department of Psychiatry and Behavioral Services, George Washington University, Washington, District of Columbia

[6]Center for Studies in Demography and Ecology, Seattle, Washington

Abstract

Objectives: Telomeres are the protective caps of chromosomes. They shorten with cell replication, age, and possibly environmental stimuli (eg, 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 statistical 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.

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 7 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, followed by saliva, and then buccal cells—possibly reflecting similarities/differences in cellular composition vs VWB.

Correspondence Peter H. Rej, Department of Anthropology University of Washington 314 Denny Hall, Box 353100 Seattle, WA 98195-3100. peterrej@uw.edu.

Conclusions: DBS DNA is the best proxy for VWB from the three minimally-invasively specimen types evaluated and can be used to expand TL research to diverse settings and populations.

1 | 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 (eg, 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, it is recommended that VWB samples be either refrigerated, frozen, or processed promptly in order to best 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 (eg, 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, Snodgrass, Maranon, Upadhyay, Granger, & Bailey, 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 5 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, de Almeida André, de 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. At least two studies have previously measured TL in DBS and matching VWB (Goldman et al., 2017; Stout et al., 2017). While Stout and colleagues concluded that DBS was superior to saliva (ie, had a stronger correlation with VWB), Goldman and colleagues reported 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.

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 additional validation (eg, statistical comparison to VWB TL and age). 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 (eg, 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 (eg, 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 (eg, 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 cohorts having 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 DBS extraction and TL assay 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, 2019; Eisenberg, 2016; Stout et al., 2017). We subsequently incorporated the published correlation values, along with those from our novel DBS work, into a meta-analysis.

2 | METHODS

2.1 | 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) was used to draw blood that was preserved on Whatman 903 protein saver cards (GE Healthcare, Chicago, IL). 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 5.5 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).

2.2 | 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 1500 to 1800 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 μL spots) onto Whatman FTA classic cards. DBS were stored in plastic containers with desiccant packs after drying at RT for at least 3 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, for DBS collected from each participant, 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.

2.3 | 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 3 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 DNA yields from the UW and Emory specimens were quantified by spectrophotometry using an Epoch Microplate Reader (BioTek, Winsooki, VT), and by Quant-iT PicoGreen dsDNA Assay Kit (Thermo-Fisher, Waltham, MA), respectively. This difference in DNA quantitation method may introduce downstream variation, as spectrophotometry measures impurities and degraded DNA in addition to double and single stranded-DNA, while PicoGreen exclusively binds to double-stranded DNA. 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.

2.4 | Telomere length measurement

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

2.4.1 | 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) using the same cycling profile reported by Eisenberg, Hayes, and Kuzawa (2012). We incorporated the primer pairs reported by Tackney, Cawthon, Coxworth, and Hawkes (2014) along with PowerUp SYBR Green master mix (Applied Biosystems, Waltham, MA). Our final reaction volume was 5 μL, which contained 1.0 μL DNA, 2.5 μL master mix, 1.1 μL primers, and 0.4 μL 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/μL) 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 3-fold serial dilution was used. Standard curve concentrations ranged from 5 ng/μl to 0.02 ng/μl. 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/μl 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. Efficiencies were automatically calculated by the software that accompanies our real-time thermocyclers/detection-systems using the following equation: efficiency $= -1 + 10^{(-1/\text{slope})}$.

2.4.2 | 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 μM 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, ~3-fold curve (ranging from 2.36 ng/μl to 0.01 ng/μl). 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.

2.5 | 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.

2.5.1 | 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: ICC(A, 1) = 0.86, 95% CI = 0.74, 0.92; ICC1(A, k) = 0.92, 95% CI = 0.85, 0.96.

2.5.2 | Emory—The T/S ratio of each positive control was divided by the average T/S ratio for the same DNA from 10 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: ICC(A, 1) = 0.97, 95% CI = 0.95, 0.99; ICC1(1, k) = 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.

2.6 | 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 (Barha et al., 2016; Chen et al., 2015; Dagan et al., 2018; Dismukes et al., 2016; Drury et al., 2014; Edmonds et al., 2015; Fillman et al., 2016; Finnicum et al., 2017; Gadalla et al., 2010; Goldman et al., 2017; Gotlib et al., 2014; Henje Blom et al., 2015; Hewakapuge et al., 2008; Kroenke et al., 2011; Lahnert et al., 2005; Lucas et al., 2018; Mitchel et al., 2014; Nelson et al., 2015; Powell et al., 2018; Stout et al., 2017, Wei et al., 2016; Woody et al., 2017; Zanet et al., 2013). We also included one of our recently published correlation values (Rej et al., 2019). A more detailed summary of our protocol is summarized in a PRISMA flow diagram (Figure S1).

2.7 | 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: (a) DBS TL vs blood TL; (b) Saliva TL vs blood TL; (c) BC TL vs blood TL; (d) DBS TL vs age; (e) Saliva TL vs age; (f) BC TL vs age. Before inclusion, we standardized all reported associations to Fisher z-transformed correlation coefficients (as in Pepper, Bateson, & Nettle, 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.

3 | RESULTS

3.1 | In-house matched specimen correlations

UW.—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 10 times more than published values (Goldman et al., 2017; Stout et al., 2017). 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).

Emory.—The average yields were: 369 ± 140 ng, 469 ± 198 ng, and 354 ± 96 for conditions A, B, and C, respectively. TL measurements were significantly different (F = 6.915, $P = 0.003$) between DBS handled under different conditions (A, B, and C). Per Tukey's HSD, condition A samples have significantly longer TL than conditions B and C, which can be observed in Figure 1B. DBS handled under condition C exhibited the strongest correlation with buffy coat TL in matched specimens, although all three have overlapping 95% confidence intervals (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). After running as simple linear regressions (ie, buffy coat TL—Condition A TL; buffy coat TL—Condition B TL; buffy coat TL—Condition C TL), we compared each model against one another using an ANOVA. This test 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.

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-Condition C vs frozen buffy coat samples: F = 6.915, $P = 0.003$).

3.2 | Description of dataset for meta- analyses

Including our new DBS data from the UW and Emory samples, our final dataset consisted of 24 studies. Of those, five reported external validation statistics for DBS, eleven for saliva, and nine for BCs. Stout et al. (2017) measured TL in VWB, DBS, and saliva, while our UW study and Goldman et al. (2017) report correlations with VWB and age. Of the 11 studies that measured salivary TL, only two did not use Oragene collection kits (Fillman, Shimizu-Furusawa, Ng, Parajuli, & Watanabe, 2016; Lahnert, 2005). All but one of these studies used a qPCR-based method to measure TL (Lahnert, 2005). Age ranges vary across the studies included, with most reporting substantial age variation. This variability in age is limitation of our TL vs age meta-analyses. Lower variation in age will tend to reduce the power to detect associations with age and make estimates of this correlation less reliable (although these estimates are expected to be unbiased). Table S3 provides details of the papers, including reported associations, DNA source, sample size, age range/variation, and method used to generate TL data.

3.3 | 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.

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 = -0.21, -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.

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).

4 | 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 correlation into our meta-analysis, along with results from 23 other studies identified by systematic literature review and one value from our own recently published 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 results 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: (a) stored for 7 days at RT and then shipped at RT; (b) stored at −30°C and shipped in dry ice; and (c) frozen immediately at −30°C and shipped in dry ice then partially extracted and stored for 7 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 6 years at −80°C, and undergoing multiple freeze-thaw cycles, the UW 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 can influence TL measurements (Cunningham et al., 2013; Dagnall et al., 2017; Denham et al., 2014; Raschenberger et al., 2016). While long-term freezing and repeated freeze-thawing are generally not recommended, there are conflicting results regarding their effects on TL measurements (Fazili, Sternberg, & Pfeiffer, 2012; Jenkins, Ye, & Silverman, 2012; Zanet et al., 2013). These factors along with the fact that the Emory blood spots were collected using venous blood applied to 903 cards vs blood collected from a finger prick for the UW (Zanet et al., 2013 observed significantly different TL measurements based on how blood was applied to a DBS card), at least in part, explain why the association of the Emory matched specimens is stronger than the UW.

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. While epithelial cells compose as much of 100% BC samples in adults, the cellular composition of saliva samples is highly heterogeneous, ranging from 5–84% (Theda et al., 2018; Thiede et al., 2000). Although white blood cells can contribute a large proportion of the DNA content in saliva, this heterogeneity likely contributes to our observed higher correlation of DBS TL 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 (eg, 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.

4.1 | Recommendations for Additional Best Practices in Measuring Telomeres

Several publications have reported different TL measurement techniques (eg, Cawthon, 2002, 2009; Kimura et al., 2010; Mender & Shay, 2016; O'Sullivan et al., 2005), their relative costs and benefits (eg, 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.

4.1.1 | Power—Previous analysts have noted that many published TL investigations are underpowered (eg, 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 $r = -0.09$ ($P = 0.04$);$0020 r = -0.03$ ($P = $ NS for their entire sample). Detecting an effect of -0.09 would require a sample size of 1007. 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 factors such as 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.

4.1.2 | 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 coefficient of variation (CV) measure of error is invalid for TL (Eisenberg, 2016; Nettle, Seeker, Nussey, Froy, & Bateson, 2019; 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. Specifically, we recommend reporting ICC(A,1), the ICC from the two-way, single measurement, absolute agreement, random effects model (McGraw & Wong, 1996).

4.1.3 | 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 trouble-shooting 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 is to assess TL is by measuring the same samples using multiple techniques (eg, 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., 2019), 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.

5 | 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

the method that closest approximates TL from whole blood. Saliva samples are among 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

REFERENCES

Aps JKM, Van Den Maagdenberg K, Delanghe JR, & Martens LC (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. 10.1016/S0009-8981(02)00062-1

Armanios M, & Blackburn EH (2012). The telomere syndromes. Nature Reviews Genetics, 13(10), 693–704. 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), e0163153. 10.1371/journal.pone.0163153 [PubMed: 27655116]

Barha CK, Hanna CW, Salvante KG, Wilson SL, Robinson WP, Altman RM, & Nepomnaschy PA (2016). Number of children and telomere length in women: A prospective, longitudinal evaluation. PLoS One, 11(1), 1–12. 10.1371/journal.pone.0146424

Bateson M, Eisenberg DTA, & Nettle D (2019). Controlling for baseline telomere length biases estimates of the effect of smoking on leukocyte telomere attrition in longitudinal studies. Royal Society Open Science, 6(10). 10.1098/rsos.190937

Bond MM, & Richards-Kortum RR (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. 10.1309/AJCP1L7DKMPCHPEH [PubMed: 26572995]

Bustin SA, Hellemans J, Pfaffl MW, Wittwer CT, Garson JA, Huggett J, … Nolan T (2009). The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. Clinical Chemistry, 55(4), 611–622. 10.1373/clinchem.2008.112797 [PubMed: 19246619]

Cawthon RM (2002). Telomere measurement by quantitative PCR. Nucleic Acids Research, 30(10), e47. 10.1093/nar/30.10.e47,47e,447 [PubMed: 12000852]

Cawthon RM (2009). Telomere length measurement by a novel monochrome multiplex quantitative PCR method. Nucleic Acids Research, 37(3), 1–7. 10.1093/nar/gkn1027 [PubMed: 19033363]

Cawthon RM, Smith KR, O'Brien E, Sivatchenko A, & Kerber RA (2003). Association between telomere length in blood and mortality in people aged 60 years or older. Lancet, 361(9355), 393–395. 10.1016/S0140-6736(03)12384-7 [PubMed: 12573379]

Ceffa S, Erba F, Brambilla D, Luhanga R, Armel Pépé M, Paturzo G, … Marazzi MC (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), 5609–5617. 10.5897/ajmr11.1469

Chatterjee D, Bhattacharjee P, Sau TJ, Das JK, Sarma N, Bandyopadhyay AK, … Giri AK (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. 10.1002/mc.22150 [PubMed: 24665044]

Chen X, Velez JC, Barbosa C, Pepper M, Andrade A, Stoner L, … Williams MA (2015). Smoking and perceived stress in relation to short salivary telomere length among caregivers of children with disabilities. Stress, 18(1), 20–28. 10.3109/10253890.2014.969704 [PubMed: 25256607]

Cong Y, Wright WE, & Shay JW (2002). Human telomerase and its regulation human telomerase and its regulation. Micro- biology and Molecular Biology Reviews, 66(3), 407–425. 10.1128/MMBR.66.3.407

Cunningham JM, Johnson RA, Litzelman K, Skinner HG, Seo S, Engelman CD, … Boardman LA (2013). Telomere length varies by DNA extraction method: Implications for epidemiologic research. Cancer Epidemiology Biomarkers and Pre- vention, 22(11), 2047–2054. 10.1158/1055-9965.EPI-13-0409

Dagan O, Asok A, Steele H, Steele M, & Bernard K (2018). Attachment security moderates the link between adverse childhood experiences and cellular aging. Development and Psycho- pathology, 30(4), 1211–1223. 10.1017/S0954579417001705

Dagnall CL, Hicks B, Teshome K, Hutchinson AA, Gadalla SM, Khincha PP, … Savage SA (2017). Effect of pre-analytic variables on the reproducibility of qPCR relative telomere length measurement. PLoS One, 12(9), e0184098. 10.1371/journal.pone.0184098 [PubMed: 28886139]

Denham J, Marques FZ, & Charchar FJ (2014). Leukocyte telomere length variation due to DNA extraction method. BMC. Res.Notes, 7, 877-. [PubMed: 25475541]

Dismukes AR, Meyer VJ, Shirtcliff EA, Theall KP, Esteves KC, & Drury SS (2016). Diurnal and stress-reactive dehydroepiandrosterone levels and telomere length in youth. Endocrine Connections, 5(3), 107–114. 10.1530/EC-16-0007 [PubMed: 27221260]

Edmonds GW, Côté HCF, & Hampson SE (2015). Childhood conscientiousness and leukocyte telomere length 40 years later in adult women—Preliminary findings of a prospective association. PLoS One, 10(7), e0134077. 10.1371/journal.pone.0134077 [PubMed: 26218760]

Eisenberg DTA, Hayes MG, & Kuzawa CW (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. 10.1073/pnas.1202092109

Eisenberg DTA, Salpea KD, Kuzawa CW, Hayes MG, & Humphries SE (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. 10.1002/ajhb.21126 [PubMed: 21319253]

Eisenberg DTA (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. 10.1093/ije/dyw191

Eisenberg DTA, Borja JB, Hayes MG, & Kuzawa CW (2017). Early life infection, but not breastfeeding, predicts adult blood telomere lengths in The Philippines. American Journal of Human Biology, 29(4), e22962. 10.1002/ajhb.22962

Eisenberg DTA, Lee NR, Rej PH, Hayes MG, & Kuzawa CW (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), 20190800. 10.1098/rspb.2019.0800

Epel ES, & Prather AA (2018). Stress, telomeres, and psychopathology: Toward a deeper understanding of a triad of early aging. Annual Review of Clinical Psychology, 14, 371–397 10.1146/annurev-clinpsy-032816-045054

Farrukh S, Baig S, Hussain R, Shahid A, & Khan ST (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 S, Baig S, Hussain R, & Lucky MH (2019). Association of Cord Blood Telomere Biology with Mother's education. International Journal of Biochemistry Research & Review, 24(3), 1–9. 10.9734/ijbcrr/2018/46192

Fazili Z, Sternberg MR, & Pfeiffer CM (2012). Assessing the influence of 5,10-methylenetetrahydrofolate reductase polymorphism on folate stability during long-term frozen storage, thawing, and repeated freeze/thawing of whole blood. Clinica Chimica Acta; International Journal of Clinical Chemistry, 413, 966–972. 10.1016/j.cca.2012.02.003 [PubMed: 22342879]

Fillman T, Shimizu-Furusawa H, Ng CFS, Parajuli RP, & Watanabe C (2016). Association of cadmium and arsenic exposure with salivary telomere length in adolescents in Terai, Nepal. Environmental Research, 149, 8–14. 10.1016/j.envres.2016.04.037 [PubMed: 27155138]

Finnicum CT, Dolan CV, Willemsen G, Weber ZM, Petersen JL, Beck JJ, … Ehli EA (2017). Relative telomere repeat mass in buccal and leukocyte-derived DNA. PLoS One, 12(1), e0170765. 10.1371/journal.pone.0170765 [PubMed: 28125671]

Gadalla SM, Cawthon R, Giri N, Alter BP, & Savage SA (2010). Telomere length in blood, buccal cells, and fibroblasts from patients with inherited bone marrow failure syndromes. Aging, 2(11), 867–874. 10.18632/aging.100235 [PubMed: 21113082]

Gautam K, Pedersen GA, Wahid SS, & Kohrt BA (2019). Ethical considerations for conducting cross-cultural biological psychiatry and prevention research on depression among adolescents in low- and middle-income countries. 10.1016/bs.dnb.2019.05.001

Goldman EA, Eick GN, Compton D, Kowal P, Snodgrass JJ, Eisenberg DTA, & Sterner KN (2017). Evaluating minimally invasive sample collection methods for telomere length measurement. American Journal of Human Biology, 30(September 2016), e23062. 10.1002/ajhb.23062

Gotlib IH, LeMoult J, Colich NL, Foland-Ross LC, Hallmayer J, Joormann J, … Wolkowitz OM (2014). Telomere length and cortisol reactivity in children of depressed mothers. Molecular Psychiatry, 20(5), 615–620. 10.1038/mp.2014.119 [PubMed: 25266121]

Greider CW, & Blackburn EH (1987). The telomere terminal transferase of tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. Cell, 51(6), 887–898. 10.1016/0092-8674(87)90576-9 [PubMed: 3319189]

Henje Blom E, Han LKM, Connolly CG, Ho TC, Lin J, LeWinn KZ, … Yang TT (2015). Peripheral telomere length and hippocampal volume in adolescents with major depressive disorder. Translational Psychiatry, 5, e676. [PubMed: 26556285]

Hewakapuge S, van Oorschot RAH, Lewandowski P, & Baindur-Hudson S (2008). Investigation of telomere lengths measurement by quantitative real-time PCR to predict age. Legal Medicine, 10(5), 236–242. 10.1016/j.legalmed.2008.01.007 [PubMed: 18378481]

Jenkins EC, Ye L, & Silverman WP (2012). Does the cryogenic freezing process cause shorter telomeres? Cryobiology, 65(1), 72–73. 10.1016/j.cryobiol.2012.03.005 [PubMed: 22465657]

Jiménez KM, Pereira-Morales AJ, Adan A, & Forero DA (2019). Telomere length and childhood trauma in Colombians with depressive symptoms. Revista Brasileira de Psiquiatria, 41 (3), 194–198. 10.1590/1516-4446-2018-0027 [PubMed: 30328966]

Kimura M, Stone RC, Hunt SC, 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. 10.1038/nprot.2010.124 [PubMed: 21085125]

Kroenke CH, Epel E, Adler N, Bush NR, Obradovi J, Lin J, … Boyce WT (2011). Autonomic and adrenocortical reactivity and buccal cell telomere length in kindergarten children. Psychosomatic Medicine, 73(7), 533–540. 10.1097/PSY.0b013e318229acfc [PubMed: 21873585]

Lahnert P (2005). An improved method for determining telomere length and its use in assessing age in blood and saliva. Gerontol- ogy, 51(5), 352–356. 10.1159/000086374

Lai TP, Wright WE, & Shay JW (2018). Comparison of telomere length measurement methods. Philosophical Transactions of the Royal Society B: Biological Sciences, 373(1741), 20160451. 10.1098/rstb.2016.0451

Lin J, Cheon J, Brown R, Coccia M, Puterman E, Aschbacher K, … Blackburn EH (2016). Systematic and cell type-specific telomere length changes in subsets of lymphocytes. Journal of Immunology Research, 2016, 1–9. 10.1155/2016/5371050

Lin J, Smith DL, Esteves K, & Drury S (2019). Telomere length measurement by qPCR – Summary of critical factors and recommendations for assay design. Psychoneuroendocrinology, 99, 271–278. 10.1016/j.psyneuen.2018.10.005 [PubMed: 30343983]

Lingner J, Cooper JP, & Cech TR (1995). Telomerase and DNA end replication: No longer a lagging strand problem? Science, 269 (5230), 1533–1534. 10.1126/science.7545310 [PubMed: 7545310]

Lucas T, Woerner J, Pierce J, Granger DA, Lin J, Epel ES, … Lumley MA (2018). Justice for all? Beliefs about justice for self and others and telomere length in African Americans. Cultural Diversity and Ethnic Minority Psychology, 24(4), 498–509. 10.1037/cdp0000212 [PubMed: 30058830]

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. 10.1002/da.20903 [PubMed: 22065550]

Mathur MB, Epel E, Kind S, Desai M, Parks CG, Sandler DP, & 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. 10.1016/j.bbi.2016.02.002

McDade TW, Williams S, & Snodgrass JJ (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. 10.1353/dem.2007.0038 [PubMed: 18232218]

McGraw KO, & Wong SP (1996). Forming inferences about some intraclass correlation coefficients. Psychological Methods, 1(1), 30.

Mender I, & Shay J (2016). Telomere restriction fragment (TRF) analysis. Bio-Protocol, 5(22), e1658. 10.21769/bioprotoc.1658

Mitchell C, Hobcraft J, McLanahan SS, Siegel SR, Berg A, Brooks-Gunn J, … Notterman D (2014). Social disadvantage, genetic sensitivity, and children's telomere length. Proceedings of the National Academy of Sciences, 111(16), 5944–5949. 10.1073/pnas.1404293111

Montpetit AJ, Alhareeri AA, Montpetit M, Starkweather AR, Elmore LW, Filler K, … Jackson-Cook CK (2014). Telomere length: A review of methods for measurement. Nursing Research, 63(4), 289–299. 10.1097/nnr.0000000000000037 [PubMed: 24977726]

Nedel F, de Almeida André D, de Oliveira IO, Tarquinio SBC, & Demarco FF (2009). Buccal cells submitted to three different storage conditions before DNA extraction. Journal of Applied Oral Science, 17(2), 113–115. 10.1590/s1678-77572009000200008 [PubMed: 19274396]

Nelson CA, Varcin KJ, Coman NK, Devivo I, & Tager-Flusberg H (2015). Shortened telomeres in families with a propensity to autism. Journal of the American Academy of Child and Adolescent Psychiatry, 54(7), 588–594. 10.1016/j.jaac.2015.04.006 [PubMed: 26088664]

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), e0216118. https://doi.org/0.1371/journal.pone.0216118 [PubMed: 31042766]

O'Sullivan JN, Finley JC, Risques R, Shen W-T, Gollahon KA, & Rabinovitch PS (2005). Quantitative fluorescence in situ hybridization (QFISH) of telomere lengths in tissue and cells. Current Protocols in Cytometry, 33, 12–16. 10.1002/0471142956.cy1206s33

Oliveira BS, Zunzunegui MV, Quinlan J, Batistuzzo De Medeiros SR, Thomasini RL, & Guerra RO (2018). Lifecourse adversity and telomere length in older women from Northeast Brazil. Rejuvenation Research, 21(4), 294–303. 10.1089/rej.2017.1937 [PubMed: 28482745]

Palm W, & de Lange T (2008). How Shelterin protects mammalian telomeres. Annual Review of Genetics, 42(1), 301–334. 10.1146/annurev.genet.41.110306.130350

Pathai S, Lawn SD, Gilbert CE, McGuinness D, McGlynn L, Weiss HA, … Shiels PG (2013). Accelerated biological ageing in HIV-infected individuals in South Africa: A case-control study. Aids, 27(15), 2375–2384. 10.1097/QAD.0b013e328363bf7f [PubMed: 23751258]

Pepper GV, 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), 180744. 10.1098/rsos.180744 [PubMed: 30225068]

Powell TR, Dima D, Frangou S, & Breen G (2018). Telomere length and bipolar disorder. Neuropsychopharmacology, 43(2), 445–453. 10.1038/npp.2017.125 [PubMed: 28621334]

Quinlan J, Tu MT, Langlois EV, Kapoor M, Ziegler D, Fahmi H, & Zunzunegui MV (2014). Protocol for a systematic review of the association between chronic stress during the life course and telomere length. Systematic Reviews, 3, 40. 10.1186/2046-4053-3-40 [PubMed: 24886862]

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, 25398. 10.1038/srep25398 [PubMed: 27138987]

Rej PH, HEAT Steering Committee, Gravlee CC, & Mulligan CM (2019). Shortened telomere length is associated with experiences of unfair treatment attributed to race in African Americans living in Tallahassee, Florida. American Journal of Human Biology, e23375. 10.1002/ajhb.23375 [PubMed: 31867825]

Rej PH, Tennyson RL, Lee NR, & Eisenberg DTA (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. 10.1016/J.PSYNEUEN.2019.01.019 [PubMed: 30711895]

Rode L, Nordestgaard BG, & Bojesen SE (2015). Peripheral blood leukocyte telomere length and mortality among 64 637 individuals from the general population. Journal of the National Cancer Institute, 107(6), djv074. 10.1093/jnci/djv074 [PubMed: 25862531]

Serrano AL, & Andrés V (2004). Telomeres and cardiovascular disease: Does size matter? Circulation Research, 94(5), 575–584. 10.1161/01.RES.0000122141.18795.9C [PubMed: 15031270]

Stout SA, Lin J, Hernandez N, Davis EP, Blackburn E, Carroll JE, & Glynn LM (2017). Validation of minimally-invasive sample collection methods for measurement of telomere length. Frontiers in Aging Neuroscience, 9, 397. 10.3389/fnagi.2017.00397 [PubMed: 29270121]

Tackney J, Cawthon RM, Coxworth JE, & Hawkes K (2014). Blood cell telomere lengths and shortening rates of chimpanzee and human females. American Journal of Human Biology, 26 (4), 452–460. 10.1002/ajhb.22538 [PubMed: 24633909]

R Development Core Team (2019). R: A language and environment for statistical computing. Retrieved from http://www.gbif.org/resource/81287.

Tennyson RL, Gettler LT, Kuzawa CW, Hayes MG, Agustin SS, & Eisenberg DTA (2018). Lifetime socioeconomic status and early life microbial environments predict adult blood telomere length in the Philippines. American Journal of Human Biology, 30, e23145. 10.1002/ajhb.23145 [PubMed: 30129279]

Theda C, Hwang SH, Czajko A, Loke YJ, Leong P, & Craig JM (2018). Quantitation of the cellular content of saliva and buccal swab samples. Scientific Reports, 8(1), 6944. 10.1038/s41598-018-25311-0 [PubMed: 29720614]

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. 10.1038/sj.bmt.1702170

Verhulst S, Susser E, Factor-Litvak PR, Simons M, Benetos A, Steenstrup T, … Aviv A (2016). Response to: Reliability and validity of telomere length measurements. Inter- national Journal of Epidemiology, 45(4), 1298–1301. 10.1093/ije/dyw194

Verhulst S, Susser E, Factor-Litvak PR, Simons MJP, Benetos A, Steenstrup T, … Aviv A (2015). Commentary: The reliability of telomere length measurements. International Journal of Epidemiology, 44(5), 1683–1686. 10.1093/ije/dyv166 [PubMed: 26403812]

Von Känel R, Malan NT, 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. 10.1097/PSY.0000000000000123 [PubMed: 25469684]

Wang Q, Zhan Y, Pedersen NL, Fang F, & Hägg S (2018). Telomere length and all-cause mortality: A meta-analysis. Ageing Research Reviews, 48, 11–20. 10.1016/j.arr.2018.09.002 [PubMed: 30254001]

Wei YB, Martinsson L, Liu JJ, Forsell Y, Schalling M, Backlund L, & Lavebratt C (2016). HTERT genetic variation in depression. Journal of Affective Disorders, 189, 62–69. 10.1016/j.jad.2015.09.025 [PubMed: 26406970]

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

Woo JS, & Lu DY (2019). Procurement, transportation, and storage of saliva, buccal swab, and oral wash specimens. Methods in Molecular Biology, 1897, 99–105. 10.1007/978-1-4939-8935-5_10 [PubMed: 30539438]

Woody A, Hamilton K, Livitz IE, Figueroa WS, & Zoccola PM (2017). Buccal telomere length and its associations with cortisol, heart rate variability, heart rate, and blood pressure responses to an acute social evaluative stressor in college students. Stress, 20(3), 249–257. 10.1080/10253890.2017.1328494 [PubMed: 28482730]

Wright WE, Tesmer VM, Huffman KE, Levene SD, & Shay JW (1997). Normal human chromosomes have long G-rich telomeric overhangs at one end. Genes and Development, 11(21), 2801–2809. 10.1101/gad.11.21.2801 [PubMed: 9353250]

Yang ZW, Yang SH, 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. 10.1046/j.1365-2257.2001.00388.x [PubMed: 11553055]

Zahran S, Snodgrass JG, Maranon DG, Upadhyay C, Granger D, & Bailey SM (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–E936. 10.1073/pnas.1411902112 [PubMed: 25730846]

Zalli A, Carvalho L, Lin J, Hamer M, Erusalimsky JD, Blackburn EH, & 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. 10.1073/pnas.1322145111 [PubMed: 24616496]

Zanet DL, Saberi S, Oliveira L, Sattha B, Gadawski I, & Côté HCF (2013). Blood and dried blood spot telomere length measurement by qPCR: Assay considerations. PLoS One, 8(2), e57787. 10.1371/journal.pone.0057787 [PubMed: 23451268]
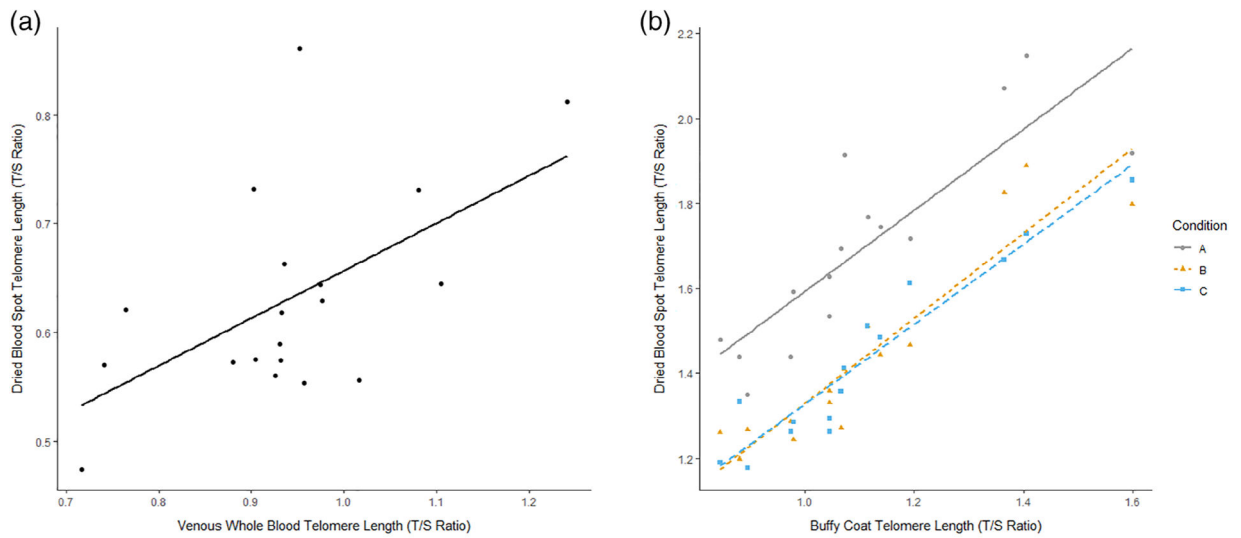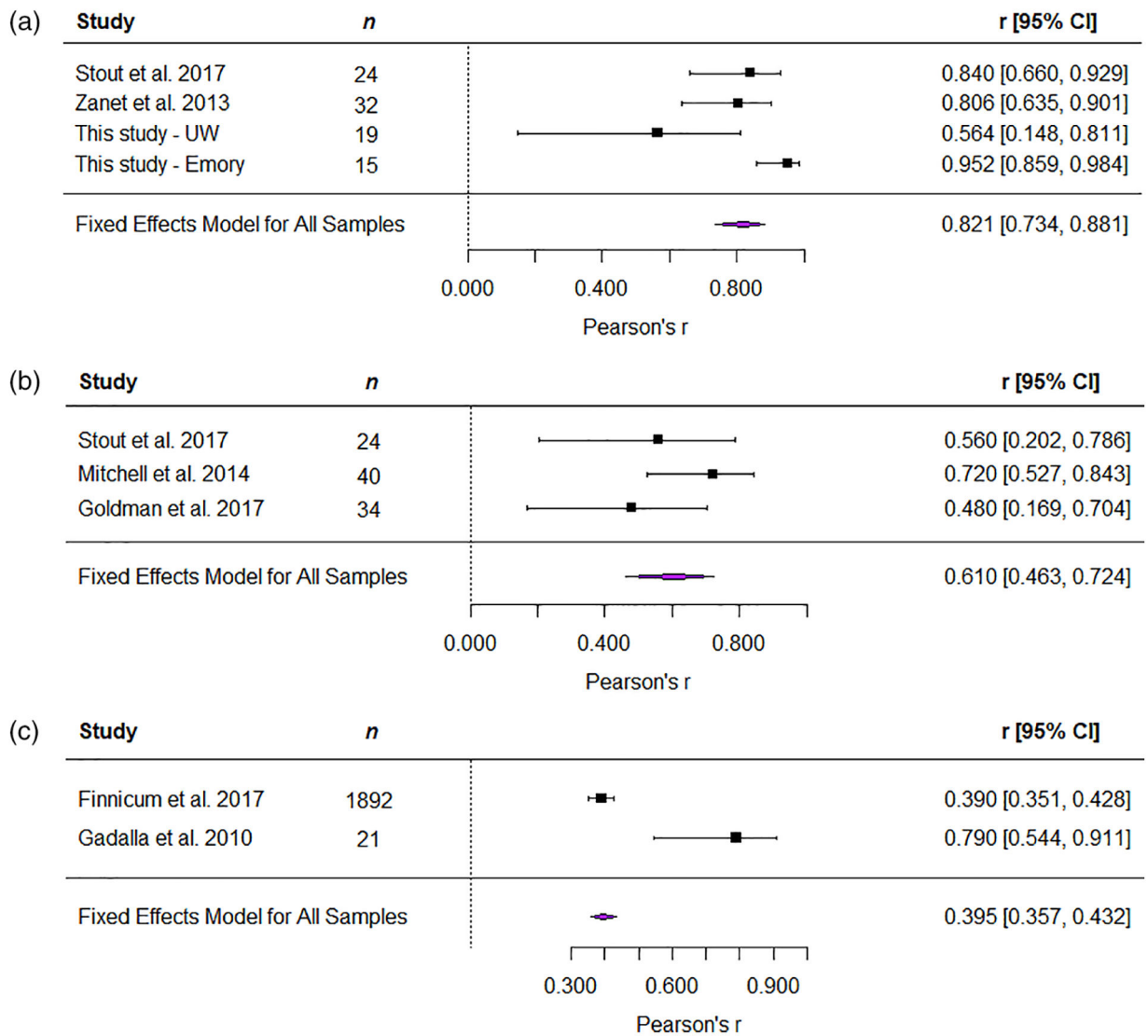
FIGURE 1.

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 (gray solid line is Condition A: $r = 0.84$; orange short dashed line is Condition B: $r = 0.93$; blue long dashed line is Condition C: $r = 0.95$)

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

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

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.