TELOMERE LENGTH: A BIOMARKER OF AGING IN NEURODEVELOPMENT AND NEUROPSYCHOLOGICAL PERFORMANCE FOR MINORITIZED POPULATIONS

AN ABSTRACT

SUBMITTED ON THE TWENTY-SEVENTH OF FEBRUARY 2023

TO THE TULANE UNIVERSITY BRAIN INSTITUTE

IN PARTIAL FULFILLMENT OF THE REQUIRMEENTS

OF THE SCHOOL OF SCIENCE AND ENGINEERING

OF TULANE UNIVERSITY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY BY

LAUREN WENDELLE YOWELUNH MCLESTER-DAVIS

Enrolled Member of the Oneida Nation of Wisconsin

APPROVED: _____

STACY S. DRURY, Ph.D., M.D. (Advisor)

_____

MAUREEN Y. LICHTVELD, M.D, M.P.H.

_____

CAREY E. GLEASON, Ph.D., M.S.

_____

RICARDO MOSTANY, Ph.D.

_____

DEMETRIUS M. MARAGANORE, M.D., F.A.A.N.

ABSTRACT

The goal of this dissertation was to investigate the utility of telomere length (TL) as a biomarker of neurodevelopment and neuropsychological performance for minoritized populations. Telomeres are the DNA and protein cap structure found at the end of all eukaryotic chromosomes which protect linear DNA from damage and preserve chromosome stability. The progressive shortening of the telomeric DNA sequence of 5'-TTAGGG-3' repeats occurs in dividing cells because of incomplete lagging-strand DNA synthesis, DNA damage, and other factors. Previous research has associated TL with various exposures, biological processes, and health outcomes. In this dissertation, the history, precision, and reliability of TL is first investigated. Following the examination of TL measurement, this dissertation applies TL as a biomarker of neurodevelopment for the diverse population sample in the Caribbean Consortium for Research in Environmental and Occupational Health Cohort in Suriname. This dissertation then applies TL as a biomarker of neuropsychological performance in American Indians / Alaska Natives and African Americans in the Wisconsin Alzheimer's Disease Research Center studies. The outcomes of this dissertation include reporting recommendations for quantitative polymerase chain reaction-based TL measurement, increasing the reproducibility of multiplex monochrome quantitative polymerase chain reaction TL measurements via visual explanation, confirmation that TL measurements within an individual from different tissues meta-analytically correlate with one another, the association of infant TL and prenatal environmental exposures with neurodevelopmental performance, an initial correlation of TL from DNA of cerebrospinal fluid to DNA of whole blood, and the first report on TL predicting neuropsychological performance in American Indians / Alaska Natives and African Americans. This dissertation expands on current literature of precision measurement of TL, TL as a biomarker of neurodevelopment, and TL as a

biomarker for neuropsychological aging. Taken together, these dissertation experiments investigate neuroscience inquires across the lifespan for minoritized populations. This dissertation expands on the sparse neuroscience literature available for these minoritized populations who are underrepresented in scientific research enrollment, retention, and consideration. In conclusion, TL can be utilized as a biomarker of neurodevelopment and neuropsychological performance across the lifespan, with inclusion of minoritized populations a priority in future work.

TELOMERE LENGTH: A BIOMARKER OF AGING IN NEURODEVELOPMENT AND NEUROPSYCHOLOGICAL PERFORMANCE FOR MINORITIZED POPULATIONS

A DISSERTATION

SUBMITTED ON THE TWENTY-SEVENTH OF FEBRUARY 2023

TO THE TULANE UNIVERSITY BRAIN INSTITUTE

IN PARTIAL FULFILLMENT OF THE REQUIRMEENTS

OF THE SCHOOL OF SCIENCE AND ENGINEERING

OF TULANE UNIVERSITY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY BY

LAUREN WENDELLE YOWELUNH MCLESTER-DAVIS

Enrolled Member of the Oneida Nation of Wisconsin

APPROVED: _____

STACY S. DRURY, Ph.D., M.D. (Advisor)

_____

MAUREEN Y. LICHTVELD, M.D, M.P.H.

_____

CAREY E. GLEASON, Ph.D., M.S.

_____

RICARDO MOSTANY, Ph.D.

_____

DEMETRIUS M. MARAGANORE, M.D., F.A.A.N.

ACKNOWLEDGEMENTS

None of this would have been possible without my first teachers, my grandparents Thelma and Wendell McLester, nor without my parents, Joan McLester and David Davis. I am very grateful for the support of my fiancé, W. Cannon Matthews, III, my friends, Drs. Baumgartner, Sawyer, Popescu, Davidson, Zuelsdorff, Singh, Bull, Lewis, Black, Hastings, Harrison, King, Febbo, Simcox, Menka, Bowman-Farrell, Monteau, Murie-Mazariegos, Suchy-Dicey, Lee Vaasjo Munoz, Cheyenne Kirkman-Reuben Thomas, Amy Katz, Miranda Horn, Kyle Conniff, McKenzie White, Dom Acri, Alden Grasse, Trev Roy, Nick Felipe, Amy DeGraff-Castro, Leila Kataria, Pedro Estrada, others, my godmothers Peggy and Chris, adoptive aunt Lora Ann, extended family, and previous trainees, Noelle Martin, Joshua Siebeneicher, and Madeline Magruder, without whom I would not have completed this dissertation. I am thankful for my mentors, Drs. Schrader, Hetzler, Hall, Debbert, Shankar, Lewis, Abdoel Wahid, Wickliffe, Parker, Hogan, Alisch, Cronin, Daniel, Piasecki, De Stasio, and my dissertation committee including principal investigator, Dr. Drury, and Drs. Gleason, Mostany, Litchveld, and Maraganore, who have made these research studies possible.

I would not have made this dissertation and research contributions without the education and encouragement from numerous teachers, educational staff, lecturers, and professors throughout my learning journey to date, including Mrs. Lueck, Mrs. Bryce, Forrest Brooks, Principals Sanders and Gerke, Ms. Buhl, Mr. Rukamp, Mr. Bobinski, Mr. Konshak, Ms. Frank, Mr. Erickson, Mr. Krhin, and all informal teachers who have contributed to the knowledge of this dissertation. I am indebted to the many individuals who encouraged me to complete this journey, including the therapists I have seen during my doctoral work, Gersovitz, Leggett, and Whiting, and the communities I gained through the Summer Internship for Indigenous Peoples in

Genomics, the Johns Hopkins Center for Indigenous Health Institutes, and the Society for the Advancement of Chicanos / Hispanics and Native Americans in Science. I also appreciate all my scientific collaborators who contributed to these findings and all funding sources that supported this work, including the Ronald E. McNair Scholar Summer Research Experience Funding, Louisiana Board of Regents Graduate Student Fellowship, National Institute on Aging Diversity Supplement, Meta-Analysis Systematic Review Support Program, and Urban Indian Health Institute / Seattle Indian Health Board Decolonizing Data Grant.

I am extremely grateful for the community to which I was born into, I appreciate the Oneida Nation of Wisconsin for financially supporting my educational journey, and I aspire that this dissertation benefits my community and others. I honor the many family members and friends who have passed away during my dissertation and the COVID-19 pandemic for their support of my work. I dedicate this work to my grandparents, Thelma (89 years old) and Wendell (92 years old) McLester, and to any Indigenous child that they may intrinsically know their contributions to science and the world are innumerable and imperative for our future.

TABLE OF CONTENTS

Abstract……**ii**

Acknowledgements...……..**vi**

Table of Contents...……..**viii**

List of Tables……...**x**

List of Figures……**xii**

Introduction to the Dissertation.……..**1**

Chapter 1: Method comparison studies of telomere length measurement using qPCR approaches: A critical appraisal of the literature.……**7**

Chapter 2: A review and meta-analysis: Cross-tissue telomere length correlations in healthy humans……...**51**

Chapter 3: Precisions MMqPCR Telomere Length Measurement Methodology..……**101**

Chapter 4: Validity, Reliability, and Transcultural Adaptations of the Bayley Scales of Infant and Toddler Development (BSIID-III-NL) for Children in Suriname.……...**136**

Chapter 5: Prenatal Environmental Exposures and Cellular Aging Impact the Social-Emotional Development of Surinamese Infants……**182**

Chapter 6: Telomere length predicts neuropsychological performance in middle-aged and older participants of underrepresented populations from the University of Wisconsin Alzheimer's Disease Research Center……..**216**

Chapter 7: The correlation of telomere length from whole blood and cerebrospinal fluid circulating white blood cells from participants of the University of Wisconsin Alzheimer's Disease Research Center……..**235**

Discussion of the Dissertation……**247**

List of Appendices and Supplemental Materials…….**257**

List of References……**259**

LIST OF TABLES

Chapter 1.

Table 1. Characteristics of included papers.……..**18**

Table 2. Results of reporting guidelines analysis for TRN and Morinha Reporting Guidelines...**19**

Table 3. Item-specific reporting results of the TRN Reporting Guidelines.……..**20**

Chapter 2.

Table 1. Characteristics of Studies in the Qualitative Review.……..**77**

Table 2. Characteristic Differences Between Studies in the Qualitative Review versus Studies in the Meta-Analysis.……**90**

Table 3. Subgroup Analysis of Correlations by Sample Size Grouping.……...**91**

Table 4. Subgroup Analysis of Correlations by Telomere Length Measurement Methodology...**91**

Table 5. Sensitivity Analysis of Correlations by Blood versus Non-Blood Sample Sources……**91**

Table 6. Sensitivity Analysis of Correlations by Peripherally Available versus Surgically Obtainable Sample Sources.……..**92**

Chapter 3.

Table 1. Table of Specific Materials and Equipment.……**104**

Table 2. Telomere and Single Copy Gene Oligonucleotide Primer Sequences……..**108**

Table 3. Final Volumes and Concentrations of Reagents……**112**

Table 4. Process for Filling Plates……...**115**

Table 5. Organization of Samples and Control Standard in Plates……..**116**

Table 6. Spectrophotometer and Qubit DNA Quality Metrics for Sample Results…….**126**

Table 7. Mean T/S Ratios and Z-scored Telomere Length for Sample Results……..**130**

Chapter 4.

Table 1. Demographics of Study Participants…….……..**142**

Table 2. Reliability of Each Subscale……..**155**

Chapter 5.

Table 1. Maternal Demographics…….**185**

Table 2. Infant Demographics……..**188**

Table 3. Significant Differences of Mean Component Levels by CCREOH Region……..**191**

Table 4. Multivariable Linear Regression Results……...**196**

Chapter 6.

Table 1. Participant Demographics……..**218**

Table 2. Multivariable Linear Regression Models with Significant Predictors' Betas and Probability Values……**226**

Chapter 7.

Not Applicable.

LIST OF FIGURES

Chapter 1.

Figure 1. Telomere publications 1988-2019.……...…**8**

Figure 2. Novel telomere methods developed from 1988 to 2020……..**9**

Figure 3. PRISMA flow diagram of literature search and article selection for inclusion…….**15**

Figure 4. Weighted average correlation (Pearson's r) and 95% CI for each method compared with qPCR, MMqPCR, or aTL……..**23**

Figure 5. The sample size required to test effect sizes of 150, 200 and 300 bp with a t-test with a power of 0.9, as a function of measurement error as expressed in the ICC…….…..**25**

Figure 6. Power to detect a 33% change of telomere shortening rate, up or down, with $p<0.05$ relative to a baseline shortening rate of 25 bp/year……...**27**

Chapter 2.

Figure 1. PRISMA Flow Diagram…….....**60**

Figure 2. Forest Plot of Correlation Coefficients Between All Tissue Telomere Lengths by Study and Year of Publication…….....**66**

Figure 3. Funnel Plot of Effect Sizes for Correlations……..**69**

Chapter 3.

Figure 1. Multiplex of Telomere and Single Copy Gene Amplicons……..**102**

Figure 2. Example of Telomere Length Data Produced in the MMqPCR Assay……**103**

Figure 3. Example Labelling for P1 and P2 Plates……..**114**

Figure 4. Thermocycling Profile of the MMqPCR Assay……...**118**

Figure 5. Example Standard Curve for the Telomere Amplicon in P1 Plate……...**127**

Figure 6. Example Standard Curve for the Single Copy Gene Amplicon in P2 Plate…….**128**

Figure 7. Example Melt Curve with Individual Peaks…….**128**

Figure 8. Example Intraclass Correlation Coefficient Sheet and Calculation…….**131**

Figure 9. Correlation of Telomere Lengths Across DNA Extractions for Sample Results…….**132**

Chapter 4.

Figure 1. Mean raw score and item clusters of the receptive communication subscale increase with age……**148**

Figure 2. Mean raw score and item clusters of the expressive communication subscale increase with age.……...**149**

Figure 3. Mean raw score and item clusters of the gross motor subscale increase with age…...**151**

Figure 4. Mean raw score and item clusters of the fine motor subscale increase with age…….**152**

Figure 5. Mean raw score and item clusters of the cognitive subscale increase with age……...**154**

Chapter 5.

Figure 1. Caribbean Cohort for Research in Environmental and Occupational Health Flow Diagram for the Current Study……**184**

Figure 2. Comparison of Mean Telomere Length by Ethnicity and Age……**194**

Figure 3. Overall Effect of the Mixture and Individual Components at Different Quantiles on Social-Emotional Development……...**197**

Figure 4. Lack of Interactive Effects Between Individual Components on Social-Emotional Development……**199**

Chapter 6.

Figure 1. Box plot of telomere length for each race and ethnicity category and the distribution of telomere length for the sample……**223**

Figure 2. Scatter plots with linear fit lines of telomere length with the Rey Auditory Verbal Learning Test sum of Trials 1-5, the Trails Making Test Trails A time and the Trails Making Test Trails B time…….**224**

Figure 3. Box plots of telomere length for each cognitive status and Apolipoprotein E genotypes…….**229**

Chapter 7.

Figure 1. Positive correlation of cerebrospinal fluid telomere length and whole blood telomere length within individuals……..**237**

Figure 2. Negative correlation of whole blood telomere length and chronological age……..**238**

INTRODUCTION TO THE DISSERTATION

Throughout an organism's lifespan, cells rely on telomeres to ensure genome stability throughout the life cycle of the cell. Telomeres are the vital ends of chromosomes that protect linear DNA from damage and preserve the stability of the chromosome. The repetitive DNA sequence of 5'-TTAGGG-3' telomere repeats are highly conserved across eukaryotic species.[1] This sequence progressively shortens across an organism's lifespan with each cell replication due to incomplete lagging-strand DNA synthesis, DNA damage, and other factors, until a certain limit, then cellular arrest and senescence, genomic instability, or cellular death can occur.[2] Telomeres also serve important roles in creating heterochromatin at chromosomal ends and the position of the repetitive DNA sequence which loops on itself has regulatory effects on other gene expression within a cell.[2,3]

A functional telomere distinguishes the chromosome end from random chromosomal breaks to ensure that DNA damage checkpoint signaling is not activated, as well as prevents chromosomal ends from fusing with one another or other strands of DNA. Telomere sequences are maintained at the ends of chromosomes through telomerase, an essential enzyme in this cellular maintenance system.[3] Telomerase is an intracellular ribonucleoprotein that can maintain or elongate telomere sequences for specific purposes. Example purposes are often seen in an organism's germ cells, for instance, spermatogonia possess high levels of telomerase to ensure continual production of sperm across an adult male's life.[4]

Telomeres are also vital to any dividing human tissue to mitigate disease and increase longevity, as an unfunctional telomere and telomerase system can be a molecular prelude to cancer and age-related diseases. For example, the length of human telomeres in red bone marrow cells, which proliferate across the lifespan to produce blood cells, are maintained by telomerase

to ensure enough 5'-TTAGGG-3' repeats for blood production.[5] Conversely, decreased levels of telomerase and shortened telomeres in hematopoietic stem cells can results in aplastic anemias and other disease syndromes. Furthermore, with the increased aging population, the public is concerned with aging-related diseases like Alzheimer's disease and related dementias (ADRD).[6] While a diverse set of disease pathologies, multiple hypotheses of ADRD etiology include dysregulation of environmental and genetic factor effects on neurological aging. Telomeres are a part of regulating environmental and genetic factors within cells and are therefore hypothesized to be involved in ADRD etiology. In support of this hypothesis, recent meta-analytical findings show shorter telomeres are associated with ADRD diagnosis.[7]

Given the important role of telomeres in the cell cycle and various diseases, many have hypothesized hijacking the telomere and telomerase system to live forever. Unfortunately, now is not the time of telomeric intervention given the probability of oncogenic side-effects. However, measuring telomere length (TL) as an indicator of the aging process is a step toward influencing the telomere and telomerase system. Studies have measured TL in relation to various age-related diseases, with conflicting findings. Given the inconsistency in findings later in the lifespan, the causes of age-related disease are hypothesized to be detectable at conception, infancy, or early in the lifespan. Therefore, a lifespan perspective on TL is necessary to fully understand the role of telomeres in age-related diseases.

The lifespan perspective posits that significant changes occur as a continuous process, such that "development" to "aging" is one mechanism where biological modifications are always occurring within an individual from conception to death. Across a lifespan, many factors influence health outcomes that are often observed at one point in the lifespan. However,

investigating a health outcome while considering the entire lifespan allows for integration of early exposures which are demonstrated to create health risks later in life.

Remarkably, studies across the human lifespan report TL as a biological indicator of development and aging. As a biomarker of aging, TL differentiates individuals of the same chronological age but different biological ages, while having a high correlation with chronological age at the population level. Biological aging, underpinned by changes in molecular pathways within organisms resulting in functional decline, is a profound risk factor for most aging-related diseases.

Leveraging TL as a biomarker of age-related disease could aide scientists in narrowing the research scope to a specific biological facet that explains the biological processes behind aging or aging-related diseases. Meta-analytic evidence of TL associated with overall mortality, cardiovascular disease, diabetes, and others positions TL to be applied in other research settings.[8-11] When applied in neuroscience research, associations with TL have been inconsistent. Furthermore, TL neuroscience research has been conducted with mostly homogenous non-Hispanic White populations in the United States, making the findings difficult to extrapolate upon for other populations. Therefore, the findings of this dissertation, when shown together, aimed to identify the utility of TL, with precision measurement, as a neuroscience biomarker in a lifespan perspective approach.

The following dissertation research contains three distinct groups of projects with seven chapters that examine the utility of TL as a biomarker in neuroscience. In this dissertation research, TL is investigated as a biomarker of neurodevelopment and of neuropsychological aging to identify the role of telomere biology in these aging processes. The following

experiments provide unprecedented information on TL, particularly for minoritized populations often underrepresented in research studies.

The first chapter evaluates studies comparing polymerase chain reaction (PCR)-based TL measurements and a different methodology and indicates the need for methodologic reporting and awareness of reporting recommendations for PCR-based TL measurements to ensure assay precision and reliability. The hypothesis for this first chapter was that studies comparing TL measurements from different methods have variability in correlations between methods, in part due to a lack of methodologic information to allow comparison of PCR-based TL measurements. The second chapter comprehensively meta-analyzes correlations of TL measured from multiple healthy human tissues within an individual. This chapter expands on available literature of the ability to infer TL of one tissue measured from another tissue within individuals. The hypothesis for the second chapter was that cross-tissue TL measurements would be positively correlated, which increases the ability to infer TL of neurological tissues from non-neurological tissues. The third chapter investigates the effect of separate DNA extractions from peripheral blood mononuclear cells on the repeatability of TL measurement and provides visual and written information on the monochrome multiplex quantitative polymerase chain reaction (MMqPCR)-based TL measurement protocol. The hypothesis for the third chapter was that the repeatability of TL measurement would not significantly differ between DNA extractions.

The fourth and fifth chapters provide unprecedented information on the long-term effects of prenatal neurotoxicant exposures on infant neurodevelopmental performance and the utility of TL in this context in the diverse Surinamese population. The fourth chapter, given that the cultural context differs between Suriname, the United States, and the Netherlands, aimed to 1) determine if any cultural adaptations of the neurodevelopmental assessment were needed for

Surinamese infants and 2) to evaluate its psychometric properties. The hypothesis for this chapter was that the neurodevelopmental assessment would be a valid measure for this population with adequate validity and reliability following cultural adaptation. The fifth chapter discerns the effect of prenatal neurotoxicant exposures and infant TL on infant neurodevelopmental performance. The hypothesis for this chapter was that greater prenatal neurotoxicant mixture concentrations and shorter TLs would be associated with worse neurodevelopmental performance in Surinamese infants, which may have long term impacts on neurological age-related disease risk for this population.

The sixth and seventh chapters provide information on neuropsychological performance and the utility TL to capture aging-related decline performance for American Indian / Alaska Native and African American / Black populations, as well as the first measurement of TL from cerebrospinal fluid DNA. The sixth chapter aimed to identify of the association between TL and neuropsychological performance in minoritized populations, where the hypothesis was that the Rey Auditory Verbal Learning Test (RAVLT) and the Trail Making Test (TMT), which both capture subtle cognitive changes associated with aging and ADRD, are associated with TL measured from peripheral white blood cells for a middle to old-age population. The seventh chapter defines the correlation of TL measured from peripheral blood to that measured from cerebrospinal fluid. The hypothesis for this chapter was that peripheral blood TL and cerebrospinal fluid TL would be significantly and positively correlated.

Taken together, these dissertation experiments investigate neuroscience inquires across the lifespan for minoritized populations. In this dissertation, TL is investigated as an indicator of biological aging across the lifespan, specifically as a biomarker of neurodevelopment and neuropsychological performance conducted with detailed scientific rigor. This dissertation

expands on the sparse neuroscience literature available for these minoritized populations who are underrepresented in scientific research enrollment, retention, and consideration, as well as defines the utility of TL as a biomarker in future neuroscience research.

CHAPTER 1

Comparison studies of telomere length measurement using qPCR approaches: A critical appraisal of the literature

Background:

Telomeres, the protective nucleic acid and protein cap found at the end of all eukaryotic chromosomes, have captured the attention of scientists, medical and public health professionals, biotechnology companies, and the media over the last two decades. In 1973, Olovnikov proposed his theory of marginotomy, which reasoned that during DNA replication, DNA polymerase would not be able to completely copy the first DNA segment and, to prevent the loss of critical DNA sequences in genes, a noncoding set of DNA nucleotides would be required to act as a buffer protecting the loss of important, gene-encoding, sequences [1]. Subsequently in 1978, Blackburn et al. first reported the actual DNA sequences of telomeres in yeast [2], followed by the first sequencing of the human telomere in 1988 [3]. The sequencing of telomeric DNA paved the way for the development of methods that measured the length of telomeres, beginning with the first report of telomere length measurement using the Southern blot method for mammalian chromosomes in 1988 [4]. Since then, thousands of papers assessing telomere length (TL) in human cells have been published across a myriad of different scientific fields (Fig 1). As a result of the broad scientific interest in both the role of TL in disease processes and the influence of environmental factors on TL dynamics, the number of studies evaluating TL in human cells continues to increase, in part facilitated by the regular development of new methods and modifications of existing assays.

Currently, over two dozen assays have been developed to measure TL (Fig 2) [4-35]. These assays are classifiable into four broad categories: hybridization-based, polymerase chain reaction (PCR)-based, sequence-based, and mixed methods (e.g. hybridization/PCR combination). These assays vary in the information they yield on TL. While most focus on the measurement of the average TL within the sample, assays also measure chromosome-specific TL [6, 17], the complete distribution of TL in a cell population [33], or the shortest TL [26]. The shortest TL has received considerable attention, given evidence from in vitro and preclinical models suggesting that the shortest TL is most predictive of cellular senescence [36, 37]. Several recent reviews have discussed the overall advantages and disadvantages of each method [38-40] focusing on cost, scalability, constraints of starting biological samples (e.g. living cells, amount of DNA, etc.) and, to some extent, inter and intra-lab precision as specific challenges facing the field, including the use of coefficient of variation (CV) compared to intraclass correlation coefficients (ICC) [41]. While studies of basic telomere biology continue to explore the complex

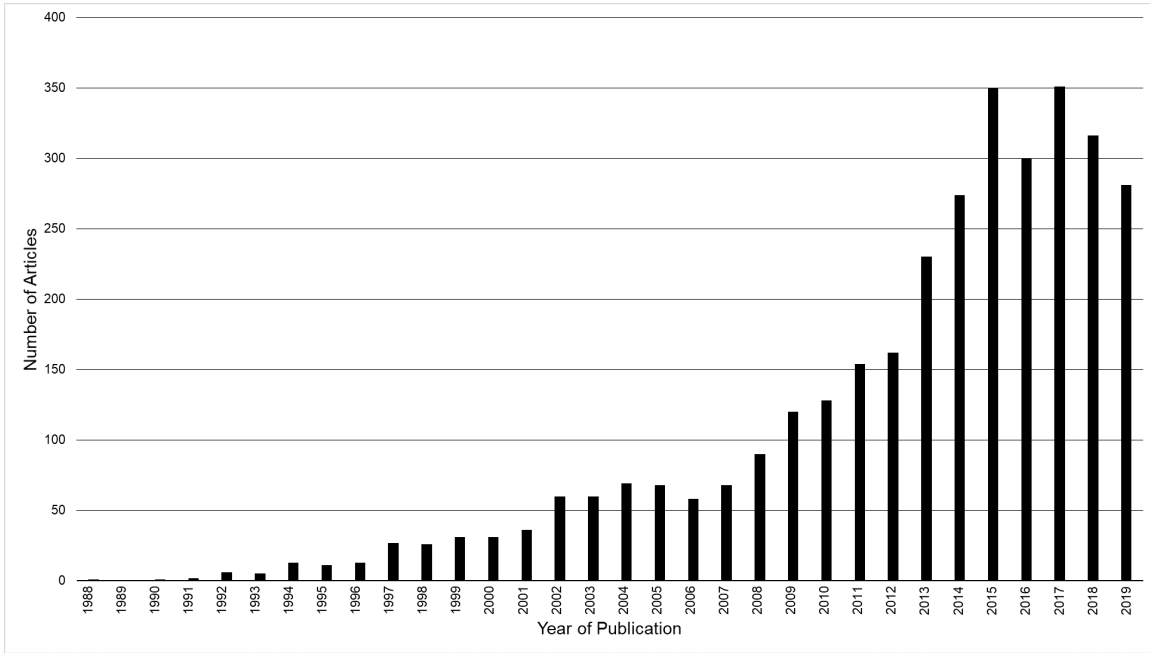


Figure 1. Telomere publications from 1988-2019. All publications mentioning telomere length of human DNA from 1988 to 2019 obtained by searching "telomere," "length," and "human" in PubMed. Search was completed on May 25, 2020.

role that telomeres play in cellular and organismal function, studies testing TL as markers of disease risk or environmental exposure must balance biological relevance, methodologic precision, and experimental practicality, similar to other epigenetic markers such as DNA methylation [42].

Over the last decade, debates have arisen over the utility and measurement of TL, particularly with regards to qPCR-based methods. This debate is partially fueled by concerns

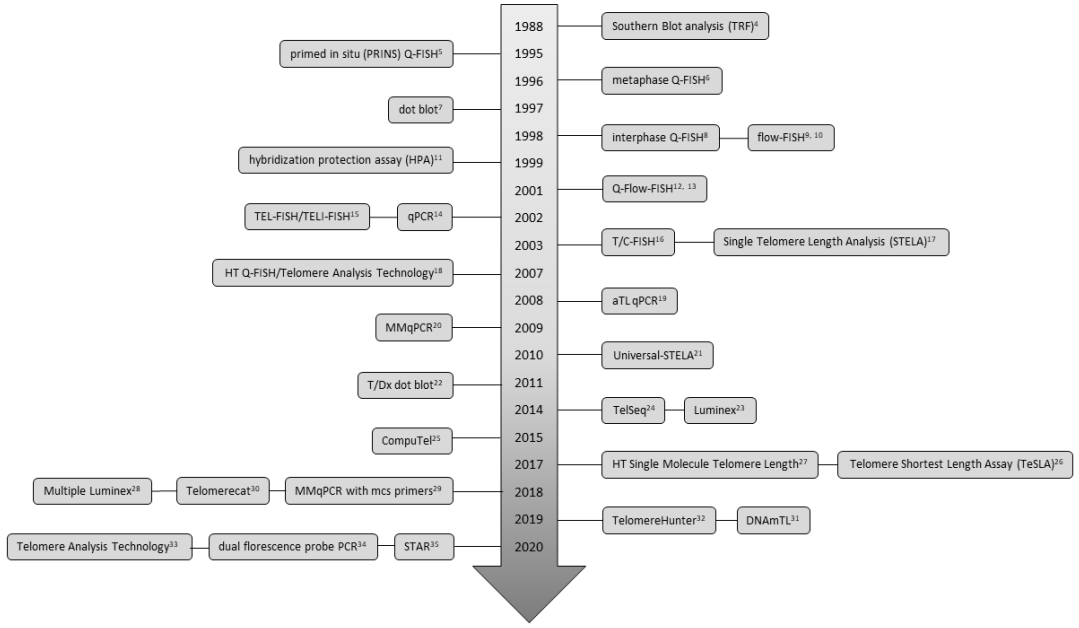


Figure 2. Novel telomere methods developed from 1988 to 2020. Data were obtained by literature search and through references of telomere review papers and method comparison papers.

related to the reproducibility and replicability of TL measurements across studies, methods, and laboratories, and is accentuated by new method development and adaptations of existing protocols without sufficient consensus on the required quality control as more laboratories begin to perform TL assays independently. In response to this debate, several studies have attempted to compare TL measured across different assay methods or laboratories. Some of these method comparison studies examined the direct correlation of TL measurement in the same sample using

different assay methods and/or tested the repeatability of TL with the same method (e.g. the amount of within assay variation) [43, 44]. Others tested the relative correlation of the TL measured by different assays with an expected phenotype (e.g. aging, parent-offspring correlation) [45], or examined the relative ability of different assays of TL to predict a specific disease or health outcome [46, 47]. Each of these approaches requires a different analytic strategy and study design and comparison of outcomes is not straightforward. To date, the existing evidence remains insufficient to answer key methodologic questions related to differences in reproducibility and replicability across measurement assays and laboratories, and how/whether these differences affect the ability of TL to serve as a biological indicator of exposure or a predictor of disease or health risk [48]. Beyond these concerns, there remains a lack of consensus as to which, if any, methodology is the "gold standard," as even the classic Southern blot method is challenged by its inability to capture potentially critical metrics (e.g. full distribution, shortest telomere length, inclusion of the subtelomeric region). To ensure reliability in the widespread utilization of TL as a biomarker of environmental exposure and/or a predictor of a disease, measured by any method that is applicable to population studies, it is critical to systematically test fundamental issues related to assay reproducibility and replicability [49].

As part of a joint National Institute of Aging and National Institute of Environmental Health Science initiative that funded a U24 cooperative award and four separate U01 awards, a Telomere Research Network (TRN) was established in 2019 (trn.tulane.edu). The TRN is coordinating cross-method comparison studies with the long-term goal of developing methodological guidelines and recommendations for telomere research applicable to population-based studies. As a first step towards the goals of this network, we undertook a systematic literature review of published studies that directly compared TL measured using at least one

PCR-based method and another approach to determine how these studies might inform the field, with particular attention to assay precision and accuracy of different measurement assays and what research gaps remain.

As defined by the Committee on Reproducibility and Replicability in Science, precision is the closeness of agreement between measured quantities obtained by replicate measurements, while accuracy is the closeness of agreement between a measured quantity and a true value [49]. Reproducibility is defined as precision in measurement under conditions that involve different locations or different measurement procedures, while repeatability is defined as precision in measurements that include the same procedures/locations. Beginning from this perspective, this systematic review evaluated the existing literature related to cross-method comparisons. This review focuses on PCR-based methodologies due to their increasing use in population-based studies, their central role in the debate related to assay precision, and the existence of two reporting guidelines - one created through the TRN (Supplemental Table 1), and a second one created by a separate group in a recently published manuscript [50]. The majority of PCR-based methods are derived from two seminal methodologic papers by Richard Cawthon, the first describing a monoplex based assay (qPCR) and the second describing a multiplex assay (MMqPCR) [14, 19].

Aim:

Our review focuses specifically on the comprehensiveness of methodologic reporting, correlation between TL measured by different assays, assay repeatability and reproducibility, and overall scientific design of methodological comparisons. Finally, we suggest areas of needed scientific examination and provide some guidance related to study design, necessary sample size, and analytic approach, to address key remaining questions: (1) What is known about the

relationship between TL measured using PCR-based methods and other assays? (2) What is known about the reproducibility and repeatability of PCR-based methods and how does this relate to other TL measurement techniques? (3) What are the implications of methodologic precision for sample size and power? (4) What are appropriate guidelines to systematically evaluate the precision of both existing and future TL assays? Addressing these important questions is a requisite step in advancing our understanding of the ability of TL, measured by any approach, to serve as a sentinel of psychosocial and environmental exposures and a predictor of future disease.

Methodology:

Manuscript search

To identify relevant papers that reported on cross method comparisons of any qPCR-based method (qPCR, absolute TL (aTL), and MMqPCR) and another method of TL assessment (PCR-based or otherwise) or the same PCR-based method conducted in separate laboratories, we conducted a critical review beginning with a literature search (Fig 3). The following key terms "telomere," "telomere length," and "human" were searched in PubMed and Web of Science. From these initial results, a second search included the keyword "PCR" to identify the initial titles for screening. Search criteria included papers published since January 1, 2002 (the year the first method to measure TL by qPCR was published) through May 10, 2020. The references of selected papers were also reviewed to identify any additional papers. A list of identified papers was presented to the TRN Steering Committee, who also suggested additional papers. Initial review of papers for inclusion was accomplished through evaluation of both the abstract and methods section, as some manuscripts were not directly focused on methodological comparison

and instead only reported the cross-method comparison on a subset of samples. Final inclusion in this review met the following baseline requirements:

1. Article published in a peer-reviewed journal (abstracts and pre-prints not included).
2. Article was not presenting the initial development of a new method or a substantial refinement of an existing methodology. This type of study was excluded due to the expectation that these new methodologic manuscripts were utilizing cross method comparison as a measure of external validity for the methodologic design. As we did not find any papers describing the failed development of a new method for TL measurement, and would not expect that to be readily publishable, to avoid any potential intrinsic bias in these highly specialized reports we opted to exclude them.
3. Included a direct comparison of TL using the same biological sample measured with two distinct TL assays or the same assay in two or more separate laboratories.

At least one of the methods used to measure TL was solely qPCR-based. TeSLA and STELA were not considered due to the additional hybridization component of the assay.

Reporting review:

Included papers were evaluated for quality of methodologic reporting using two different indices of reporting guidelines for PCR-based telomere studies. The first was created through consensus of the initial participants in the TRN (S1 Table). The second was derived from recommendations published by Morinha et al. 2020 (S2 Table) [50]. We included both guidelines for two reasons. First, there is not empirical data to distinguish between the two guidelines in terms of ensuring rigor and reproducibility for the field. Second, as several of the papers reviewed were authored by participants involved in the creation of the TRN guidelines, the inclusion of both guidelines provided some degree of impartiality. Both guidelines contain

overlap with the MIQE guidelines and include characterization of the importance of each recommendation [51]. In terms of specific differences, the Morinha recommendations included several pre-analytic considerations not included in the TRN guidelines (e.g. volume of sample processed, robotic instrumentation vs manual), while storage buffer and the percentage of samples tested for DNA integrity were included in the TRN guidelines but not in the Morinha guidelines. The latter also required greater detail for qPCR validation such as the standard curve and calibration samples, as well as a requirement for the melt curve and Ct of the negative control, which were not included in the TRN guidelines. Both guidelines assess the comprehensiveness of the information describing the PCR assay itself as well as analytic considerations for final TL determination. A grading rubric for each set of guidelines was developed to reduce subjective reviewer interpretation (S3 Table). A composite assessment for each index was divided into three subcategories for the TRN guidelines and five subcategories for the Morinha guidelines. These broadly encompassed sample collection and processing, DNA quality metrics and storage; PCR assay components and quality control; and data analysis. Two of three reviewers independently assessed each article for fulfillment of reporting guidelines (ARL, LWYM, and SSD). The scores for each individual item were compared and discrepancies resolved by the third reviewer.

Additional characteristics assessed included sample blinding prior to analyses, single lab or multi-lab testing, conversion/transformation of raw TL measurement prior to comparison, and whether the study design evaluated repeatability and/or reproducibility. Lastly, when available, sample size, means and standard deviation of TL measurement are included to assess study power. Although several studies included means and SD of the entire sample, only a subset reported the means and SD of the samples utilized in the method comparison analyses. This

review only included method comparisons that involved at least one PCR-based method as

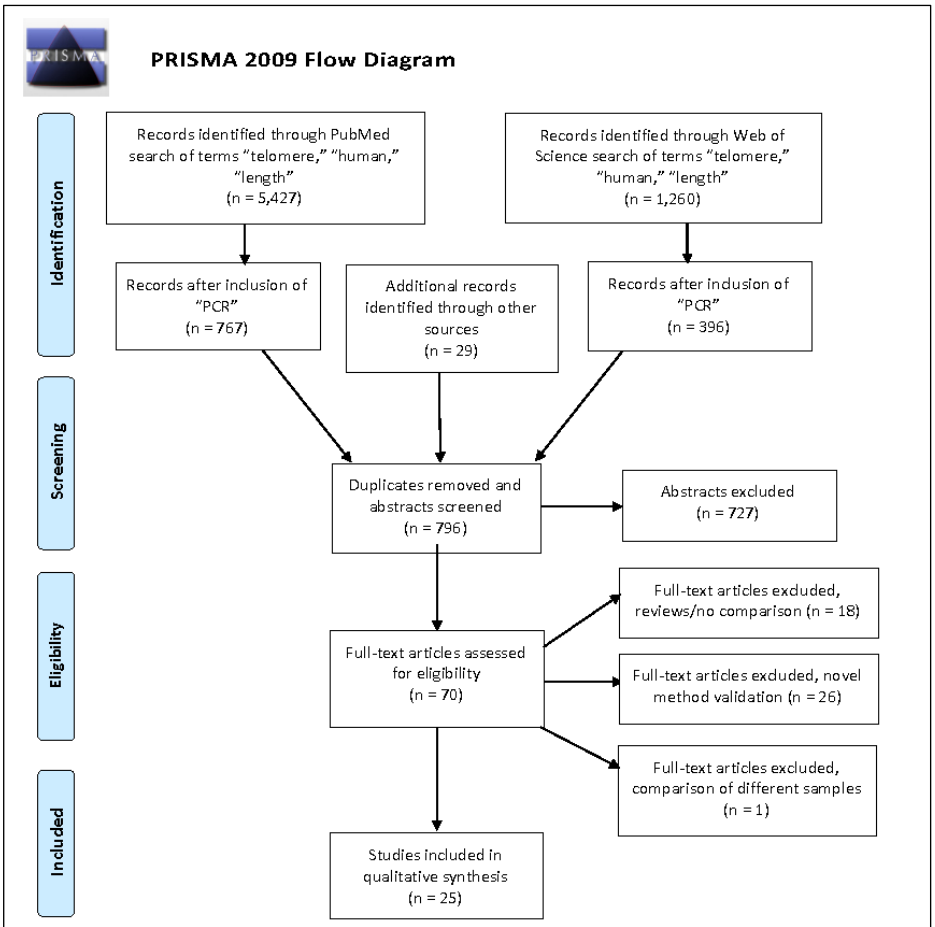


Figure 3. PRISMA flow diagram of literature search and article selection for inclusion. Date range searched was between Jan 1 2002 to May 10 2020. Other sources of article identification included suggestions from participants of the TRN and review of reference lists of selected papers.

currently reporting guidelines are only available for PCR-based methods. As the majority of PCR-based measurements of TL are relative, and no current assays measure the true TL, it was not possible to address accuracy.

Correlation between methods was assessed by using Pearson's r or r^2 values where provided. Weighted average correlation coefficients were determined for each type of comparison by converting reported Pearson's r values (or the square root of reported r^2 values) to Fisher's z values, and weighting by sample size. A forest plot was generated from the weighted r

average, total sample size for that correlation, and 95% confidence interval (CI) range using DistillerSR Forest Plot Generator from Evidence Partners (https://www.evidencepartners.com/resources/forest-plot-generator/).

ICC calculation:

Given established analytic shortcomings related to the use of the CV as a metric of testing the repeatability of TL, or the correlation of TL measurement between assays, raw data from cross method comparison studies was used to calculate ICCs for comparison between methods where available [41,52,53]. ICCs for one study were also calculated using a two-way, single measurement, absolute agreement, random effects model, known as ICC(A,1) and for average measurements ICC(A,k) in McGraw & Wong's (1996) terminology [54]. The R script used for calculating ICC and associated instructions can be found in the Supporting Methods. To provide guidance for future study design, we present several different power analyses outlining the relation between sample size, ICC, and ability to detect group differences. These calculations assumed a realistic (true) standard deviation of 650 base pairs (bp), an estimate routinely found in adult studies [55,56], and N is the combined n of the two groups to be compared and was assumed to be equally distributed among the two groups. We acknowledge that not all TL estimates produce base pair (bp) measurements, as such the graphs are provided based on ICC and sample size to ensure guidance to research studies utilizing TL assays that generate both relative and bp based estimates of TL. Power analysis for cross-sectional comparisons was done using G*Power [57], while power of longitudinal comparisons was estimated through simulations. To examine the impact of variation in ICC on longitudinal TL studies, the statistical power to detect a significant change (paired-t-test) in telomere length of 25 bp/year for sample sizes of 25, 50 and 100 individuals, and an interval of 8 years between

baseline and follow-up (i.e. on average 200 bp in total), as a function of measurement repeatability (e.g. reliability) expressed as the ICC. Measurement error was simulated by adding a random number from a normal distribution to the true TL, with the error set at different levels to generate variation in ICC between simulations. Population SD of telomere length was assumed to be 650 bp at both time points and telomere shortening was simulated assuming a Poisson distribution with mean/variance of 25bp/year. This is close to the mean shortening rate typically observed in adults in studies where the age-dependent SD is estimated to be close to 650 bp, and thus the scaling of shortening rate to the overall variance is realistic. Furthermore, power of comparisons using data with another SD can be read from the graphs after rescaling the data to have an SD of 650.

Results:

The initial search revealed 5427 articles and the inclusion of "PCR" as a search term limited the results to 767 articles, whose abstracts and methods were read (Fig 3). An additional 30 articles were identified through assessment of the references of method validation papers and other included cross-method validation studies. A review of these 797 abstracts identified 70 articles for assessment of the full text and supplemental information to determine inclusion in this review. Twenty-six articles were determined to be novel method validation and excluded. Eighteen articles were excluded as they were either reviews or did not include direct method comparisons. One article was excluded due to the determination that the DNA samples used for cross-method comparison were obtained at different time points. We also included nine papers that, while not specifically designed as a cross method comparison, included sufficient details comparing TL measurement using different assays. This resulted in a total of 25 articles included in this review (Table 1).

Paper Characteristics

Table 1. Characteristics of Included Papers.

Study	Year	Methods	Multi-lab?	Sample Type	Sample Size	Age Range (years)	TL Mean ± SD
Martin-Ruiz* [58]	2004	qPCR x TRF		cell lines	22	N/A	NR
Gardner* [59]	2007	qPCR x TRF		kidney	32	0.1-71.4	NR
Hunt* [60]	2008	qPCR x TRF		blood	72, 162	19-93	T/S = 2.71 ± 0.08 (n=36); 2.71 ± 0.05 (n=81); 3.01 ± 0.10 (n=36); 3.05 ± 0.06 (n=81)
Salpea* [61]	2008	qPCR x TRF	Yes	blood	32	24-54	T/S = 1.38 (1.31-1.44) (n=765)
Ehrlenbach*† [62]	2009	qPCR x TRF		blood	56	51-81	T/S = 1.31 (0.95-1.90) (n=669)
Pavesi* [63]	2009	qPCR x TRF		blood, cord blood	28	0-97	NR
Aviv [43]	2011	qPCR x TRF	Yes	blood	50	41-70	NR
Imam* [64]	2012	qPCR x FlowFISH		blood, cord blood, DBS	29	0	T/S = 6.23 ± 1.54 (n=35); 5.64 ± 1.54 (n=51)
Zanet [65]	2013	qPCR x MMqPCR		blood	32	2-59	NR
Gutierrez-Rodrigues [47]	2014	qPCR x FlowFISH x TRF		blood, cord blood	70, 51	0-88	NR
Martin-Ruiz [44]	2015	qPCR x MMqPCR x TRF x STELA	Yes	cell lines, placenta	12	N/A	NR
Eisenberg [45]	2015	MMqPCR x TRF	Yes	blood	190	NR	NR
Jodcyzyk* [66]	2015	qPCR x TRF		blood	20	0-35/50	T/S = 1.184 ± 0.371 (n=677); 1.104 ± 0.153 (n=351)
Panero* [67]	2015	aTL x TRF		bone marrow cells	102	30-87	7.2 ± 0.46 kb (n=48); 6.15 ± 0.41 kb (n=54)
Hsieh [68]	2016	qpCR x MMqPCR x TRF x flowFISH		blood, cord blood, DBS, placenta, cell lines	33 - 84	NR	NR
Lynch*† [69]	2016	qPCR x TRF	Yes	blood	101, 111	53-63, 26-43	6.4 kb (4.3-8.3) (n=101); 6.3 kb (5.1-7.9) (n=111)
Wand [70]	2016	qPCR x TRF x FlowFISH		blood	11	24-69	NR
Behrens [71]	2017	qPCR x FlowFISH x T/C-FISH x TRF		blood, cord blood	83 - 99	0-81	NR
Khincha† [72]	2017	qPCR x TRF x FlowFISH		blood	35, 53	3-69	T/S = 0.50 (0.21-1.48) (n=35); 0.91 (0.45-1.27) (n=53)
Lee [73]	2017	qPCR x TRF x WGS		cell lines	20	N/A	NR
Tarik [74]	2018	MMqPCR x TRF		blood	94	NR	T/S = 1.02 ± 0.32
Wang [75]	2018	qPCR x FlowFISH		blood	181	19-53	NR
Gadalla* [46]	2018	qPCR x qPCR	Yes	blood	100	NR	NR
Ventura Ferreira [76]	2019	MMqPCR x FlowFISH		blood	105	24-66	NR
Ropio [77]	2020	qPCR x TRF x aTL		cell lines	20	N/A	4.32 ± 0.143 kb

Sample size reflects the n used for method comparisons. If the TL mean and standard deviation (SD) reflects reported values for the entire sample size, the n is included in parentheses.
* Papers not specifically intended to compare correlation of methods.
† TL median and range were provided in lieu of mean and SD.
N/A: not applicable to this sample type.
NR: not reported
DBS: dried blood spot

The most common methods comparison among the 25 papers evaluated in this review was monoplex quantitative PCR (qPCR) and the telomere restriction fragment (TRF) method by Southern blot (n=17). Four studies compared multiplex qPCR (MMqPCR) with TRF. Seven studies compared qPCR with the flow-FISH method, and two studies compared MMqPCR with

flow-FISH. Three studies involved comparisons of MMqPCR and qPCR, and two studies involved direct comparisons of qPCR methods or separate laboratories. Other methods examined included STELA (compared with qPCR, MMqPCR, and TRF in one study) [44], and absolute TL (two studies, one comparing aTL to TRF and one comparing to aTL to qPCR) [67,78]. T/C-FISH was also examined in one study as it related to qPCR [71]. Additionally, one paper compared the correlation of several whole genome sequencing (WGS) platforms to qPCR-based measurement [73]. Note some studies compared more than two methods [44, 47, 68, 70-73, 75, 78].

Table 2. Results of reporting guidelines analysis for TRN and Morinha reporting guidelines.

	Category	**Average Score**
TRN Reporting Guidelines	Sample	37%
	qPCR assay	75%
	Data analysis	39%
	Total	**52%**
Morinha Reporting Guidelines	Sample	46%
	DNA extraction	32%
	qPCR validation	14%
	qPCR protocol	66%
	Data analysis	78%
	Total	**50%**
	Average of Both	**51%**

Average score reflects the average score across all papers for each reporting category. Total scores reflect the average score across all papers from all reporting items.

Whole blood was the most common sample type used (n=19), but cord blood (n=5), peripheral blood mononuclear cells (PBMCs) (n=4), and cell lines (n=5) were also utilized as well as a range of other sample types. Several studies reported on more than one sample type. The reported sample size for the cross-method comparisons ranged from 12 to 181 and only 7 papers had a total sample size greater than 100. Five studies reported the means and standard

deviations and two reported the median and range of TL measurements for the study. Two studies provided the raw values of the PCR-based TL measurements.

Table 3. Item-specific reporting results of the TRN Reporting Guidelines.

Category	Reporting Item	# Papers Reported
Sample type, storage, DNA extraction and integrity	Sample type	25
	Sample storage temperature	5
	Sample storage time before extraction	2
	Sample storage buffer	6
	DNA extraction method	20
	DNA storage conditions	4
	DNA freeze-thaw cycles	4
	Method of documenting DNA quality/integrity	8
	% of samples tested for DNA quality/integrity	6
qPCR assay	State type of PCR method	25
	PCR machine	21
	Source of master mix & reagents	22
	Final reaction volume	19
	Telomere primer sequences	21
	Telomere primer concentration	21
	Single copy gene name	25
	Single gene primer sequence	22
	Singe gene primer concentration	20
	Full PCR program	18
	PCR efficiency of both primers	3
	Source of control samples	14
	Concentration of DNA standard	13
Data analysis	Mean and standard deviation/median and range of TL	13
	Number of sample replicates	24
	Level of independence of replicates	14
	Analytic method to determine final TL	23
	Method of accounting for variation between replicates	11
	Method of accounting for well position effects	5
	Method of accounting for between-plate effects	12
	% of samples repeated as a result of failed QC	3
	% of samples excluded from analysis	3
	Acceptable range of PCR efficiency for primers	6
	ICCs of sample/study groups	2
	T/S ratio transformed to Z score prior to analysis	2

Overall quality of reporting of PCR assay methodology

Of the 25 studies included in this systematic review, the average completion score across both reporting guidelines was 51%, with an average of 52% for the TRN guidelines and 50% for the Morinha guidelines (Table 2). Overall, papers included between 26-75% and 29-78% of the recommended reporting metrics for the TRN and Morinha guidelines, respectively. Some metrics were consistently reported in nearly all papers, including the sample type, single copy gene name, and type of PCR method utilized. However, only about 10% of the included papers reported on sample storage, PCR efficiencies, or the number of samples excluded due to quality concerns with the assay.

DNA Processing

For both the TRN guidelines and Morinha guidelines, reporting of sample type, storage, DNA extraction, and DNA quality/integrity was poor, with an average of 37% for the TRN Sample/DNA category, 46% for the Morinha Sample category, and 32% for the DNA category of the Morinha guidelines. Storage conditions for both the biological samples and extracted DNA were poorly reported, with 24% or less of studies providing this information (Table 3). Fewer than half of the studies reported on metrics related to DNA integrity.

PCR Assay

Reporting on PCR assay conditions and quality control varied. While many metrics of the PCR assay were well-reported, only 18 of 25 studies reported the full cycling conditions. The lowest reporting metric related to PCR was experimental efficiency, with only 12% reporting actual PCR efficiencies. Additionally, just over half (56%) of studies reported the source of their control samples.

Analytic Approaches

Several key reporting gaps were noted in relation to assay quality control and analytic approaches to determining final TL. As with any biologic assay, there is the potential that a specific sample will fail quality control metrics and need to be repeated. Only six papers reported on the number of samples repeated and/or the number of samples that failed quality control. While all but one study reported the numbers of sample replicates, surprisingly, only 14 studies reported on the level of independence of sample replicates (e.g. replicates run on the same plate or on different plates/different times) and only half of the studies reported the means and standard deviations (or median and range) of the T/S ratio.

Cross-Laboratory Studies

Only six studies compared analyses across more than one laboratory (Table 1) [43-46]. Of these studies, three described how samples were blinded before analyses. Further, of these cross-laboratory studies, only three studies included the same assay performed in different laboratories [44, 46, 69].

Reproducibility

Reproducibility, a critical criterion for biologic assays, refers to the relation between measurements using the same assay in different locations or the comparison of values generated using different measurement procedures. This systematic review attempted to assess the relative reproducibility of PCR-based measure of TL in different laboratories as well as the reproducibility precision, e.g. the closeness of two or more measurements, in TL measurement using different methods.

Relative Reproducibility

The current literature does not provide sufficient data to address the relative reproducibility, as, to date, only three studies have tested this directly by performing the same

assay in different laboratories or settings. In one study that blinded comparison samples before

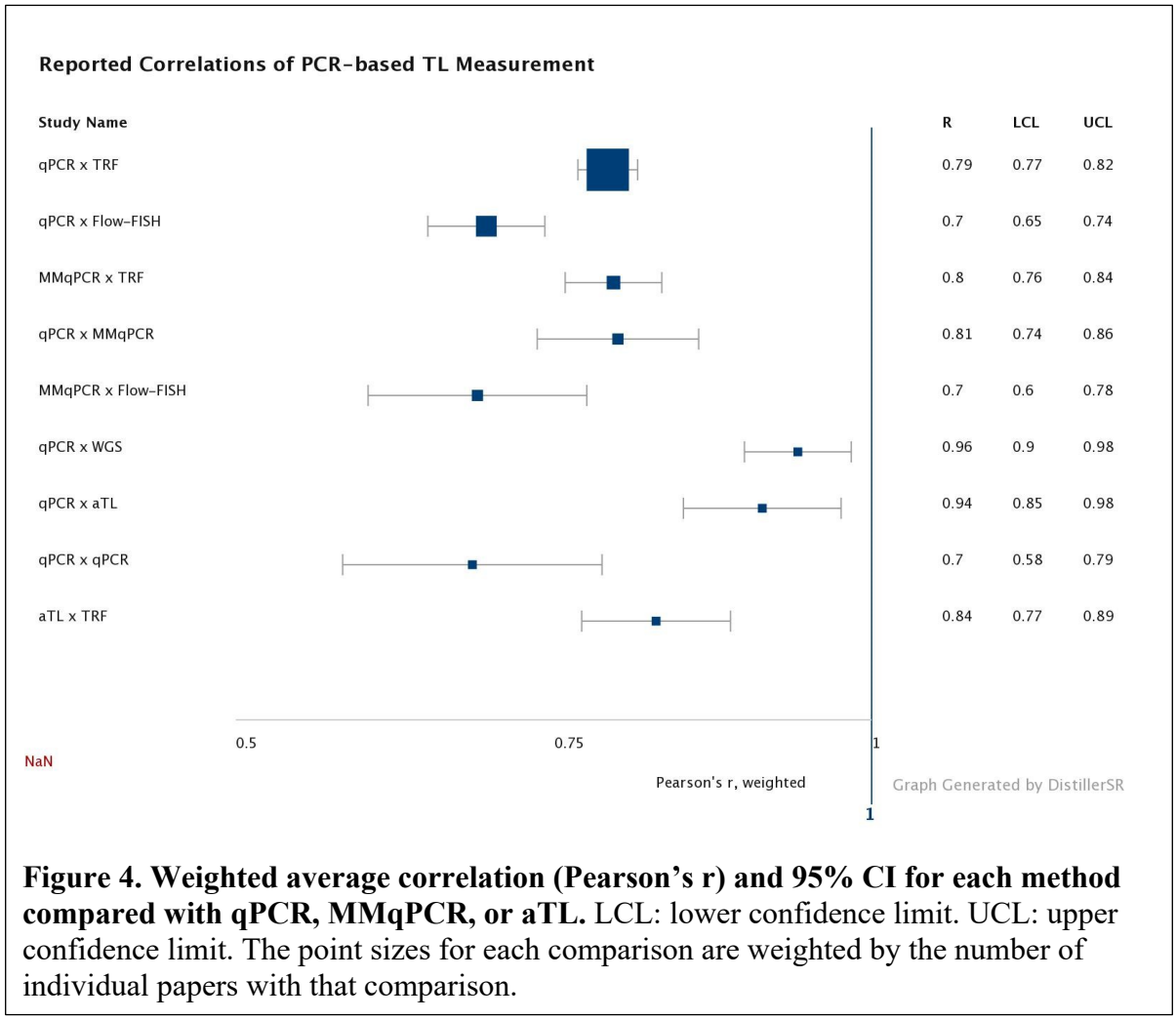


Figure 4. Weighted average correlation (Pearson's r) and 95% CI for each method compared with qPCR, MMqPCR, or aTL. LCL: lower confidence limit. UCL: upper confidence limit. The point sizes for each comparison are weighted by the number of individual papers with that comparison.

they were sent to the external laboratories, the median CV across laboratories for qPCR was 18.3%, while the median CV for STELA/TRF based TL measurement was 9.2% [44]. However given the dependence of CVs on the y-intercept, the interpretation of these CVs remains challenging [41]. In the second study, where samples were not blinded before being assayed, the reported within-lab CVs for replicate qPCR measurements were 2.5% and 8.6% [46]. As the laboratories involved utilized different PCR primers, and slightly different methods, it was not possible to directly compare cross-laboratory reproducibility. The third study found inter-assay CVs of 12.0 and 1.2% in two participating labs performing qPCR, but an additional laboratory's

results were excluded from analysis due to an extremely high CV of 27%. Correlation between each laboratory and TRF results were calculated, but no correlation results were provided for the two qPCR assays, and ICC estimates were not reported.

Reproducibility Precision

Reproducibility precision, i.e. the closeness of two or more measurements using different techniques, was addressed to some extent in 19 of the 25 studies reviewed. However, only six studies involved assays performed in different laboratories. The correlation of TL measurement with qPCR-based assays to other assays was, for the most part, reported as linear regression and correlation coefficients (Fig 4). Other papers reported Bland-Altman analyses or did not report a measure of correlation at all. One paper reported mean LTL values for both TRF and qPCR but did not report a measure of correlation [69]. When examining these results, it should be kept in mind that these methodological studies were generally done in laboratories with extensive experience in the focal technique, and as such are unlikely to be representative of the field at large.

As qPCR and TRF were the most common methods compared, these studies typically reported high correlation, with a weighted correlation coefficient for all studies around 0.75. Correlation of other methods with qPCR or MMqPCR were more variable. No studies have compared MMqPCR to aTL, or whole genome sequencing (WGS). Only one paper each compared qPCR and qPCR (in separate labs), aTL and TRF, qPCR and aTL, or qPCR with WGS. To our knowledge no studies have compared WGS data with aTL, although studies have compared TRF and WGS [24].

In five of the papers in this review, linear regression was used to extrapolate TL into kilobases (kb) from the T/S ratio using TRF values. One paper converted T/S ratio to kb before

analysis of the correlation between methods [47]. In two cases, T/S ratio were converted to kb prior to TL comparison utilizing Bland-Altman analysis [71,76]. In two of the five papers, the conversion of the T/S ratio to bp was based on analyses extrapolated from different data or measured on a different sample type, raising substantial concerns on the true measurement with uncertain implications for the r value [47,76]. Beyond concerns related to the source of the data utilized for conversion from T/S to bp before comparison across methods, this analytic approach likely to leads to inaccurate reporting [47]. Only two studies utilized the TRN-recommended procedure of transformation to z-score before comparison [44,69]. When comparing relative TL estimates such as the T/S ratio generated from qPCR, transformation of these values to z-scores will yield more informative results and improve ability to compare results between laboratories or assays [79].

Repeatability

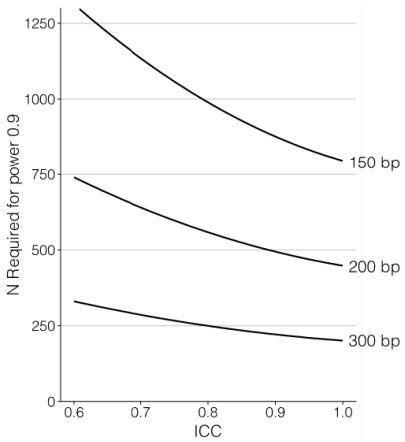


Figure 5. The sample size required to test effect sizes of 150, 200 and 300 bp with a t-test with a power of 0.9, as a function of measurement error as expressed in the ICC. Calculations assumed a realistic (true) standard deviation of 650 bp and power analysis was done using G*Power [57]. N is the combined n of the two groups to be compared and was assumed to be equally distributed among the two groups.

Repeatability, the precision in measurements that include the same procedure/locations, revealed the greatest variation in both lab and assay specific precision and between methods. In

these studies, the number of replicates for a specific DNA sample ranged from 1 replicate (i.e. sample analyzed twice) to five replicates (each sample analyzed six different times). Additionally, only four reviewed studies reported the number of samples that were repeated due to within replicate variance, despite clear acknowledgement in the field that a proportion of all studies will ultimately require repeated assays of telomere length as a result of between-replicate variance. While of limited utility in confirming precision, the intra-assay reported CVs for PCR-based methods (qPCR, MMqPCR) in this review ranged from 2.5 to 12%, and the inter-assay CVs ranged from 3.97 to 15.9%. Inter-assay CVs for TRF ranged from 1.25 to 6.3%, with intra-assay CV reported in only one paper as 1.20%. Inter-assay CVs for flow-FISH were reported as 9.3% and 10.8% in two papers, with only one reporting an intra-assay CV of 9.6%. One paper examining the aTL assay reported its inter-assay CV as 6.7% and intra-assay CV as 2.5%. We emphasize, however, that there are analytic concerns related to the use of CVs for cross laboratory comparisons [41,53], and directly converting CVs to ICC values is not possible.

Only two studies utilized ICC analyses to examine the repeatability of replicates, reporting ICCs of 0.89 and 0.92 for PCR-based measurement [72, 74]. To expand data on the repeatability of PCR-based and other TL methods, raw data was obtained from authors of a subset of these papers, and ICCs independently calculated. Calculated/reported ICC for TRF methods ranged from 0.92- 0.99 in the studies included in this review and are consistent with the ICCs reported in existing studies utilizing the TRF (0.95 to 0.99). However, it is of note that these ICCs were almost entirely the result of TRF measurement in one laboratory. The ICCs for qPCR-based methods in reviewed papers ranged from 0.89-0.92, including the two reported in manuscripts and an additional ICC calculated from raw data (ICC = 0.915, SE = 0.023, 95% confidence interval: [0.860, 0. 946], P<0.001; reported CV 6.5%) [41,80]. ICCs for MMqPCR

(triplicates on the same plate) from one study were run separately based on year of analyses. In one set (n=873) run across different PCR plates in initial and duplicate runs, the ICCs were ICC(A,1) = 0.82 (95% CI 0.79-0.84) and ICC(A,k)=0.90 (95% CI 0.88-0.91). Because these samples were re-run due to initially high intra-assay CVs, this is possibly an under-estimate of the true ICC value. For these same samples, TRF ICCs were calculated from duplicate gels on a subset (n=159) and the inter-gel ICC = 0.96 (95% CI 0.94-0.97). However, we note that these TRF analyses were conducted by a trainee which likely decreased repeatability compared to what is typical of experienced technicians. Given the significant variation in methodologic and raw data reporting, and the wide variability in published CVs, it is likely that the majority of existing TL studies not specifically comparing methods would have significantly lower ICCs.

Determination of effects of ICC variability on sample size and study power

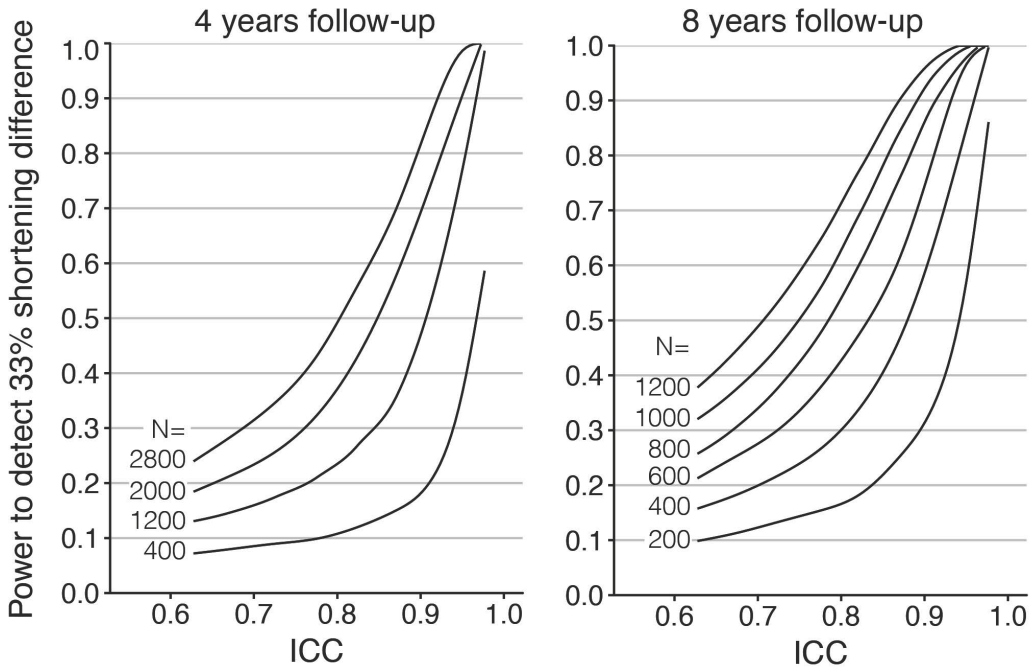


Figure 6. Power to detect a 33% change of telomere shortening rate, up or down, with p<0.05 relative to a baseline shortening rate of 25 bp/year. D. Four-year follow-up period. E. Eight-year follow-up period. Power was calculated for sample sizes as shown (200 – 2800), equally divided over the two levels of telomere shortening rate. Baseline telomere shortening was simulated assuming a Poisson distribution with mean/variance of 25, and population SD of telomere length was maintained at 0.65 kb at both time points.

Our systematic review revealed wide variation in TL measurement repeatability. No biologic assay is perfect, and laboratories measuring any biologic substrate vary in their own internal quality control and repeatability. To provide general guidance for investigators, we therefore conducted analyses to evaluate the impact on power and sample size across a range of ICCs.

In Fig 5, we present the sample size required to test effect sizes of 150, 200, and 300 bp with a t-test with a power of 0.9, as a function of measurement error as expressed in the ICC. To contextualize the differences: 150 bp is the approximate difference found between the sexes, and 300 bp is the approximate difference observed between individuals with and without atherosclerotic cardiovascular disease [81]. As directly converting bp to T/S ratios is not feasible in this analysis and the analyses below, we suggest that investigators using T/S or other relative TL measurements use standard deviation (SD) differences to estimate power. For example. a difference of 150 bp is equal to 150 / 650 = 0.23 SD, which can be converted to a T/S difference when the SD of the T/S measurements is known. Estimates of potential difference can be extracted from existing literature related to their exposure or outcome of interest when considering study design and sample size.

Finally, we present the statistical power of different sample sizes to detect a significant difference in telomere shortening rate of 33% using longitudinal data, as a function of measurement reliability expressed as ICC (Fig 6). This analysis revealed that even with a high ICC (>0.9), large sample sizes are required to yield sufficient statistical power to detect differences in telomere shortening rate, in particular when the follow-up period is short. This is due to the mean rate of telomere shortening being low (here 25 bp/year) compared to the TL

variation between individuals (here an SD of 650 bp). The rate of base pair loss in infants and children is likely significantly different and, but as of yet is poorly characterized (but see [82]).

Discussion:

This systematic review found a total of 25 papers documenting comparison between TL measured using a PCR-based methodology and another TL assay. Until recently, no publication reporting guidelines existed for qPCR-based TL measurement. Our review focused on method comparison studies with the expectation that critical assay parameters and methodologic description would be more detailed and specific. Our review, using two separately developed reporting guidelines, found that, on average, only half of the recommended factors were documented, indicating the need for increased methodologic reporting and wider awareness of reporting recommendations. The lowest reporting was related to information about the validation of PCR-based assays outlined in the Morinha guidelines, with only seven papers including any of the recommended factors. PCR efficiencies, a key reporting requirement in both guidelines and the MIQE guidelines, was absent from the majority of papers with only six mentioning the PCR efficiency parameters and only three documenting the actual PCR efficiencies. Given that all PCR-based methods either assume or specifically calculate the PCR efficiency when determining the T/S ratio, and that, in general, the determination of the T/S ratio assumes similar efficiencies for the single gene and the telomeric primers, the absence of this key metric is concerning. Fewer than half of studies failed to comment on key pre-analytic factors, specifically sample storage time and conditions, freeze-thaw cycles, and evaluation of DNA quality and integrity, all potential sources of assay variability for both PCR and non-PCR-based TL assays that may contribute to current debates in the field about the utility of TL [77,83,84]. Lastly, the reporting of the number of samples failing initial quality control, repeated, or unable to be assayed was

low. In laboratories routinely performing TL measurement using any assay, a certain percentage of samples for each study will require repeating and regularly a small subset may be unanalyzable for various reasons. While it is possible that these factors were considered and monitored, the lack of reporting for this metric heightens the need for increased attention to the proposed reporting recommendations. Moving forward, the widespread dissemination of these qPCR reporting guidelines to study sections, peer reviewers, and scientists represents an important next step in enhancing the scientific rigor of the field.

At this time, evaluation of the existing literature fails to provide sufficient evidence of the relative or precision reproducibility of different TL assays. Our review identified only six studies that included cross laboratory comparisons and, of these, only three evaluated PCR-based assays performed in more than one lab. As the number of laboratories performing TL studies using PCR and other methods continue to increase, the lack of clear data about cross-laboratory reproducibility and the absence of existing DNA standards or other methods to account for cross-laboratory variation substantially limits the ability to characterize relative reproducibility. In terms of reproducibility across different methods (e.g. PCR and TRF, or PCR and FISH), the current variability in findings, particularly when coupled with limited methodologic reporting, highlights the need for additional rigorous and blinded cross laboratory studies that are adequately powered to accurately determine how TL in a population measured using different assays truly relates. Although 17 studies evaluated the relationship between qPCR and TRF, due to the wide range in reported correlations between TL measurements, the relatively small samples sizes, and the insufficient analytic and assay blinding, there is currently insufficient data to draw firm conclusions on the general correlation between TL measured with different assays. The analytic consequences of using CVs to test the relationship between TL measured using

different assays has been discussed previously, as has the issues caused by the use of analytic strategies such as conversion to base pairs instead of z scores, especially when extrapolating from data produced in different laboratories or using different samples [41,52]. In this review, we utilized existing raw data from 3 included studies to provide preliminary data about precision reproducibility for PCR and TRF studies. The wide range of ICCs calculated from these few studies, particularly for PCR-based methods, and the low reporting of ICCs in the papers included in this review highlights the need to increase attention to the importance of reporting ICC statistics. For many of the existing studies, the small sample size and the lack of reporting of the means and standard deviations of TL prevents objective determination of whether any of the current studies were adequately powered. Beyond these concerns, the over-representation of data from specialized laboratories, particularly for TRF, the applicability of much of the existing data to the wider telomere field is uncertain. For aTL and MMqPCR-based TL measurements, the current paucity of published cross method comparisons limits the ability to form an opinion of how TL measured with these assays relates to other methods.

Measurement precision is critical, in particular for longitudinal studies. Methodologies that are low cost, practical, and simple to implement with standard laboratory equipment, especially when they are innovative or high impact, are often rapidly implemented across laboratories with various levels of expertise in the new methodology. Invariably this results in diverse protocols, analyses, and methodologic reporting – consequences that are even more problematic when there is an absence of consensus on best practices [51]. As with many other biologic assays, the development of reporting guidelines for TL measurement has lagged behind the broad implementation of the methods themselves [85-89]. The lack of consolidated guidance about factors, both pre-analytic and within the assay itself, that contribute to measurement error

when combined with the wide popularity of PCR-based TL measurement undoubtedly contributed to discrepancies in the existing literature and failed study replications. Similar to the MIQE guidelines, the reporting guidelines presented and tested in this systematic review for PCR-based TL assays are meant as minimal reporting recommendations focused on enhancing the reliability of results, consistency between different laboratories performing the same assay, and increased experimental transparency and accuracy [51]. To assist investigators and reviewers we highlight the overlapping recommendations with the MIQE guidelines, indicate whether a particular requirement is desirable or essential, and provide references that support the selection of the particular reporting requirement. Over the course of the next four years, the TRN expects to develop similar reporting recommendations for other types of TL assays while conducting adequately powered and scientifically rigorous studies to support these reporting guidelines, recognizing that individual recommendations have varying levels of initial empirical support [90].

Despite the strengths of this review, there are several limitations. First, this review only focused on assays applicable to population-based studies in humans. It does not address issues in other species or assays that may have clinical utility but for which the requirements for sample types (e.g. fresh tissue and/or live cells) or the cost/labor/expertise requirements (e.g. TeSLA, STELA) limit utilization in population based studies. A second limitation is that we utilized reporting guidelines for qPCR-based assays only. To date, specific protocol recommendations and reporting guidelines have not been published for other TL assays (e.g. TRF, FISH) although detailed methodologic protocols do exist [91]. Additionally, it is possible that additional articles comparing TL assays may be available in other databases or pre-print servers. However, many of the articles included in this review were not specifically designed solely to compare TL

measurement methods and would not be found through standardized database searches. Further, it is unlikely that additional articles would change the general picture emerging from this review. Finally, we note that while this article focused on precision and reproducibility, accuracy of measurement is as important. Precisely inaccurate measures will be of limited use to the scientific field, a factor that becomes more problematic when using relative estimates and not true values as is the case in many TL assays. In the absence of a clear gold standard measurement technique, accuracy is difficult to discern.

Conclusions

After careful examination of the existing literature, it is apparent that rigorous cross laboratory and methodological studies must be an immediate priority for the field. To assist the field moving forward, we include reporting guidelines for PCR-based TL assays and indicate specific scientific papers that support these recommendations originally developed through consensus of the initial TRN members and consultants. These guidelines do not outline a specific PCR methodology and, at this time, we do not believe there is sufficient data to provide guidance on specific assay approaches or components. Rather, these guidelines are provided to ensure reviewers and readers can adequately assess the methodology and consider the implications of these factors for each study's findings. The consistency in results across reporting guidelines (TRN, Morinha, MIQE) related to the integrity and quality of both the initial biological sample and the DNA itself support the critical nature of this reporting metric. In terms of assay reporting, increased attention of investigators and reviewers to ensuring complete reporting of assay reagents and PCR efficiencies is also expected to enhance the rigor of the field. TRN investigators are currently testing the impact of different pre-analytic factors, DNA integrity, and PCR conditions to provide evidence of the importance of these parameters in relation to

precision and reproducibility. We recommend that studies be required to report ICCs in lieu of CVs, as well as either the median or mean and standard deviation of TL. We also provide specific guidance related to sample size and power that is contingent upon the ICC given the substantial impact of differences in assay precision on the ability to determine true relationships and with the expectation that this will be of use for investigators as they embark on new research studies. It is important to balance assay cost, in terms of both time and reagents, with the needed sample size and statistical power. Moving forward, investigators should carefully consider study design from this perspective, recognizing that there is currently no "ideal" approach. Telomere research offers significant potential across a diverse range of scientific fields with potential mechanistic insight into overlapping biological pathways contributing to many of the leading causes of morbidity and mortality. Ensuring the highest scientific rigor and precision, through accurate methodological reporting and rigorous testing of the factors that contribute to assay variability, are requisite steps to ensuring that potential is achieved.

References:

1. Olonikov A. A theory of marginotomy. J Theor Biol. 1971;41:181-90.

2. Blackburn E, Gall J. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. Journal of molecular biology. 1978;120(1):33-53.

3. Moyzis R, Buckingham J, Cram L, Dani M, Deaven L, Jones M, et al. A highly conserved repetitive DNA sequence,(TTAGGG) n, present at the telomeres of human chromosomes. Proceedings of the National Academy of Sciences. 1988;85(18):6622-6.

4. Allshire RC, Dempster M, Hastie ND. Human telomeres contain at least three types of G-rich repeat distributed non-randomly. Nucleic Acids Res. 1989;17(12):4611-27. doi: 10.1093/nar/17.12.4611. PubMed PMID: 2664709.

5. Therkelsen A, Nielsen A, Koch J, Hindkjaer J, Kølvraa S. Staining of human telomeres with primed in situ labeling (PRINS). Cytogenet Cell Genet. 1995;68(1-2):115-8. Epub 1995/01/01. doi: 10.1159/000133903. PubMed PMID: 7525160.

6. Lansdorp P, Verwoerd N, van de Rijke F, Dragowska V, Little M, Dirks R, et al. Heterogeneity in telomere length of human chromosomes. Hum Mol Genet. 1996;5(5):685-91. Epub 1996/05/01. doi: 10.1093/hmg/5.5.685. PubMed PMID: 8733138.

7. Bryant J, Hutchings K, Moyzis R, Griffith J. Measurement of telomeric DNA content in human tissues. Biotechniques. 1997;23(3):476-8, 80, 82, passim. Epub 1997/09/23. doi: 10.2144/97233st05. PubMed PMID: 9298219.

8. de Pauw E, Verwoerd N, Duinkerken N, Willemze R, Raap A, Fibbe W, et al. Assessment of telomere length in hematopoietic interphase cells using in situ hybridization and digital fluorescence microscopy. Cytometry. 1998;32(3):163-9. Epub 1998/07/17. doi: 10.1002/(sici)1097-0320(19980701)32:3<163::aid-cyto1>3.3.co;2-v. PubMed PMID: 9667504.

9. Hultdin M, Grönlund E, Norrback K, Eriksson-Lindström E, Just T, Roos G. Telomere analysis by fluorescence in situ hybridization and flow cytometry. Nucleic Acids Res. 1998;26(16):3651-6. Epub 1998/08/01. doi: 10.1093/nar/26.16.3651. PubMed PMID: 9685479; PubMed Central PMCID: PMCPMC147775.

10. Narath R, Lörch T, Greulich-Bode K, Boukamp P, Ambros P. Automatic telomere length measurements in interphase nuclei by IQ-FISH. Cytometry Part A: the Journal of the International Society for Analytical Cytology. 2005;68(2):113-20.

11. Nakamura Y, Hirose M, Matsuo H, Tsuyama N, Kamisango K, Ide T. Simple, rapid, quantitative, and sensitive detection of telomere repeats in cell lysate by a hybridization protection assay. Clin Chem. 1999;45(10):1718-24. Epub 1999/10/03. PubMed PMID: 10508116.

12. Law H, Lau Y. Validation and development of quantitative flow cytometry-based fluorescence in situ hybridization for intercenter comparison of telomere length measurement. Cytometry: The Journal of the International Society for Analytical Cytology. 2001;43(2):150-3.

13. Rufer N, Dragowska W, Thornbury G, Roosnek E, Lansdorp P. Telomere length dynamics in human lymphocyte subpopulations measured by flow cytometry. Nat Biotechnol. 1998;16(8):743-7. Epub 1998/08/14. doi: 10.1038/nbt0898-743. PubMed PMID: 9702772.

14. Cawthon R. Telomere measurement by quatitative PCR. Nucleic Acids Res. 2002;30(10):e47- 53.

15. Meeker AK, Gage WR, Hicks JL, Simon I, Coffman JR, Platz EA, et al. Telomere length assessment in human archival tissues: combined telomere fluorescence in situ hybridization and immunostaining. The American Journal of Pathology. 2002;160(4):1259-68. doi: 10.1016/s0002-9440(10)62553-9. PubMed PMID: 11943711; PubMed Central PMCID: PMC1867217.

16. Perner S, Brüderlein S, Hasel C, Waibel I, Holdenried A, Ciloglu N, et al. Quantifying Telomere Lengths of Human Individual Chromosome Arms by Centromere-Calibrated Fluorescence in Situ Hybridization and Digital Imaging. The American Journal of Pathology. 2003;163(5):1751-6. doi: https://doi.org/10.1016/S0002-9440(10)63534-1.

17. Baird D, Rowson J, Wynford-Thomas D, Kipling D. Extensive allelic variation and ultrashort telomeres in senescent human cells. Nat Genet. 2003;33(2):203-7. Epub 2003/01/23. doi: 10.1038/ng1084. PubMed PMID: 12539050.

18. Canela A, Vera E, Klatt P, Blasco M. High-throughput telomere length quantification by FISH and its application to human population studies. Proc Natl Acad Sci U S A. 2007;104(13):5300-5. Epub 2007/03/21. doi: 10.1073/pnas.0609367104. PubMed PMID: 17369361; PubMed Central PMCID: PMCPMC1828130.

19. Cawthon R. Telomere length measurement by a novel monochrome multiplex quantitative PCR method. Nucleic Acids Res. 2009;37:e21.

20. Bendix L, Horn P, Jensen U, Rubelj I, Kolvraa S. The load of short telomeres, estimated by a new method, Universal STELA, correlates with number of senescent cells. Aging Cell. 2010;9(3):383-97. Epub 2010/03/25. doi: 10.1111/j.1474-9726.2010.00568.x. PubMed PMID: 20331440.

21. O'Callaghan N, Fenech M. A quantitative PCR method for measuring absolute telomere length. Biological procedures online. 2011;13:3-. doi: 10.1186/1480-9222-13-3. PubMed PMID: 21369534.

22. Kimura M, Aviv A. Measurement of telomere DNA content by dot blot analysis. Nucleic Acids Res. 2011;39(12):e84-e. doi: 10.1093/nar/gkr235.

23. Kibriya M, Jasmine F, Roy S, Ahsan H, Pierce B. Measurement of telomere length: a new assay using QuantiGene chemistry on a Luminex platform. AACR; 2014.

24. Ding Z, Mangino M, Aviv A, Spector T, Durbin R. Estimating telomere length from whole genome sequence data. Nucleic Acids Res. 2014;42(9):e75. Epub 2014/03/13. doi: 10.1093/nar/gku181. PubMed PMID: 24609383; PubMed Central PMCID: PMCPMC4027178.

25. Nersisyan L, Arakelyan A. Computel: Computation of Mean Telomere Length from Whole-Genome Next-Generation Sequencing Data. PLOS ONE. 2015;10(4):e0125201. doi: 10.1371/journal.pone.0125201.

26. Lai T, Zhang N, Noh J, Mender I, Tedone E, Huang E, et al. A method for measuring the distribution of the shortest telomeres in cells and tissues. Nature communications. 2017;8(1):1-14.

27. McCaffrey J, Young E, Lassahn K, Sibert J, Pastor S, Riethman H, et al. High-throughput single-molecule telomere characterization. Genome research. 2017;27(11):1904-15.

28. Jasmine F, Shinkle J, Sabarinathan M, Ahsan H, Pierce B, Kibriya M. A novel pooled-sample multiplex luminex assay for high-throughput measurement of relative telomere length. American Journal of Human Biology. 2018;30(4):e23118.

29. Dahlgren P, Bishop K, Dey S, Herbert B, Tanaka H. Development of a New Monochrome Multiplex qPCR Method for Relative Telomere Length Measurement in Cancer. Neoplasia. 2018;20(5):425-31. Epub 2018/03/25. doi: 10.1016/j.neo.2018.02.007. PubMed PMID: 29573637; PubMed Central PMCID: PMCPMC5915991.

30. Farmery JHR, Smith ML, Huissoon A, Furnell A, Mead A, Levine AP, et al. Telomerecat: A ploidy-agnostic method for estimating telomere length from whole genome sequencing data. Scientific Reports. 2018;8(1):1300. doi: 10.1038/s41598-017-14403-y.

31. Lu A, Seeboth A, Tsai P, Sun D, Quach A, Reiner A, et al. DNA methylation-based estimator of telomere length. Aging. 2019;11(16):5895-923. Epub 08/18. doi: 10.18632/aging.102173. PubMed PMID: 31422385.

32. Feuerbach L, Sieverling L, Deeg K, Ginsbach P, Hutter B, Buchhalter I, et al. TelomereHunter–in silico estimation of telomere content and composition from cancer genomes. BMC bioinformatics. 2019;20(1):1-11.

33. de Pedro N, Díez M, García I, García J, Otero L, Fernández L, et al. Analytical Validation of Telomere Analysis Technology® for the High-Throughput Analysis of Multiple Telomere-Associated Variables. Biological Procedures Online. 2020;22(1):2. doi: 10.1186/s12575-019-0115-z.

34. Sethi I, Rasool Bhat G, Kumar R, Rai E, Sharma S. Dual Labeled Fluorescence Probe Based qPCR Assay to Measure the Telomere Length. Gene. 2020:145178. doi: https://doi.org/10.1016/j.gene.2020.145178.

35. Luo Y, Viswanathan R, Hande MP, Loh AHP, Cheow LF. Massively parallel single-molecule telomere length measurement with digital real-time PCR. Science Advances. 2020;6(34):eabb7944. doi: 10.1126/sciadv.abb7944.

36. Hemann M, Strong M, Hao L, Greider C. The shortest telomere, not average telomere length, is critical for cell viability and chromosome stability. Cell. 2001;107(1):67-77.

37. Xu Z, Duc K, Holcman D, Teixeira M. The length of the shortest telomere as the major determinant of the onset of replicative senescence. Genetics. 2013;194(4):847-57.

38. Montpetit A, Alhareeri A, Montpetit M, Starkweather A, Elmore L, Filler K, et al. Telomere length: a review of methods for measurement. Nursing research. 2014;63(4):289.

39. Shay JW, Wright WE. Telomeres and telomerase: three decades of progress. Nat Rev Genet. 2019;20(5):299-309. Epub 2019/02/15. doi: 10.1038/s41576-019-0099-1. PubMed PMID: 30760854.

40. Aubert G, Hills M, Lansdorp P. Telomere length measurement-caveats and a critical assessment of the available technologies and tools. Mutat Res. 2012;730(1-2):59-67. Epub 2011/06/15. doi: 10.1016/j.mrfmmm.2011.04.003. PubMed PMID: 21663926; PubMed Central PMCID: PMCPMC3460641.

41. Eisenberg D. 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. 2016;45(4):1295-8.

42. Ehrlich M. DNA hypermethylation in disease: mechanisms and clinical relevance. Epigenetics. 2019;14(12):1141-63.

43. Aviv A, Hunt SC, Lin J, Cao X, Kimura M, Blackburn E. Impartial comparative analysis of measurement of leukocyte telomere length/DNA content by Southern blots and qPCR. Nucleic Acids Res. 2011;39(20):e134. Epub 2011/08/10. doi: 10.1093/nar/gkr634. PubMed PMID: 21824912; PubMed Central PMCID: PMC3203599.

44. Martin-Ruiz C, Baird D, Roger L, Boukamp P, Krunic D, Cawthon R, et al. Reproducibility of telomere length assessment: an international collaborative study. International journal of epidemiology. 2015;44(5):1673-83.

45. Eisenberg D, Kuzawa C, Hayes MG. Improving qPCR telomere length assays: Controlling for well position effects increases statistical power. American Journal of Human Biology. 2015:n/a-n/a. doi: 10.1002/ajhb.22690.

46. Gadalla S, Wang T, Loftus D, Friedman L, Dagnall C, Haagenson M, et al. No association between donor telomere length and outcomes after allogeneic unrelated

hematopoietic cell transplant in patients with acute leukemia. Bone marrow transplantation. 2018;53(4):383-91.

47. Gutierrez-Rodrigues F, Santana-Lemos B, Scheucher P, Alves-Paiva R, Calado R. Direct Comparison of Flow-FISH and qPCR as Diagnostic Tests for Telomere Length Measurement in Humans. PLoS ONE. 2014;9(11):e113747. doi: 10.1371/journal.pone.0113747.

48. Bodelon C, Savage S, Gadalla S. Telomeres in molecular epidemiology studies. Prog Mol Biol Transl Sci. 2014;125:113-31. Epub 2014/07/06. doi: 10.1016/b978-0-12-397898-1.00005-0. PubMed PMID: 24993700; PubMed Central PMCID: PMCPMC6737930.

49. Plant A, Hanisch R, National Institute of Standards and Technology. Reproducibility and Replicability in Science, A Metrology Perspective A Report to the National Academies of Sciences, Engineering and Medicine Committee on Reproducibility and Replicability in Science: 2018 June 15 2018. Report No.

50. Morinha F, Magalhães P, Blanco G. Standard guidelines for the publication of telomere qPCR results in evolutionary ecology. Molecular Ecology Resources. 2020;20(3).

51. Bustin S, Benes V, Garson J, Hellemans J, Huggett J, Kubista M, et al. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clinical Chemistry. 2009;55(4):611-22. doi: 10.1373/clinchem.2008.112797.

52. Verhulst S, Susser E, Factor-Litvak P, Simons M, Benetos A, Steenstrup T, et al. Response to: Reliability and validity of telomere length measurements. International journal of epidemiology. 2016;45(4):1298-301.

53. Nettle D, Seeker L, Nussey D, Froy H, Bateson M. Consequences of measurement error in qPCR telomere data: A simulation study. PLoS One. 2019;14(5):e0216118.

54. McGraw K, Wong S. " Forming inferences about some intraclass correlations coefficients": Correction. 1996.

55. Factor-Litvak P, Susser E, Kezios K, McKeague I, Kark JD, Hoffman M, et al. Leukocyte Telomere Length in Newborns: Implications for the Role of Telomeres in Human Disease. Pediatrics. 2016;137(4). Epub 2016/03/13. doi: 10.1542/peds.2015-3927. PubMed PMID: 26969272; PubMed Central PMCID: PMCPMC4811318 conflicts of interest to disclose.

56. Benetos A, Toupance S, Gautier S, Labat C, Kimura M, Rossi PM, et al. Short Leukocyte Telomere Length Precedes Clinical Expression of Atherosclerosis. Circulation Research. 2018;122(4):616-23. doi: doi:10.1161/CIRCRESAHA.117.311751.

57. Faul F, Erdfelder E, Buchner A, Lang A. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. Behavior research methods. 2009;41(4):1149-60.

58. Martin-Ruiz C, Saretzki G, Petrie J, Ladhoff J, Jeyapalan J, Wei W, et al. Stochastic variation in telomere shortening rate causes heterogeneity of human fibroblast replicative life span. J Biol Chem. 2004;279(17):17826-33. Epub 2004/02/14. doi: 10.1074/jbc.M311980200. PubMed PMID: 14963037.

59. Gardner JP, Kimura M, Chai W, Durrani JF, Tchakmakjian L, Cao X, et al. Telomere dynamics in macaques and humans. J Gerontol A Biol Sci Med Sci. 2007;62(4):367-74. Epub 2007/04/25. doi: 10.1093/gerona/62.4.367. PubMed PMID: 17452729.

60. Hunt SC, Chen W, Gardner JP, Kimura M, Srinivasan SR, Eckfeldt JH, et al. Leukocyte telomeres are longer in African Americans than in whites: the National Heart, Lung, and Blood Institute Family Heart Study and the Bogalusa Heart Study. Aging Cell. 2008;7(4):451-8. Epub 2008/05/09. doi: 10.1111/j.1474-9726.2008.00397.x. PubMed PMID: 18462274; PubMed Central PMCID: PMCPMC2810865.

61. Salpea KD, Nicaud V, Tiret L, Talmud PJ, Humphries SE. The association of telomere length with paternal history of premature myocardial infarction in the European Atherosclerosis Research Study II. J Mol Med (Berl). 2008;86(7):815-24. Epub 2008/04/17. doi: 10.1007/s00109-008-0347-x. PubMed PMID: 18414821; PubMed Central PMCID: PMCPMC2480609.

62. Ehrlenbach S, Willeit P, Kiechl S, Willeit J, Reindl M, Schanda K, et al. Influences on the reduction of relative telomere length over 10 years in the population-based Bruneck Study: introduction of a well-controlled high-throughput assay. International Journal of Epidemiology. 2009;38(6):1725-34. doi: 10.1093/ije/dyp273.

63. Pavesi E, Avondo F, Aspesi A, Quarello P, Rocci A, Vimercati C, et al. Analysis of telomeres in peripheral blood cells from patients with bone marrow failure. Pediatric Blood & Cancer. 2009;53(3):411-6. doi: 10.1002/pbc.22107.

64. Imam T, Jitratkosol MHJ, Soudeyns H, Sattha B, Gadawski I, Maan E, et al. Leukocyte Telomere Length in HIV-Infected Pregnant Women Treated With Antiretroviral Drugs During Pregnancy and Their Uninfected Infants. JAIDS Journal of Acquired Immune Deficiency Syndromes. 2012;60(5):495-502. doi: 10.1097/QAI.0b013e31825aa89c. PubMed PMID: 00126334-201208150-00008.

65. Zanet DL, Saberi S, Oliveira L, Sattha B, Gadawski I, Côté HCF. Blood and Dried Blood Spot Telomere Length Measurement by qPCR: Assay Considerations. PLOS ONE. 2013;8(2):e57787. doi: 10.1371/journal.pone.0057787.

66. Jodczyk S, Pearson JF, Aitchison A, Miller AL, Hampton MB, Kennedy MA. Telomere length measurement on the Roche LightCycler 480 Platform. Genet Test Mol Biomarkers. 2015;19(2):63-8. Epub 2014/12/24. doi: 10.1089/gtmb.2014.0208. PubMed PMID: 25535668.

67. Panero J, O'Callaghan N, Fenech M, Slavutsky I. Absolute qPCR for measuring telomere length in bone marrow samples of plasma cell disorders. Molecular biotechnology. 2015;57(2):155-9.

68. Hsieh A, Saberi S, Ajaykumar A, Hukezalie K, Gadawski I, Sattha B, et al. Optimization of a relative telomere length assay by monochromatic multiplex real-time quantitative PCR on the LightCycler 480: sources of variability and quality control considerations. The Journal of Molecular Diagnostics. 2016;18(3):425-37.

69. Lynch SM, Peek MK, Mitra N, Ravichandran K, Branas C, Spangler E, et al. Race, Ethnicity, Psychosocial Factors, and Telomere Length in a Multicenter Setting. PLOS ONE. 2016;11(1):e0146723. doi: 10.1371/journal.pone.0146723.

70. Wand T, Fang M, Chen C, Hardy N, McCoy J, Dumitriu B, et al. Telomere content measurement in human hematopoietic cells: Comparative analysis of qPCR and Flow-FISH techniques. Cytometry A. 2016;89(10):914-21. Epub 2016/10/22. doi: 10.1002/cyto.a.22982. PubMed PMID: 27717244; PubMed Central PMCID: PMCPMC6482817.

71. Behrens Y, Thomay K, Hagedorn M, Ebersold J, Henrich L, Nustede R, et al. Comparison of different methods for telomere length measurement in whole blood and blood cell subsets: recommendations for telomere length measurement in hematological diseases. Genes, Chromosomes and Cancer. 2017;56(9):700-8.

72. Khincha P, Dagnall C, Hicks B, Jones K, Aviv A, Kimura M, et al. Correlation of leukocyte telomere length measurement methods in patients with dyskeratosis congenita and in their unaffected relatives. International journal of molecular sciences. 2017;18(8):1765.

73. Lee M, Napier C, Yang S, Arthur J, Reddel R, Pickett H. Comparative analysis of whole genome sequencing-based telomere length measurement techniques. Methods. 2017;114:4-15.

74. Tarik M, Ramakrishnan L, Sachdev H, Tandon N, Roy A, Bhargava S, et al. Validation of quantitative polymerase chain reaction with Southern blot method for telomere length analysis. Future Science OA. 2018;4(4):FSO282.

75. Wang Y, Savage S, Alsaggaf R, Aubert G, Dagnall C, Spellman S, et al. Telomere Length Calibration from qPCR Measurement: Limitations of Current Method. Cells. 2018;7(11):183. doi: 10.3390/cells7110183. PubMed PMID: 30352968.

76. Ferreira M, Kirschner M, Halfmeyer I, Estrada N, Xicoy B, Isfort S, et al. Comparison of flow-FISH and MM-qPCR telomere length assessment techniques for the screening of telomeropathies. Annals of the New York Academy of Sciences. 2019:1-11.

77. Ropio J, Chebly A, Ferrer J, Prochazkova-Carlotti M, Idrissi Y, Azzi-Martin Z, et al. Reliable blood cancer cells' telomere length evaluation by qPCR. Cancer Med. 2020;9(9):3153-62. Epub 2020/03/07. doi: 10.1002/cam4.2816. PubMed PMID: 32142223; PubMed Central PMCID: PMCPMC7196062.

78. Ropio J, Chebly A, Ferrer J, Prochazkova-Carlotti M, Idrissi Y, Azzi-Martin L, et al. Reliable blood cancer cells' telomere length evaluation by qPCR. Cancer Medicine. 2020.

79. Verhulst S. Improving comparability between qPCR-based telomere studies. Molecular Ecology Resources. 2020;20(1):11-3. doi: 10.1111/1755-0998.13114.

80. Verhulst S, Susser E, Factor-Litvak P, Simons M, Benetos A, Steenstrup T, et al. Commentary: The reliability of telomere length measurements. International Journal of Epidemiology. 2015;44(5):1683-6. doi: 10.1093/ije/dyv166.

81. Benetos A, Kark J, Toupance S, Verhulst S, Aviv A. Response by Benetos et al to letter regarding article,“Short leukocyte telomere length precedes clinical expression of atherosclerosis: the blood-and-muscle model”. Circulation research. 2018;122(8):e73-e4.

82. Benetos A, Verhulst S, Labat C, Lai TP, Girerd N, Toupance S, et al. Telomere length tracking in children and their parents: implications for adult onset diseases. Faseb j. 2019;33(12):14248-53. Epub 2019/10/28. doi: 10.1096/fj.201901275R. PubMed PMID: 31652401; PubMed Central PMCID: PMCPMC6894096.

83. Dlouha D, Maluskova J, Kralova Lesna I, Lanska V, Hubacek J. Comparison of the relative telomere length measured in leukocytes and eleven different human tissues. Physiol Res. 2014;63 Suppl 3:S343-50. Epub 2014/11/28. PubMed PMID: 25428739.

84. Dagnall C, Hicks B, Teshome K, Hutchinson A, Gadalla S, Khincha P, et al. Effect of pre-analytic variables on the reproducibility of qPCR relative telomere length measurement. PLOS ONE. 2017;12(9):e0184098. doi: 10.1371/journal.pone.0184098.

85. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, et al. Minimum information about a microarray experiment (MIAME)—toward standards for microarray data. Nature genetics. 2001;29(4):365-71.

86. Taylor C, Paton N, Lilley K, Binz P, Julian R, Jones A, et al. The minimum information about a proteomics experiment (MIAPE). Nature biotechnology. 2007;25(8):887-93.

87. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, et al. The minimum information about a genome sequence (MIGS) specification. Nature biotechnology. 2008;26(5):541-7.

88. Taylor C, Field D, Sansone S, Aerts J, Apweiler R, Ashburner M, et al. Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. Nature biotechnology. 2008;26(8):889-96.

89. Fiehn O, Robertson D, Griffin J, van der Werf M, Nikolau B, Morrison N, et al. The metabolomics standards initiative (MSI). Metabolomics. 2007;3(3):175-8.

90. Brazma A. Minimum information about a microarray experiment (MIAME)–successes, failures, challenges. TheScientificWorldJournal. 2009;9.

91. Kimura M, Stone R, Hunt S, Skurnick J, Lu X, Cao X, et al. Measurement of telomere length by the Southern blot analysis of terminal restriction fragment lengths. Nature protocols. 2010;5(9):1596-607.

Supplemental Table 1. Telomere Research Network Reporting Guidelines

Reporting items of the TRN Guidelines by category, with assigned importance for each item and comparison with the related assigned importance of the MIQE guidelines. E: essential information, should be submitted with the manuscript. D: desired information, should be submitted if available.

Category (score)	Reporting Item	Importance	MIQE Importance
Sample type, storage, DNA extraction and integrity (score out of 9 or 10)	Sample type	E	E
	Sample storage temperature	E	E
	Sample storage time before extraction	D	E
	Sample storage buffer	D	E
	DNA extraction method	E	E
	DNA storage conditions	E	N/A
	DNA freeze-thaw cycles	E	N/A

	Method of documenting DNA quality/integrity	D	E
	% of samples tested for DNA quality/integrity	D	N/A
	**For studies with repeated measures design report the above for all time points*	E	N/A
qPCR assay (score out of 13 or 15)	State type of PCR method	E	N/A
	PCR machine	D	E
	Source of master mix & reagents	E	E
	Final reaction volume	E	E
	Telomere primer sequences	E	E
	Telomere primer concentration	E	N/A
	Single copy gene name	D	N/A
	Single gene primer sequence	E	E
	Singe gene primer concentration	E	N/A
	Full PCR program description including temperature, times, and cycle numbers	E	E
	PCR efficiency of single copy gene and telomere primers	E	E
	Source of control samples	E	N/A
	Concentration of DNA standard	E	N/A
	For aTL PCR measurement only: sequence of oligo standards	E	E
	For aTL PCR measurement only: concentration of oligo standards	E	N/A
Data analysis (score out of 12 or 13)	Mean and standard deviation or median and range of telomere lengths	D	N/A
	Number of sample replicates	E	E
	Level of independence of the replicates (plate vs day vs extraction)	E	E
	Analytic method, considering replicate measurements, to determine final TL	E	N/A
	Method of accounting for variation between sample replicates	E	N/A
	Method of accounting for well position effects	E	N/A
	Method of accounting for between-plate effects	E	N/A
	% of samples repeated due to failed QC	D	E
	% of samples excluded from analysis due to failed QC	E	E
	Acceptable range of PCR efficiency for primers	D	N/A
	ICCs of sample/study groups	E	N/A
	T/S ratio transformed to Z score prior to analysis	D	N/A
	For studies with family samples or repeated measures design: analytic method to account for this	E	N/A

Supplemental Table 2. Morinha Reporting Guidelines

Reporting items of the guidelines developed by Morinha et al. by category, with assigned importance for each item and comparison with the related assigned importance of the MIQE guidelines. A: very important. B: somewhat important. C: somewhat important, should be submitted if available. E: essential information, should be submitted with the manuscript. D: desired information, should be submitted if available.

Category (score)	Reporting Item	Significance	MIQE Importance
Sample (score out of 5)	Experimental and control groups characteristics	A	E
	Tissue sampled	A	E
	Volume/mass of sample processed	B	D
	Storage conditions in the field and lab (including freeze-thaw cycles)	A	E
	Storage time before DNA extraction	B	E
DNA extraction (score out of 7)	Name of kit and details of any modifications	A	E
	Procedure and/or instrumentation	A	E
	Extraction method of the calibrator sample (if applicable)	A	N/A
	DNA quantification method	B	E
	DNA quality and purity (integrity, yield, 260/280 and 260/230 ratios)	A	D
	Storage conditions	A	N/A
	Freeze-thaw cycles before qPCR	B	N/A
qPCR validation (score out of 7)	Calibration curves with slope and y-intercept	A	E
	R2 of calibration curve	A	E
	qPCR efficiencies (telomere and reference genes)	A	E
	Cq of the NTC	B	E
	Linear dynamic range	B	E
	Specificity (e.g. gel, sequence, melt)	A	E
	Amplification and melting curve plots	B	N/A
qPCR protocol (score out of 12 or 13)	Primer sequences	A	E
	PCR Master mix (manufacturer and name of the products used in the assay)	A	E
	Complete reaction conditions (dNTP, MG2+, primer and polymerase concentrations, DNA amount, other components, reaction volume	A	E
	Additives (SYBR Green I, ROX, DMSO, etc.)	A	E
	Calibrator sample used (e.g. synthetic, pool)	A	N/A
	Singleplex or multiplex	A	N/A
	If singleplex, were both reactions run on the same plate or not	A	N/A
	Number of replicates (technical and biological)	A	E

	Were different groups (age/sex/etc.) run on the same plates or randomized	A	N/A
	Manufacturer of plates/tubes and catalog number	C	D
	Complete thermocycling parameters	A	E
	Reaction setup (manual/robotic)	C	D
	qPCR instrument	A	E
Data analysis (score out of 13)	Quality control steps for data	A	E
	Analysis program (source, version)	A	E
	Method used in the data analysis	A	E
	Choice of reference genes	B	E
	Concentrations of the calibrator sample used to create standard curves	A	N/A
	Normalization and adjustment methods	A	E
	Repeatability (intra-assay variation)	A	E
	Reproducibility (inter-assay variation statistics)	A	D
	Acceptance and rejection criteria	A	E
	How was telomere length calculated?	A	N/A
	Statistical methods for results significance	A	E
	Software used for statistical analysis (source, version)	A	E
	Cq or raw data submission	B	D

Supplemental Table 3. Reporting guidelines rubric

For all reporting guidelines:	▪ Record information as reported only if it is provided directly in the paper itself or in the 1st order of references cited (with exceptions related to Cawthon papers – see below). Do not include information that extends beyond one previous citation. For example, if the methods section states that the detailed methods are described in a previous paper, report only the information from that cited paper. If the referenced paper itself has an additional reference for methods, mark this information as not provided. o Applies for both methods and sample demographics (i.e. if further analysis is done on samples from a previously published study) ▪ Methods sections that state "followed Cawthon 2002 and/or 2009" and failing to provide any additional information about the specific assay performed are not counted as reporting assay parameters and conditions ▪ If sample demographics for the overall population in the study are provided, and only a subset of samples were involved in cross method analyses, reviewer can consider the demographic characteristics as present for the cross-method sample ▪ Information in supplemental material or supplemental methods is reported as present. ▪ qPCR efficiencies in qPCR assay/validation section: this refers to actual efficiencies of the assays performed
TRN reporting guidelines:	▪ Studies without repeated measures design: do not include "for studies with repeated measures…." in score for sample type, storage, extraction, and integrity (mark N/A) or "method of accounting for within-family samples or repeated measures design" if it does not apply (mark N/A, exclude from score) ▪ Studies without aTL measurement: mark N/A for "for aTL PCR measurement…" and do not include in score ▪ DNA quality and integrity: report as yes if methods section describe how DNA quality was assessed; actual DNA values not required (note this differs from the Morinha requirement) ▪ Acceptable range of PCR efficiency: Note this differs from actual PCR efficiency values in the experiment, and should express the lab's typical acceptability/exclusion criteria for PCR assay efficiency
Morinha reporting guidelines:	▪ Experimental and control group characteristics: if cross-method analysis is performed on only a subset of samples, report this as present only if characteristics are provided specifically for the subset of samples (not just the entire experiment) ▪ DNA quality and purity: report yes only if actual DNA integrity, yield, 260/280, 260/230 ratios are provided (ranges/mean are acceptable) ▪ Choice of reference genes: Name of single-copy gene

Supplemental Table 4. Individual paper results of reporting guidelines

Scores for each category are provided as a percentage of items in that category sufficiently reported. A total score for each set of reporting guidelines are calculated from all items as well as an average overall score between the two reporting guidelines.

Author	TRN Reporting Guidelines				Morinha Reporting Guidelines						***Average Overall Grade***
	Sample	qPCR assay	Data analysis	***Overall Grade***	Sample	DNA extraction	qPCR validation	qPCR protocol	Data analysis	***Overall grade***	
Hsieh	44%	92%	58%	68%	40%	71%	71%	85%	92%	78%	**73%**
Eisenberg	67%	85%	58%	71%	40%	86%	29%	83%	85%	70%	**71%**
Zanet	56%	92%	50%	68%	80%	29%	43%	77%	92%	69%	**68%**
Martin-Ruiz (b)	50%	92%	77%	75%	40%	57%	0%	62%	100%	60%	**68%**
Tarik	56%	92%	58%	71%	40%	29%	71%	75%	77%	64%	**67%**
Jodcyzk	67%	92%	58%	74%	40%	43%	0%	69%	77%	53%	**63%**
Gutierrez-Rodrigues	78%	77%	42%	65%	60%	57%	0%	77%	77%	60%	**62%**
Ropio	67%	67%	33%	56%	80%	71%	0%	69%	85%	64%	**60%**
Imam	33%	77%	33%	50%	80%	29%	14%	77%	69%	58%	**54%**
Ehrlenbach	30%	46%	69%	53%	20%	43%	0%	62%	83%	49%	**51%**
Aviv	33%	85%	25%	50%	40%	14%	0%	69%	69%	47%	**48%**
Gadalla	56%	69%	25%	50%	60%	29%	0%	54%	54%	42%	**46%**
Khincha	22%	69%	42%	47%	40%	14%	0%	54%	69%	42%	**45%**
Wang	22%	69%	42%	47%	40%	14%	0%	46%	54%	36%	**41%**
Panero	11%	73%	18%	40%	40%	0%	43%	77%	31%	42%	**41%**
Behrens	33%	62%	33%	44%	40%	29%	0%	38%	62%	38%	**41%**
Lee	11%	69%	17%	35%	40%	0%	0%	69%	38%	36%	**35%**
Pavesi	11%	77%	8%	35%	40%	0%	0%	54%	46%	33%	**34%**
Wand	11%	46%	25%	29%	40%	29%	0%	38%	54%	36%	**32%**
Ventura Ferreira	22%	31%	25%	26%	40%	29%	0%	25%	46%	29%	**28%**
Gardner	44%	92%	25%	56%	40%	43%	0%	77%	77%	56%	**56%**
Hunt	11%	92%	33%	50%	40%	0%	0%	85%	85%	53%	**52%**
Martin-Ruiz (a)	22%	85%	33%	50%	20%	33%	0%	69%	31%	36%	**43%**
Salpea	22%	92%	50%	56%	40%	29%	71%	69%	85%	60%	**58%**
Lynch	33%	46%	42%	41%	80%	29%	0%	78%	69%	49%	**45%**
Average per Category	**37%**	**75%**	**39%**	**52%**	**46%**	**32%**	**14%**	**66%**	**68%**	**50%**	**51%**

Supplemental Methods. Calculation of the repeatability (ICC) of telomere length measures:

Calculation of the repeatability (ICC) of telomere length measures

Dan Eisenberg (dtae@uw.edu), Daniel Nettle (daniel.nettle@newcastle.ac.uk), and Simon Verhulst (s.verhulst@rug.nl)

Measurement repeatability is a critical component of assessing whether any measurement is reliable and for discerning statistical power to detect associations with TL (Nettle *et al*, 2019). In the past, TL measurement repeatability has typically been assessed using the coefficient of variation (standard deviation divided by the mean). However, for multiple reasons, the coefficient of variation (CV) is an invalid statistic for TL measurement and should not be used (Verhulst *et al*., 2015; Eisenberg, 2016; Verhulst *et al*., 2016). Instead, we urge the use of the repeatability, also known as intra-class correlation coefficient (ICC) for discerning assay repeatability. To this end, we provide detailed instructions how to calculate the ICC using R, which is freely available, and an example data set.

Note that measurement validity is determined by both precision (the closeness of measurements to each other) and accuracy (the closeness of measurements to a specific 'true' value). The ICC is informative on precision, but provides no information on the accuracy.

Unlike the CV statistics, the ICC cannot be calculated for each biological sample individually, but instead is calculated for a set of biological samples. Thus, to be able to calculate the repeatability, (a subset of) samples should be re-measured one or more times in an identical fashion as TL is typically measured in one's laboratory. Two important considerations when selecting these samples are the following. Firstly, variation in TL measurement can arise at any point between sample collection and arrival at your TL estimate. Therefore, the ICC will be overestimated when, for example, the ICC is estimated over repeated measurements of extracted DNA relative to measurements where the DNA extraction is also repeated (but by how much is not known). Thus, the ICC you estimate from the data will be closer to the true ICC when more of the measurement process is repeated independently in the successive measurements. Secondly, the TL of the subset of samples needs to be representative of the complete set of samples that is analysed. This is true in particular for the range of TL in the sample – having a larger range in the subset than in the ultimate sample will overestimate the 'true' ICC, while having a smaller range in the subset than in the ultimate sample will result in underestimation of the ICC.

The ICC can be calculated in different ways, depending on how variation between batches is accounted for (i.e. not at all, as random effect, or as fixed effect), where 'batches' can be thought of as different plates, or gels, or measurement sessions that differ in any respect (e.g. day, person, equipment, laboratory). The best way to account for batch in the ICC calculation is to use the same approach as in the ultimate analyses in which hypotheses are tested.

The text below is an R script, mixing instructions with the actual script. When a line is precede by '#', this indicates it is a comment – and it will not be executed. It is left in the text here to make it possible to copy all text below (up to the References) to an R-script – see instructions below.

```
# R script to calculate the ICC (IntraClass Correlation), also known as 'repeatability'.
# The text below assumes you are new to R....
# '#' before text indicates it is a comment - will not be executed
# The other lines need to be 'run' and results will show in the console window.

# Before you start
# 1. it is advisable to do the analysis in Rstudio (freely downloadable and works on
# multiple platforms), which serves as a 'shell' to R, and copy this text to a new R-script.
# 2. A useful introduction by the authors of the package used in this script to calculate
```

```
# the ICC is recommended reading:
# https://cran.r-project.org/web/packages/rptR/vignettes/rptR.html
# The text below is only enough to get you started
# 3. Data format
# The data need to be in the 'long' format. This implies ALL the telomere estimates are
# in one column, with sample identity in another column. Data may have been collected in
# different 'batches' (plates, gels, days, labs) and batch identity is coded in a separate
# column.
# When your data is in wide format, this can easily be changed to a long format in Excel.
# So the data file usually has a minimum of three columns (with variable names in brackets):
# sample identity (id), telomere estimate (TL), batch (batch) [when you use different
# variable names, the names in the code below need to be replaced with those names].

# The packages listed below are needed and you will probably need to install them first.
# At the top of the bottom right panel in RStudio there is a tab 'packages' you can use.
library(readxl) #for when your data are in Excel format
library(MASS)
library(lme4)
library(rptR)

# To clear lists of objects – useful to run whenever you start an analysis:
rm(list=ls())
rm(list = ls(all = TRUE))

# Reading in the data.
# In the example below, the data were stored in Excel, but Rstudio reads many formats.
# Other formats may require loading another package - Rstudio will tell you this.

# Note that the first bit of the code below, 'd <-', you can read as 'd becomes'.
# We here arbitrarily name the data set 'd' (commands / names in R are case sensitive!).

# You can get the import code for your file location and correct command for your file format
# using the 'import dataset' tab in RStudio (above the top right frame on a mac).

# When you import a data set using RStudio, the dataset will have a name different from
# what is in the script below. I recommend copying the code you see after the "<-" to the
# script just before actually importing the data and copy this below after the text "d <-".
# Alternatively, you replace 'd' in the code with the name you have given your data set.

d <- read_excel("FILE LOCATION/example.xls") #importing the example data from Excel file
# The file 'example.xls' is available on the telomere network site.

# Calculating the ICC
# The ICC can be calculated in different ways, depending on how variation between
# batches is accounted for (i.e. not at all, as random effect, or as fixed effect).
# The best way to account for batch in the ICC calculation is to use the same approach
```

```
# as in the analysis for which the data were collected.
# Below is the code for different ways to include 'batch' in the analysis.

# When running the scripts below, depending on details,
# there may be 'Singularity' issues that are reported as errors.
# You can safely ignore these (see information on rptR package for details.)

# 1. No correction for batch
# Including 'id' only - i.e. batch is not in the model
# Note that the (1|something) codes for a random intercept for levels of 'something'
rpt(TL ~ (1|id), grname = "id", data=d, datatype = "Gaussian", nboot = 1000, npermut=0)

# 2. Including 'batch' as random effect
rpt(TL ~ (1|id) + (1|batch), grname = "id", data=d, datatype = "Gaussian", nboot = 1000,
npermut=0)

# 3. Including 'batch' as fixed effect
rpt(TL ~ batch + (1|id), grname = "id", data=d, datatype = "Gaussian", nboot = 1000, npermut=0)

# Note that these models can be extended with other factors and covariates.
# See example below that includes age (as fixed effect)
rpt(TL ~ age + (1|id) + (1|batch), grname = "id", data=d, datatype = "Gaussian", nboot = 1000,
npermut=0)
# When running the model including age, the ICC is likely to become lower. The new (lower)
# ICC estimate is the more relevant estimate. This is so, because you will probably
# take age into account in your statistical analyses also. The remaining variation in the
# data will decrease as a consequence, and the ICC is calculated over the remaining variation.
# Adding other factors to the model, e.g. batch identity or procedure characteristics
# that you also include in the model with which you test hypotheses can also be added as
# factors to the model, and may increase the ICC.

# The script above assumes the TL measurements are normally distributed, but the
# rptR package can handle other distributions.

# Extrapolated repeatability – an important extension!
# Suppose your protocol includes rTL measurement in duplicate, using for example two plates
# with samples in triplicate on each plate, and you use the average of the two plates in
# the data analysis. You can then calculate the ICC over the two plates, but this will
# underestimate the ICC of the average of the two plates, which is based on more measurements.
# However, the ICC of the mean of the two plates can be calculated with a simple equation
# once the ICC over the two plates is known: when r is your repeatability (ICC) estimate,
# and n is the number of replicate measurements then the extrapolated repeatability (re):
# re = r / (r+1/n*(1-r))
# This is equation 37 in: Nakagawa, S. & Schielzeth 2010, Biological Reviews 85, 935-956.
```

Supplemental References

Eisenberg, D. T. (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, 45(4), 1295-1298. doi:10.1093/ije/dyw191

Lindrose, A.R., McLester-Davis. L.W.Y., Tristano, R.I., Kataria, L., Gadalla, S.M., Eisenberg, D.T.A., Verhulst, S., & Drury, S. Method comparison studies of telomere length measurement using qPCR approaches: a critical appraisal of the literature. bioRxiv preprint doi: https://doi.org/10.1101/2020.09.04.282632.

Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biological Reviews, 85(4), 935-956. doi:10.1111/j.1469-185X.2010.00141.x

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. doi:10.1371/journal.pone.0216118

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. doi:10.1093/ije/dyw194

Verhulst, S., Susser, E., Factor-Litvak, P. R., Simons, M. J., Benetos, A., Steenstrup, T., Aviv, A. (2015). Commentary: The reliability of telomere length measurements. International Journal of Epidemiology, 44(5), 1683-1686. doi:10.1093/ije/dyv166

CHAPTER 2

A review and meta-analysis: Cross-tissue telomere length correlations in healthy humans

Background:

Over the last four decades there has been an increasing interest in the role of telomeres in human health outcomes (Lindrose et al., 2021). Telomeres were first identified in 1938, with the first telomere sequenced in *Tetrahymena* in 1978, and the human telomere sequenced in 1988 (Muller, 1938; Blackburn et al., 1978; Moyzis et al., 1988). Telomeres are the protective nucleic acid and protein cap at the end of all eukaryotic chromosomes. The telomeric DNA sequence of 5'-TTAGGG-3' repeats are highly conserved across eukaryotic species and the telomere protects linear DNA from damage and preserves chromosome stability. Progressive shortening of the telomere sequence occurs in dividing cells as a result of incomplete lagging-strand DNA synthesis, DNA damage, and other factors. Since 1988, measuring the length of human telomeres has been of interest throughout diverse scientific disciplines.

Telomere length (TL) has been associated with various exposures, biological processes, and health outcomes. Studies across the human lifespan report TL as a biomarker of aging, and TL is increasingly recognized as a biomarker of environmental exposures and stress. Additionally, TL is a predictor of aging-related health outcomes including obesity, cardiovascular disease, and overall mortality (Mundstock et al., 2015; Khosravaniardakani et al., 2022; Haycock et al., 2014; Zhan and Hägg, 2019; De Meyer et al., 2018; Wang et al., 2018; Schneider et al., 2022). While there is significant convergence of these findings, conflicting reports exist and are hypothesized to result from key variances in study design, including age, TL measurement methodology, sample size, and tissue source. Information on the effects of the age of participants on TL associations can be found in other publications (Steenstrup et al., 2017;

Dorado-Correa et al., 2018). Similarly, information on the effects of TL measurement methodology is summarized elsewhere (Horn et al., 2010; Haussmann et al., 2011; Lin et al., 2019). For guidelines regarding sample size in TL association studies refer to Lindrose et al., 2021. The focus of this meta-analysis is to examine the influence of tissue selected for TL measurement as a source of variation in cross-study findings.

As TL is measurable from any cell containing DNA, published studies have varied in the source tissue for TL measurement. Studies of small to moderate size often focused on individual health outcomes, such as case-control studies, usually select source tissues specific to the health outcome of interest, such as biopsy samples. In contrast, population-based studies, such as large observational cohorts, tend to use minimally invasive tissues, such as peripheral blood, saliva, and / or buccal epithelial cells given practical constraints. Using TL measured from one tissue as a proxy for the tissue of interest assumes similarities of telomere dynamics across tissues. However, certain differences in telomere dynamics are necessary for specific tissues to function, for example TL differences between somatic and germ cells, meaning not all proxy tissues may be equally appropriate for TL measurement (Rollings et al., 2019).

An illustration of the effect of source tissue selection is seen in a meta-analysis on the association of TL and depression (Ridout et al., 2016). In this meta-analysis, 35 studies measured TL in a peripherally available tissue, specifically leukocytes or saliva, while the remaining 3 studies measured TL in various brain tissues. Overall, this meta-analysis reported tissue type as a significant moderator of the association between depression and TL. For example, in a case-control study included in the meta-analysis, shorter TL measured in leukocytes was associated with diagnosed major depressive disorder (Garcia-Rizo et al., 2013). However, a study utilizing

post-mortem occipital cortex tissue failed to detect an association between shorter TL and depression (Teyssier et al., 2011; Ridout et al., 2016).

While the effects of depression on biological aging was detectable in peripheral tissues, this is not consistent for other outcomes or exposures. In a meta-analysis on the association between TL and prostate cancer, utilization of different human tissues across studies also contributed to variability in results (Hu et al., 2019). In this meta-analysis, studies of small size (less than 100 participants) utilized biopsy samples of prostate stromal and epithelial cells, while larger (more than 100 participants) studies selected peripheral blood for TL measurement. However, the opposite trend in relation to source tissue and health outcome was observed. Specifically, meta-analytic findings supported a predictive relation between TL measured in stromal and epithelial cells and prostate cancer, however this association was not present when TL was measured from peripheral blood, even though these studies had more participants. Bias to publish significant effects seen in the smaller studies was not reported in the meta-analysis, suggesting that the more disease-relevant tissue source (e.g., prostate tissue) was more critical for the association than sample size.

A more recent study reported on the relative TL for multiple post-mortem tissues within 952 humans providing the largest study to date of TL measured across tissues (Demanelis et al., 2020). In this study, the correlation of TL measured from prostate tissue and whole blood (n=138 healthy individuals) was 0.29 (p=0.0016), suggesting that although TL is correlated between prostate tissue and whole blood, the correlation could be stronger (Demanelis et al., 2020). Defining the relation of TL measured in normal tissues is a needed first step before extrapolations to diseased tissues can occur. As such, this meta-analytic review is focused on the correlations between TL measured in different tissues in healthy human subjects. The results of

this meta-analysis are expected to provide guidance on the ability to rely on measurement of TL in one tissue to infer TL from another tissue to inform designs of future TL studies. The importance of reporting sample size, TL measurement methods, and demographics of participants in future TL studies are also discussed.

Aim:

To qualitatively and meta-analytically review literature examining correlations of TL measured from multiple tissues from the same individual. Secondary analyses of meta-correlations by sample size, quality assessment, telomere measurement methodology, age group of participants, and sex of participants were performed to identify the role of these methodological factors on the correlation of TL measured in different tissues within an individual.

Methodology:

Study Inclusion Criteria:

Inclusion criteria for studies included in the systematic review consisted of the following:

1) The study was published after 1987 or before 2023 in English;

2) The study collected more than one tissue from the same individual;

3) TL was measured in more than one tissue within the same individual using the same TL assay;

4) The tissues were not in a known disease state defined as (a) obtained from a healthy control, (b) presumed healthy living individual, or (c) presumed healthy at time of death prior to sample collection, according to WHO definition; and

5) The study reported statistical and demographic information like sample size and sample age range, sex, and racial demographics.

If the study met the above inclusion criteria and 6) the study reported a quantitative correlation of TL for more than one tissue, the study was included in the meta-analysis. If the study met inclusion criteria one through five, but did not include a quantitative correlation, corresponding authors were contacted via email for this additional information. If the corresponding author responded and provided the correlation, the data was included. If the study did not report a quantitative cross-tissue TL correlation, the study was included in the qualitative review only.

Information Retrieval:

A scientific librarian (RPH) constructed a search strategy and performed a computerized bibliographic search of PubMed (see Appendix A). This search strategy was adapted to search Embase and Web of Science. The searches were completed on November 21st, 2020. Searches were restricted to English results in humans from January 1st, 1988 to December 31st, 2022. The following keywords were employed: “telomere length” and “tissues” or “tissue.” Appendix A describes all terms included in the PubMed, Embase, and Web of Science searches. All full text studies were eligible to enter the meta-analytic workflow, including studies available online, in print, in press, or ahead of publication.

Study Selection Procedures:

All screening and selection of studies were performed in duplicate by two authors (LWYMD, PE, LAK, NAM, JTS, SD, RIT, and / or CVM). Abstract, title, and screening of the methodology sections identified papers containing TL measurements of multiple healthy human tissues. During full text screening, the studies were selected for extraction if they met the above listed inclusion criteria. Any discrepancies were settled by consensus for two or more authors. Following identification of studies for extraction, duplications of studies were defaulted to the

PubMed result and non-duplicate studies reporting any data from the same cohort were minimized to one study with the largest sample size.

Data Extraction:

Information from the included studies was recorded in a prespecified data abstraction spreadsheet in duplicate, see Appendix B. The information from each study was extracted independently by two reviewers and compared for discrepancies which were settled by consensus. In studies where the correlation data for TL measured in different tissues was not reported, the corresponding authors were contacted for data. Final data extracted for included studies encompassed: title; authors; year of publication; sample size; sample demographics including race, sex, age, and country of study; tissue samples analyzed; sample collection conditions; sample storage conditions; TL measurement methodology (e.g., qPCR, TRF, FISH, etc.); statistical methodology of TL correlations; and reported TL correlations between tissues (see Appendix B).

Data Coding and Qualitative Analysis:

Qualitative review of studies emphasized country of origin, sample demographics, telomere measurement methodology, and quality of methodology reporting. Coding for sample size was conducted by categorizing individual pairwise correlations of two tissues within studies. A correlation with an n of 4 to 75 as *small,* correlations with an n of 76 to 258 were categorized as *medium*, and correlations with an n greater than 258 were categorized as *large* in alignment with the histogram distribution of study sample sizes. For the quality assessment, the number of completed items recommended in the TRN Telomere Length Data Quality Assurance Checklist for each study was completed by consensus by a minimum of two reviewers (Lindrose et al., 2021). As some studies did not have all quality assessment criteria apply, e.g. DNA was not

extracted for FISH analysis studies, the number of possible completed criteria in the checklist was adjusted and the quality assessment score was subsequently converted into a percentage value for all studies to permit cross-study comparisons (Appendix C).

Telomere measurement methodology was categorized from the following methods utilized in the studies included in the meta-analysis: quantitative polymerase chain reaction (qPCR), absolute telomere length qPCR (aTL), monochrome multiplex qPCR (MMQPCR), dot blot, flow cytometry fluorescence in situ hybridization (flow-FISH), hybridization protection assay, interphase quantitative fluorescence in situ hybridization (Q-FISH), metaphase Q-FISH, Luminex, Southern Blot / Telomere Restriction Fragment (TRF), single telomere length analysis (STELA), and studies that utilized multiple methods. These various assays were subsequently categorized based on underlying methodology into the following four categories for moderator analysis: *FISH analyses* (Q-FISH, Flow-FISH, interphase Q-FISH, metaphase Q-FISH), *Hybridization analyses* (dot blot, hybridization protection assay, Southern Blot / TRF), *Luminex analyses*, and *PCR analyses* (aTL qPCR, MMQPCR, qPCR). For studies that included cross-tissue TL correlations across multiple assays, individual pairwise correlations were placed into the category corresponding to the specific method for that unique pairwise correlation.

For age group of participants, categories were *lifespan* if the reported age range of participants was greater than a 30-year span, *adult* if the reported median or mean age of participants included 50-years of age, *young adult* if the reported age median or mean age of participants was less than 50-years of age and a minimum age of participants of 18 years, *neonatal* if the reported age of participants was less than 5-years of age, and *unknown/other* if not stated in the study demographics or did not fit into the aforementioned categories. For sex of participants, correlations from all female participants were categorized as *100% female*,

correlations from mixed female and male participants were categorized as *mixed*, correlations from all male participants were categorized as *100% male*, and correlations from participants of unknown sex were categorized as *unknown*. Coding correlations based on the sample types in the pairwise correlation was conducted for sensitivity analysis.

Meta-Analysis and Statistical Approach:

For the primary objective, the correlations, including Pearson product-moment correlation coefficient, Kendall tau rank correlation, and Spearman rank correlation, for each tissue pairwise comparison were collected or computed and analyzed following Fisher's z-transformation by LWYMD and WJH. Correlations with an n of 3 or less were unable to be z scored and were excluded from analysis. To examine effects by moderator variables, the coded variables for the moderators were used in the Assink and Wibbelink approach for sample size, quality assessment, telomere measurement methodology, age group of participants, and sex of participants (Assink and Wibbelink, 2016).

The meta-analytic dataset contained multiple effect sizes within studies making the effects interdependent and requiring a three-level nesting model to cluster the results of individual, participant-level effects in each study (level 1) into effect sizes within studies (level 2) and between studies (level 3). R Studio program was used for all analyses. A multi-level random-effects meta-analysis was utilized to account for dependency in effect sizes and enable the meta-analysis of all eligible effect sizes. Multi-level models were then compared to a traditional random effect model using the ANOVA function to determine best fit. The {metafor} package was used to fit the three-level meta-analysis model, and the {dmetar} and {meta} packages were used for the forest plot and tests for heterogeneity and publication bias. The overall correlation for all tissue TLs was reported as an r value which was converted from the

overall estimated effect. The overall model was fitted using the restricted maximum likelihood estimation method and Knapp and Hartung's adjustment (Röver et al., 2015). A forest plot was constructed from the weighted effect estimates and 95% confidence intervals (CIs).

Moderator Analyses

Moderator analyses were conducted using a three-level mixed-effects model to assess potential moderators of the overall effect. Moderators included sample size (3 categories), quality assessment (continuous variable), telomere measurement methodology (4 categories), the age group of participants (4 categories), and the sex of participants (4 categories). In multi-level moderator analyses with multiple categories (i.e., 3 sample size categories), one category functions as the reference category, and the other categories are compared against the reference category. Large sample size was tested as the reference category for sample size, FISH analyses as the reference category for telomere measurement methodology, neonatal as the reference category for age group, and 100% female as the reference for sex.

Sensitivity Analyses

Two post hoc sensitivity analyses were conducted to investigate the effect of the tissues in which TLs were measured in the meta-correlation. First, pairwise correlations were separated into three categories by the types of samples in the pairwise correlation: a pairwise correlation where the samples were both a type of or found within blood (e.g. B Cells, CD34+CD38- Cells, Cord Blood, etc.), a pairwise correlation where one sample was a type of or found within blood and the other sample was not (e.g. Buccal Epithelium, Cerebellum, Kidney, etc.), or a pairwise correlations where both samples were not a type of or found within blood. In the second post hoc sensitivity analysis, pairwise correlations were separated into three categories based on the availability of a sample given collection measures. The three categories were: both samples

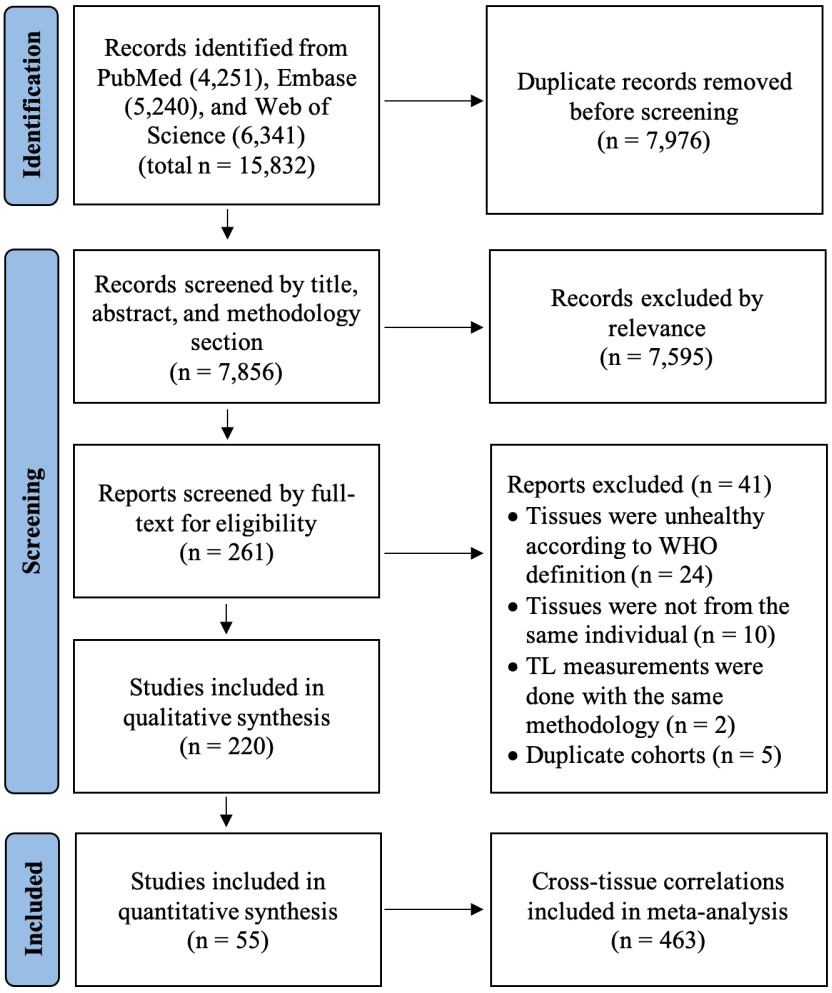


Figure 1. PRISMA Flow Diagram. Template taken from The PRISMA 2020 statement (Page et al., 2020).

being peripherally accessible (e.g. Buccal Epithelium, Peripheral Blood Mononuclear Cells, Skin, etc.), a peripherally accessible sample to a surgically obtainable sample (Bone Marrow, Small Intestine, Spleen, etc.), and a surgically obtainable sample to another surgically obtainable sample. The assigned groups in the sensitivity analyses by sample are listed in Appendix D.

Additional post hoc sensitivity analyses were conducted for the sample size within each category of sample size and TL measurement methodology pairwise correlations following the moderator analyses of sample size and TL measurement methodology. Individual pairwise correlations were categorized as described previously. Three models were completed to assess

the effect of correlations with *small*, *medium*, or *large* sample sizes on the overall effect. Similarly, four models were completed to assess the effect of correlations with only *FISH analyses*, *hybridization analyses*, *PCR analyses*, or *Luminex analyses* on the overall effect.

Assessment of Publication Bias and Heterogeneity:

Assessment for publication bias was carried out by assessing funnel plot asymmetry using data points derived from estimates and standard errors from individual studies in relation to the pooled effect estimate. In addition to visual inspection of the funnel plot for asymmetry, Egger's test was applied to statistically test for publication bias where a p-value of <0.1 was considered to represent significant asymmetry. Assessment of heterogeneity of effect sizes within the meta-analysis was calculated using the Higgins & Thompson's I^2 statistic using a random effects model, where a value of more than 75% indicates high heterogeneity.

Results

Qualitative Review

A total of 15,832 studies were identified (Figure 1). After removal of 7,976 duplicates, reviewers screened the abstracts and methods of 7,856 studies for inclusion and exclusion resulting in the removal of 7,595 articles. Examples of removed studies included meeting abstracts, review articles, animal studies, and studies where TL was measured in only one tissue. Two hundred and sixty-one articles underwent full-text review (Figure 1). From those, 5 studies were condensed to one study as the data was from the same cohort, where individual pairwise correlations were only taken from one study if applicable; another 36 studies which were screen by full-text review were excluded for not meeting inclusion criteria.

A total of 220 articles were included in the qualitative analysis. The studies in the qualitative review were conducted in 37 countries (Table 1). The reporting quality of the TL

measurement methodology ranged from 12.50% to 100% with a mean of 75.04%. The 220 studies in the review used one of four different categories of telomere measurement methodologies. Seventy-five studies (34.09%) utilized a qPCR analysis, 61 studies (27.73%) utilized a hybridization-based analysis, 83 studies (37.73%) utilized a FISH analysis, and 1 study (0.45%) utilized Luminex analysis.

The reporting of demographics differed by study, for example, 3 studies did not report the sample size of the study. The majority of studies (n = 169, 76.82%) had a small sample size (n < 75). Thirty-eight studies (17.27%) had a medium sample size (76 to 258) and ten studies (4.55%) had a large sample size (n > 258). For reporting the age of participants in the study, 20 studies (9.09%) did not report on the age of participants or the age grouping fell outside of the defined categories in this review. Seventeen studies (7.73%) were from neonatal populations, 52 studies (23.64%) had young adult participants, 46 studies (20.91%) had adult participants, and 85 studies (38.63%) had participants across the lifespan.

Over a third of studies (n = 75, 34.09%) in this review did not report the sex of participants in their study. One hundred and three studies had a mix of female and male participants, ranging from 4.17% female to 94.12% female participants. Twenty studies had only female participants and twenty-two studies had only male participants. No studies reported non-binary sex of study participants (Table 1).

Lastly, the reporting of racial /ethnic demographics was only done in 16.82% of the studies. Of the studies (n = 37) that reported racial / ethnic demographics, 10 studies reported only White participants, 2 studies had all Asian participants, 1 study had all Lebanese participants, 4 studies reported a mix White and Other participants, 3 studies reported a mix of Black and White participants, 4 studies reported a mix of Black, White, and Other participants, 1

study reported a mix of Black, White, and Asian participants, 1 study reported a mix of Black, White, and Latino participants, 1 study reported a mix of White, Asian, and American Indian / Alaska Native participants, and the remaining 9 studies reported a mix of participants from four to seven racial / ethnic categories (Black, White, White Latino, Latino, Non-Latino, Asian, Asian or Pacific Islander, American Indian / Alaska Native, Native Hawaiian / Other Pacific Islander, or Other).

Comparison of Studies Only in the Qualitative Review versus Studies Included in the Meta-Analysis

Only 55 studies provided correlations for the meta-analysis from 102 individual tissue types. While all authors not reporting correlations were contacted for cross-tissue TL correlations, only five authors were able to provide a correlation as many authors had moved labs, disposed of data, and / or only retained data that is unmatched between sample types since 1996. From these 55 studies, there were 463 individual pairwise correlations reported from 4,324 unique individuals. Included studies were published since 1996 and were conducted in 20 different countries.

The average quality assessment score for the studies included in the meta-analysis (84.97%) was higher than the average for studies only in the qualitative review (75.73%; Table 2). The 55 studies used one of four different types of telomere measurement methodologies: FISH based analyses, hybridization-based analyses, qPCR based analyses, or Luminex analysis. The 165 studies in the qualitative review used FISH based analyses, hybridization-based analyses, qPCR based analyses, but no Luminex analysis (Table 2).

Three studies in the qualitative review did not report the sample size of their study, while all of the studies in the meta-analysis included sample size. Of the 55 studies in the meta-

analysis, the study sample sizes ranged from 2 to 381 (Table 1). Sixty-seven-point-twenty-seven percent of studies in the meta-analysis had sample sizes less than 75 participants, while 80.00% of studies only in the qualitative review had small sample sizes. Twenty-seven-point-twenty-seven percent of studies in the meta-analysis had a medium sample size (75 to 258) of participants compared to 13.94% of studies in the qualitative review had a medium sample size. Only 4.24% of studies in the qualitative review had a large sample size (n > 258) while 5.46% of studies in the meta-analysis had a large sample size.

More studies in the meta-analysis reported age of participants than studies only included in the qualitative review (Table 2). Similarly, more studies in the meta-analysis reported sex of participants compared to studies only in the qualitative review. Furthermore, more studies in the meta-analysis reported racial / ethnic demographics of participants and reported more racial / ethnic categories.

Overall Models

The overall association for between-tissue TLs was significant ($t(416) = 10.71$, $p < 0.0001$) with a meta-analytic estimate of $z = 0.66$ (95% CI: 0.54, 0.78; Figure 2). The estimate equated to a correlation coefficient is $r = 0.58$. The estimated variance of effect sizes measured between studies at the third level of the model was 87.48%, while the estimated variance between effect sizes within studies was negligible (5.31%) despite significant study publication bias (Figure 3).

Moderator Analyses

Sample Size

The study sample size was investigated as a moderator of the overall association between all tissue TLs because the sample size affects the statistical power to detect an association with

TL, particularly given variability in TL assay precision (Lindrose et al., 2021). The overall association between all tissue TLs was moderated by sample size (i.e., large vs. not large), $F(1,415) = 14.73$, $p < 0.001$. After accounting for sample size, significant residual heterogeneity remained, $QE = 2266.79$, $p < 0.001$ (Table 3). As seen in the funnel plot, there was an increased number of studies published with small sample sizes in this meta-analysis, which is reproduced in the subgroup analysis by sample size (Table 4). While all models across sample sizes were significant ($p < 0.0001$), the *small* group had the largest number of correlations ($n = 253$) and the highest estimate $z = 0.75$ (95% CI: 0.59, 0.91). The correlations of *medium* sample size had $n = 146$ and the overall effect was significant, $z = 0.49$ (95% CI: 0.31, 0.67). The pairwise correlations of *large* sample size had the least number of correlations in this meta-analysis ($n = 17$) with an overall significant effect of $z = 0.45$ (95% CI: 0.30, 0.60).

TRN Quality Assessment

The study quality assessment score was investigated as a moderator of the overall association between all tissue TLs as reported methodological differences between studies within and between methods have moderated meta-analytically associations between TL and health outcomes (Ridout et al., 2016; Hu et al., 2019). The overall association between all tissue TLs was not moderated by the quality assessment of methods in studies, $F(1,415) = 0.28$, $p = 0.59$; there was significant residual heterogeneity after accounting for quality assessment score, $QE = 2080.66$, $p < 0.001$.

Telomere Measurement Methodology

There is a body of evidence that the precision of TL measurement differs by methodology, and studies using methods with greater precision present different effects in meta-analyses and may show differences in cross-tissue correlations compared to methods with lesser

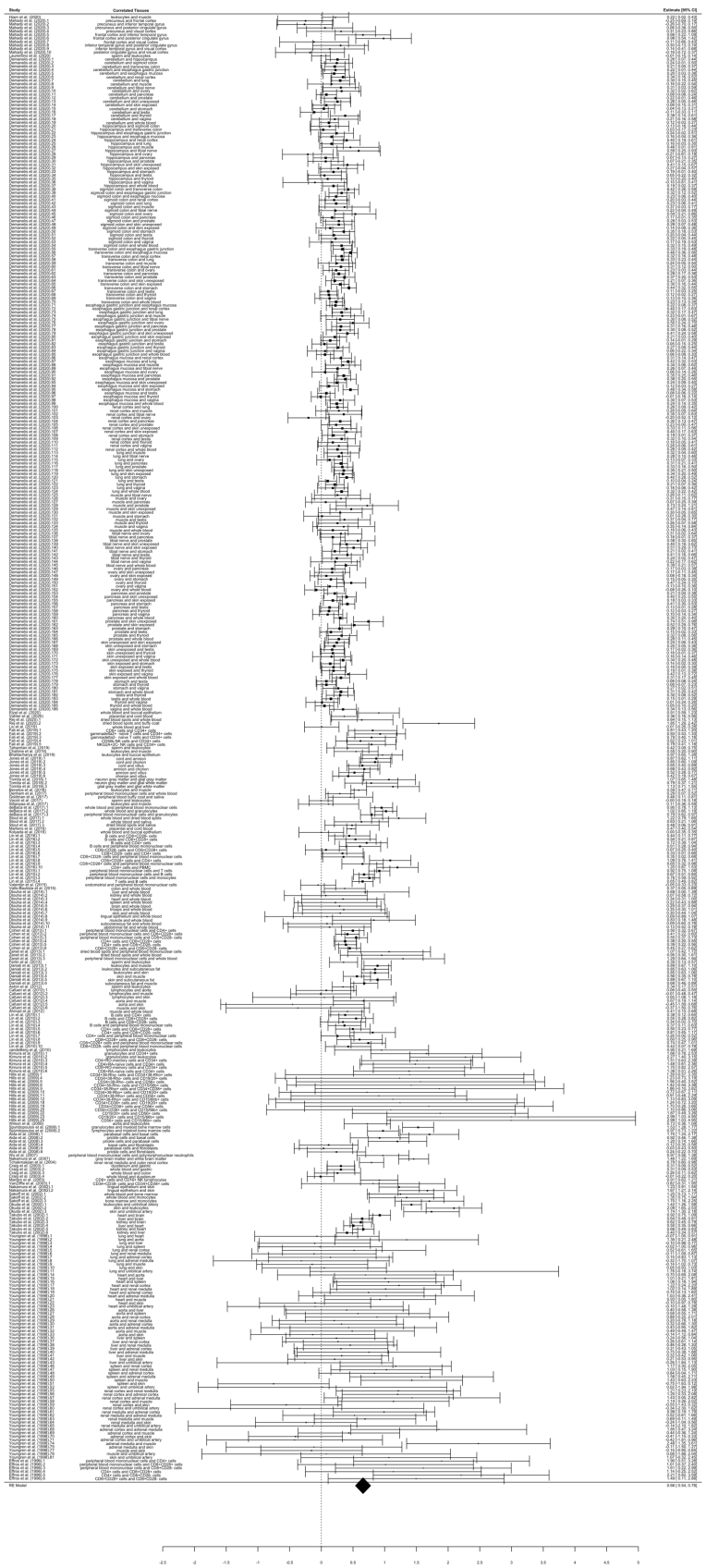


Figure 2. Forest Plot of Correlation Coefficients Between All Tissue Telomere Lengths by Study and Year of Publication. Numbers after the year of publication indicate the number of that correlation collected within the study, i.e. Effros et al. (1996) 6 indicates the sixth correlation collected from this study.

precision (Horn et al., 2010; Haussmann et al., 2011; Lin et al., 2019). In line with previous results, the overall association between all tissue TLs was moderated by telomere measurement methodology when comparing across the four assay categories: FISH vs. hybridization vs. PCR vs. Luminex assays, $F(1,415) = 7.58$, $p = 0.0062$. After accounting for telomere measurement methodology, significant residual heterogeneity remained, $QE = 2180.61$, $p < 0.001$ (Table 4). All models across telomere measurement methodology categories were significant ($p < 0.0001$; Table 4). While the Luminex analyses had the largest number of correlations (n = 186), it had the smallest overall effect, $z = 0.25$ (95% CI: 0.23, 0.27) and narrowest confidence interval. The correlations from hybridization analyses had the second largest sample size (n = 113) and the greatest overall effect, $z = 0.86$ (95% CI: 0.63, 1.10). The PCR analyses had a sample size of 98 with a significant estimate of 0.51 (95% CI:0.37, 0.65), and the FISH analyses had the smallest number of correlations (n = 20) with a significant estimate of 0.77 and the widest confidence interval (95% CI: 0.39, 1.14). Despite difference in the magnitude of the effect as a function of each methodology, meta-analytic findings were still significant when each method was analyzed independently (Table 4).

Age Group of Participants

Given the evidence of age-related decline in TL and that factors contributing to TL shortening (e.g. cellular division, DNA damage, etc.) occur differently across tissues, the age of participants was investigated as a moderator of the overall association between all tissue TLs (Steenstrup et al., 2017; Dorado-Correa et al., 2018). The overall association between all tissue TLs was not moderated by the age group of participants $F(1,398) = 3.10$, $p = 0.08$; there was significant residual heterogeneity after account for the age group of participants $QE = 1790.33$, $p < 0.001$.

Sex of Participants

The sex of participants was investigated as a moderator of the overall association between all tissue TLs as sex has been observed to moderate meta-analytic findings of TL associated with health outcomes, and sex is meta-analytically associated with TL (Wang et al., 2018; Gardner et al., 2014). The overall association between all tissue TLs was not moderated by the sex of participants $F(1,307) = 0.75$, $p = 0.39$; there was significant residual heterogeneity after account for sex of participants, $QE = 1945.05$, $p < 0.001$.

Sensitivity Analyses

To provide insight on the correlation across different cellular components of blood as well as the ability of TL measured in a blood sample to inform TL derived from a non-blood tissue source, sensitivity analysis was performed (Table 5). For the model only including correlations between cell types and samples derived from whole blood ($n = 76$), $z = 0.93$, (95% CI: 0.74, 1.12), the association between cross-tissue TLs correlations was significant with the largest effect ($p < 0.0001$). For the model only including correlations between non-blood samples and a blood sample ($n = 67$), $z = 0.47$, (95% CI: 0.31, 0.63), the association between cross-tissue TLs correlations was significant ($p < 0.0001$) with the smallest effect. Lastly, for the model only including correlations between non-blood and non-blood samples ($n = 274$), $z = 0.74$, (95% CI: 0.51, 0.98), the association between cross-tissue TLs correlations was also significant ($p < 0.0001$).

Additionally, sensitivity analysis was conducted to evaluate the correlation between peripherally available and surgically obtainable samples. The sensitivity analysis tested the correlation between two different peripherally available samples, the correlation between a peripherally available and surgically obtainable samples, and the correlation between two

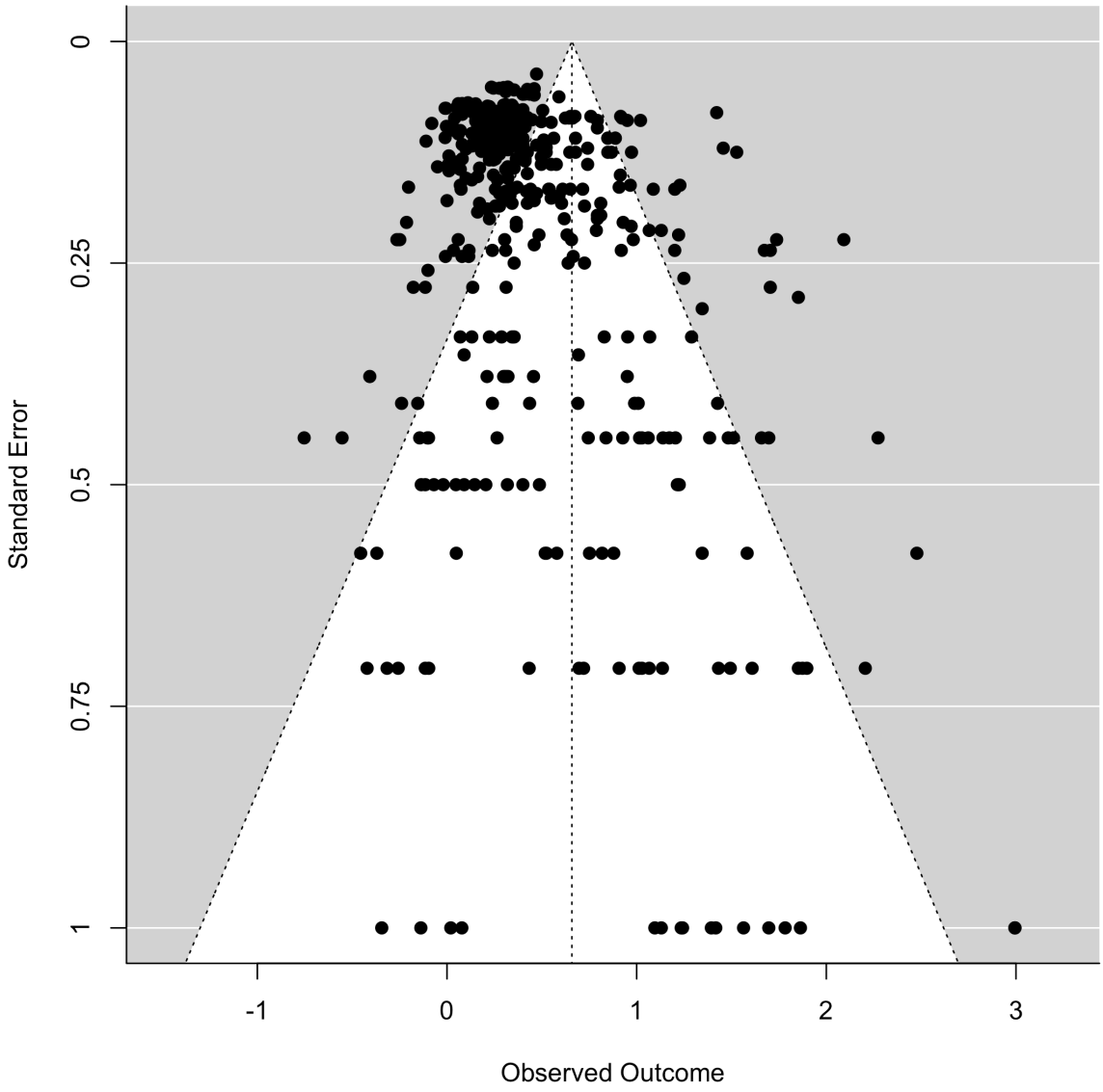


Figure 3. Funnel Plot of Effect Sizes for Correlations. Funnel plot for the included studies reporting the calculated effect sizes (observed outcome) in relation to the calculated standard error.

surgically obtainable samples (Table 6). For all models in this sensitivity analysis, the association between cross-tissue TLs correlations was significant ($p < 0.0001$). For the model with correlations between different surgically obtainable samples (n = 214), the overall effect was z = 0.65 (95% CI: 0.41, 0.88); for the model with correlations between surgically obtainable samples and peripherally available samples (n = 100), the overall effect was lower than for between surgically obtainable samples z = 0.47 (95% CI: 0.26, 0.68); while the highest estimate

was for the model of correlations between different peripherally available tissues (n = 103), where the overall effect was z = 0.76 (95% CI: 0.59, 0.93).

Publication Bias and Heterogeneity

Evaluation of the funnel plot both visually and via Egger's regression test indicated evidence of significant asymmetry in the funnel plot with an intercept of 0.26 (95% CI: 0.23, 0.28) suggesting the presence of publication bias. Specifically, the results suggest increased propensity to publish studies with positive correlations, t(415) = 4.61, p = <0.0001 (Figure 3). The I2 statistic was 81.69%, indicating high heterogeneity and a need for random effects models to be utilized.

Discussion:

The current qualitative review and meta-analysis were conducted to provide guidance on the correlation of TL measurements from different source tissues in healthy individuals across the lifespan. In the 165 studies only included in the qualitative review, most studies (88.49%) did not report on race and ethnicity of participants, more than a third (37.58%) did not report on sex of participants, and 8.33% did not report on details of age of participants. However, studies included in the meta-analysis reported more information on demographics of participants on average compared to studies included in the qualitative review which did not provide a cross-tissue TL correlation. For example, all 55 studies included in the meta-analysis reported the sample size, while 3 studies in the qualitative review did not report sample size.

The meta-analysis conducted here included over four hundred correlations from 4,324 unique individuals across 102 different tissues and demonstrated a significant overall meta-correlation coefficient of 0.58. Furthermore, these results report a significant positive association of cross-tissue TL correlations across the lifespan for both males and females, as age and sex of

participants were not significant moderators of this meta-correlation. In contrast, study level factors, specifically sample size and telomere measurement methodology, were significant moderators of the overall effect estimate.

Consistent with other meta-analyses on TL associations, sample size was a significant moderator of this meta-correlation (Ridout et al., 2016). However, studies with a smaller number of participants had an overall larger estimate than correlations with a larger number of participants. This may be a result of the fact that 54.64% of the correlations in the meta-analysis had small sample sizes. In other words, studies investigating a higher number of pairwise correlations between individuals tended to be those with the smallest number of unique individuals, likely resulting from the increased time and cost of collecting multiple tissue types from each individual participant. The higher effect size may be due to an inadequate adjustment for the multiple pairwise correlations assessed within an individual, thereby overestimating the true cross-tissue correlation. To ensure appropriate estimation of TL association in studies with minimally acceptable sample sizes, future studies should follow TRN guidelines on sample sizes for cross-sectional or longitudinal study designs. Studies with a sample size of 75 are unlikely to be powered to provide reproducible associations with health outcomes and exposures, due to factors like assay precision (Lindrose et al., 2021). However, recent literature shows the sample size needed to detect a TL cross-tissue correlation may be as small as 12, in line with sample sizes of this meta-analysis (Hastings et al., 2021).

Assay methodology was a significant moderator of the cross-tissue TL meta-correlation. The hybridization analyses, e.g. Southern blot, produced the highest correlated TL measurements between tissues, followed by FISH analyses, then qPCR analyses, and finally correlations from Luminex analyses. The qPCR analyses were the most highly represented methodology in this

meta-analysis, followed by hybridization analyses, with only 8 studies using a FISH analysis, and only one study using Luminex analysis. It is a limitation of this work that only one study used Luminex analysis for TL measurement, however this study produced many correlations as the first concerted effort to identify cross-tissue TL correlations for over 25 tissues (Demanelis et al., 2020). These findings on the significant effect of TL measurement methodology are consistent with other meta-analyses which reported a significant influence of the measurement methodology on the overall associations (Wang et al., 2018). Given the significant effect of methodology on the meta-correlation of TL between tissues within an individual, future research should ensure extremely detailed protocols of the utilized TL measurement methodology are included in each publication to increase the reproducibility of findings across measurement methodologies (Lindrose et al., 2021).

The moderating effect of racial and ethnic demographics on the overall model was unable to be tested given the low reporting of these demographics in the studies included in this qualitative review and meta-analysis. While more of the studies included in the meta-analysis (32.73%) reported racial / ethnic demographics compared to studies included only in the qualitative review (11.51%) and reported more racial and ethnic categories – 7 categories in the meta-correlation studies to 5 categories in the review studies – most studies did not report these essential demographics of the study participants. Given evidence of differential associations between TL and health outcomes by race and ethnicity, to provide meaningful results, studies conducting human telomere research must report racial and ethnicity. The sex of participants was not reported for 37.58% of studies in the qualitative review, while only 23.64% of the studies in the meta-correlation did not report this demographic. None of the 220 studies reported inclusion of any non-cissexual or transgender participant. Studies included in the meta-correlation had a

more balanced representation of the neonatal, young adult, adult, and lifespan age groups compared to studies only included the qualitative review. Only one study from 1996 did not report age of participants that was included in this meta-correlation, whereas 8.33% of studies in the review that did not report on this fundamental demographic for telomere research studies. Future studies should include detailed demographics of participants in order to identify the effects of demographic differences on TL associations and increase equitable research findings.

The average score on the TRN quality assessment of methodology reporting criteria was greater (84.97%) in studies in the meta-analysis compared to studies in the review (75.73%). The increased quality of studies included in the meta-analysis is not surprising, however, given that studies have reported cross-tissue correlations as an indicator of study assay validity, meaning that the same studies reporting these correlations would be likely to also report demographics of participants and other methodological features present on the TRN Telomere Length Data Quality Assurance Checklist. Even so, the TRN identified these criteria to ensure reproducibility of telomere research findings across studies, and only half of the studies in the meta-analysis (n = 32; 58.18%) reported 100.00% of the criteria. It is important to note that the TRN Telomere Length Data Quality Assurance Checklist reflects the details presented in a publication and not the quality of the assay itself. Previous work highlights the wide variability in the reporting of assay conditions, as well as actual assay precision, assay replicability within the same assay type, and this meta-analysis displays the need to increase the rigor of methodological reporting in TL research across assay types (Morinha et al., 2020; Lindrose et al., 2021).

For comparability across sample types, multiple studies reported on the correlation of TL within individuals from within or across blood samples (e.g. B cells or dried blood spots) with findings ranging from a perfect negative correlation (-1) to perfect positive correlations (1). The

findings of the sensitivity analyses support a proposed biological model of coupled TL shortening across tissue types of different replicative histories, in which samples of hemopoietic origin are more closely correlated compared to samples of non-hemopoietic origin (Rollings et al., 2019). The overall effect of correlations of TL between non-blood samples and blood samples was of $z = 0.47$, which, given that blood to blood and non-blood to non-blood correlations were greater, investigators can consider these meta-analytic relationships when deciding a proxy tissue where samples peripherally available may not be the best selection for TL measurement. Future research should select samples for TL measurement that are, or reflect, the sample of biological consequence in the study to detect TL findings.

Coupling of TL shortening across tissues is further demonstrated by the differences in overall effect seen in the sensitivity analysis comparing samples peripherally available to samples surgically obtainable. As seen in the blood and non-blood correlation, the overall effect ($z = 0.47$) for correlations of TL between peripherally available and surgically obtainable samples is lower than correlations of samples that are both peripherally available ($z = 0.76$) or both surgically obtainable ($z = 0.65$). Future research should consider these correlations when choosing a surgically obtainable or peripherally available sample for TL measurement and the relationship of this sample source on the biological mechanisms of the outcome of interest, perhaps conducting small pilot studies to identify the potential of a sample source prior to study design.

Peripherally available samples may be appropriate when an exposure or outcome is expected to exert systemic effects. For example, the meta-analysis on the association of TL and depression reported a larger association of TL with depression in samples from peripherally available samples. This is consistent with a biological model of depression where the effects of

the disease go beyond neurological tissue to a systemic impact of depression, which is reflected in evidence of physiologic and immune consequences of depression (Branchi et al., 2021; Lee and Giuliani, 2019). In other words, if the predictor variable in a study affects the biology of peripheral mononuclear blood cells, or the outcome of interest is based in these cells, then selecting this sample for TL measurement is sound. However, in other research queries, such as prostate cancer, selecting a peripheral sample is less advisable as the disease may or may not initiate systemic effects. Therefore, sample source selection for future TL studies requires an advanced comprehension of the outcome of interest on individual samples and may still require further work to understand how the correlations of cross-tissue TLs differ in non-healthy samples.

Despite the significant strengths of this meta-analysis and qualitative review, there are limitations. First, while the focus of this meta-analysis was to evaluate correlations in healthy individuals or presumed healthy at time of death prior to sample collection according to WHO definition, individuals may have been included who had unrecognized illnesses or conditions, such as cytomegalovirus (re)infection (van de Berg et al., 2010). Even within these criteria, because the primary focus of the majority of these studies was not a cross-tissue TL correlation in healthy individuals there may be exposures or confounding factors that influence the correlations that are included in this meta-analysis. Further, there are known effects of degradation of tissues post-mortem on TLs that are not addressed in this work. These results are limited to TL correlations in healthy individuals and how different disease states influence the cross-tissue TL meta-correlation is a critical next step which will likely vary based on the disease or exposure. Despite this limitation, this meta-correlation provides a reference for researchers aiming to understand the pathological effects of a disease on the biological aging of an affected

tissue sampled in concert with a non-affected tissue for TL measurement. Future work should identify how specific exposures, biological processes, and health outcomes alter this cross-tissue TL meta-correlation.

An additional limitation to this work is that a significant number of studies could not be included in the meta-correlation due to a lack of reporting of correlations and inability to receive correlations from 165 corresponding authors for various reasons. To address this limitation, a qualitative review of studies without correlations was conducted to report on the difference of quality of studies with and without correlations included in this work. Studies included in the meta-analysis that did provide correlations, either through the original publication or by request, were of greater quality than studies only included in the review. For example, the age distribution was more equitable in the meta-analysis studies resulting in a cross-tissue TL meta-correlation that is applicable across the lifespan. Moreover, studies which provided correlations also provided a greater number of racial and ethnic groups and possessed higher TRN quality assessment scores than studies in the qualitative review. Thus, there is reason to believe that the meta-correlation derived from these higher quality studies is more reliable than one derived from studies of lower quality. This meta-analysis included correlations from 4,324 unique individuals which greatly increases the sample size of any previously published cross-tissue TL correlation.

Conclusions

In conclusion, these results support, across age groups, biological sex, measurement methodologies, and sample sizes, that TL measured in one tissue correlates with TL measured in other tissues within healthy individuals. Future TL research needs to be intentional about study design and the specific exposure or outcome that is being investigated to ensure the sample selected for TL is biologically relevant. The reader is also directed to supplementary material

that provides correlations by each of the 102 sample types included in this meta-correlation to identify samples that may be utilized when a biologically relevant tissue is unavailable for a given research project. Previous research may not have shown a correlation between TL and an outcome or exposure due to the tissue selected for the TL assay. These findings, or lack thereof, warrant critical appraisal of existing literature and how different source tissues could better inform the role of TL in a disease pathology or effects of exposures on biological aging. Furthermore, these findings identify the need for more rigorous methodologic reporting and reporting of demographics for human research participants in future TL studies. In conclusion, there is a robust correlation of TL across healthy human tissues. Future work should focus on defining how biological processes and pathologies affect the relationship between TL of different tissues within individuals.

Table 1. Characteristics of Studies in the Qualitative Review

Author and Year	Country	Study Participants or Correlation Sample Size	Race & Ethnicity	TL Measurement Methodology	Age Group	Sex of Participants	Quality Rating
Ahmad 2012	**Sweden and Finland**	**58**	**unknown**	**qPCR analysis**	**adult**	**27.59% Female, 72.41% Male**	**100.00%**
Ahmed 2016	United Kingdom	7	unknown	hybridization analysis	lifespan	unknown	50.00%
Aida 2008	**Japan**	**21**	**unknown**	**FISH analysis**	**lifespan**	**42.85% Female, 57.15% Male**	**85.71%**
Aida 2011	Japan	24	unknown	FISH analysis	lifespan	50.00% Female, 50.00% Male	85.71%
Aida 2012	Japan	17	unknown	FISH analysis	lifespan	41.17% Female, 58.83% Male	87.50%

Aida 2015	Japan	68	unknown	FISH analysis	lifespan	4.17% Female, 95.83% Male	75.00%
Akhtar 2013	India	50	unknown	hybridization analysis	lifespan	40.00% Female, 60.00% Male	75.00%
Albuquerque 2017	Portugal	unknown	unknown	FISH analysis	unknown / other	unknown	62.50%
Alder 2018	United States	192	36.45% B, 38.00% W, 25.55% A	FISH analysis	lifespan	53.65% Female, 46.35% Male	87.50%
Allsopp 1995	Canada	61	unknown	hybridization analysis	lifespan	unknown	62.50%
Allsopp 2007	United States	10	unknown	hybridization analysis	neonatal	unknown	62.50%
Alrefaei 2019	Saudi Arabia	48	unknown	qPCR analysis	young adult	100% Female	50.00%
Armengol 2008	Spain	48	unknown	FISH analysis	lifespan	87.50% Female, 12.50% Male	57.14%
Aston 2012	**United States**	**135**	**14.60% B, 77.10% W, 2.10% L, 4.20% A, 2.10% O**	**hybridization analysis**	**lifespan**	**100% Male**	**50.00%**
Aubert 2012	Canada	835	unknown	FISH analysis	lifespan	unknown	75.00%
Azevedo 2013	Portugal	10	unknown	FISH analysis	lifespan	30.00% Female, 70.00% Male	85.71%
Baerlocher 2003	Canada	5	unknown	FISH analysis	lifespan	unknown	85.71%
Baerlocher 2004	Canada	1	unknown	FISH analysis	adult	100% Female	28.57%
Baerlocher 2009	Switzerland	44	unknown	FISH analysis	lifespan	54.55% Female, 45.45% Male	62.50%
Ball 1998	United Kingdom	60	unknown	hybridization analysis	lifespan	unknown	87.50%
Bastos 2020	Brazil	7	unknown	FISH analysis	adult	100% Male	62.50%
Batliwalla 2000	United States	8	unknown	FISH analysis	unknown / other	unknown	57.14%
Beier 2007	Germany	20	unknown	FISH analysis	young adult	unknown	85.71%
Beier 2015	Germany	104	unknown	FISH analysis	lifespan	unknown	50.00%

Benetos 2018	**France**	**259**	**unknown**	**hybridization analysis**	**adult**	**46.88% Female, 53.12% Male**	**62.50%**
Bestilny 2000	Canada	27	unknown	hybridization analysis	unknown / other	unknown	87.50%
Betjes 2011	The Netherlands	144	unknown	FISH analysis	adult	36.11% Female, 63.89% Male	75.00%
Bhattacharya 2019	**India**	**41**	**unknown**	**hybridization analysis**	**neonatal**	**39.02% Female, 60.98% Male**	**50.00%**
Bijnens 2017	Belgium	184	unknown	qPCR analysis	young adult	49.46% Female, 50.54% Male	100.00%
Biron-Shental 2015	Israel	16	unknown	FISH analysis	neonatal	unknown	100.00%
Boeck 2018	Germany	21	unknown	FISH analysis	adult	100% Female	75.00%
Boeck 2018	Germany	15	100% W	FISH analysis	young adult	100% Female	87.50%
Brummendorf 2001	Canada	4	unknown	FISH analysis	young adult	unknown	12.50%
Burns 2000	United States	5	unknown	hybridization analysis	lifespan	unknown	87.50%
Butler 1998	United States	8	100% W	hybridization analysis	lifespan	37.50% Female, 62.50% Male	75.00%
Cafueri 2012	**Italy**	**6 to 20**	**unknown**	**FISH analysis**	**adult**	**unknown**	**100.00%**
Chahine 2019	**Lebanon**	**35**	**100% Lebanese**	**hybridization analysis**	**adult**	**5.71% Female, 94.29% Male**	**100.00%**
Chen 2019	China	20	unknown	FISH analysis	young adult	50.00% Female, 50.00% Male	87.50%
Cohen 2013	**United States**	**111 to 138**	**22.40% B, 71.70% W, 5.90% O**	**qPCR analysis**	**lifespan**	**41.45% Female, 58.55% Male**	**87.50%**
Colonna-Romano 2009	Italy	88	unknown	FISH analysis	lifespan	unknown	71.43%
Costa del Amo 2018	United Kingdom	5	unknown	qPCR analysis	lifespan	unknown	100.00%
Craig 2003	**United Kingdom**	**32 to 88**	**unknown**	**hybridization analysis**	**adult**	**unknown**	**100.00%**
Cross 2009	United Kingdom	23	unknown	FISH analysis	young adult	47.83% Female,	62.50%

						52.17% Male	
Damjanovic 2007	United States	82	unknown	hybridization analysis	adult	73.17% Female, 26.83% Male	87.50%
Damle 2004	United States	15	unknown	FISH analysis	lifespan	unknown	100.00%
Daniali 2013	**United States**	**87**	**40.20% B, 59.80% W, 37.90% L, 21.80% NL**	**hybridization analysis**	**lifespan**	**70.11% Female, 29.89% Male**	**62.50%**
Darmishonnejad 2019	Iran	10	unknown	qPCR analysis	young adult	100% Male	100.00%
Davy 2009	United States	36	unknown	hybridization analysis	neonatal	unknown	87.50%
de Baca 2017	**United States**	**129**	**5.00% B, 78.00% W, 7.00% L, 9.00% A, 1.00% AIAN**	**qPCR analysis**	**young adult**	**100% Female**	**100.00%**
de Beer 2015	Switzerland	5	unknown	FISH analysis	adult	40.00% Female, 60.00% Male	75.00%
De Pauw 2002	The Netherlands	10	unknown	hybridization analysis	young adult	unknown	62.50%
De Vusser 2015	Belgium	30	unknown	qPCR analysis	young adult	43.60% Female, 56.40% Male	100.00%
Demanelis 2020	**United States**	**18 to 381**	**12.70% B, 84.90% W, 2.40% O**	**Luminex analysis**	**lifespan**	**67.00% Male, 33.00% Female**	**100.00%**
Denham 2017	**Australia**	**79**	**100% W**	**qPCR analysis**	**young adult**	**27.50% Female, 72.50% Male**	**100.00%**
Desgarnier 2016	Canada	52	unknown	qPCR analysis	lifespan	unknown	100.00%
Dlouha 2014	**Czech Republic**	**7 to 12**	**unknown**	**qPCR analysis**	**lifespan**	**47.62% Female, 52.38% Male**	**100.00%**
Domogala 2017	United Kingdom	6	unknown	qPCR analysis	unknown / other	unknown	75.00%

Effros 1996	**United States**	**5**	**unknown**	**hybridization analysis**	**unknown / other**	**unknown**	**87.50%**
Engelhardt 2000	United States	207	unknown	hybridization analysis	lifespan	unknown	62.50%
Etzel 2020	**United States**	**58**	**46.55% B, 50% W, 3.45% L**	**qPCR analysis**	**young adult**	**100% Female**	**100.00%**
Everaerts 2018	Belgium	13	unknown	qPCR analysis	lifespan	23.08% Female, 76.92% Male	100.00%
Fali 2019	**France**	**27 to 29**	**100% W**	**qPCR analysis**	**unknown / other**	**unknown**	**100.00%**
Ferlin 2013	**Italy**	**81**	**unknown**	**qPCR analysis**	**young adult**	**100% Male**	**87.50%**
Fernandez-Rozadilla 2018	United Kingdom	109	unknown	qPCR analysis	lifespan	51.89% Female, 48.11% Male	87.50%
Ferraris 2005	Italy	13	unknown	hybridization analysis	lifespan	100% Female	25.00%
Fessler 2016	Austria	50	unknown	qPCR analysis	young adult	66.00% Female, 34.00% Male	75.00%
Finnicum 2017	The Netherlands	1892	unknown	qPCR analysis	lifespan	68.97% Female, 31.03% Male	87.50%
Fritsch 2005	United States	5	unknown	FISH analysis	adult	60.00% Female, 40.00% Male	75.00%
Fukunaga 2007	Japan	15	unknown	FISH analysis	lifespan	unknown	87.50%
Gadalla 2018	United States	197	75.60% W, 24.40% O	FISH analysis	young adult	47.21% Female, 52.79% Male	71.43%
Garcia-Martin 2017	United Kingdom	42	92.85% W, 7.15% O	hybridization analysis	young adult	100% Female	100.00%
Gardner 2007	United States	12	unknown	hybridization analysis	lifespan	unknown	100.00%
Garrett-Bakelman 2019	United States	1	100% W	FISH analysis	adult	100% Male	37.50%
Gemetzi 2012	Greece	15	unknown	qPCR analysis	lifespan	81.82% Female, 18.18% Male	100.00%
Goldman 2018	**United States**	**34**	**unknown**	**qPCR analysis**	**unknown / other**	**100% Female**	**100.00%**
Grady 2016	Germany	15	100% W	FISH analysis	young adult	100% Female	87.50%
Hagman 2020	Denmark	140	unknown	qPCR analysis	lifespan	100% Male	100.00%

Halaschek-Wiener 2008	Canada	181	unknown	FISH analysis	adult	67.16% Female, 32.84% Male	87.50%
Harley 2011	United States	114	unknown	FISH analysis	adult	28.07% Female, 71.93% Male	50.00%
Hastie 1990	United Kingdom	1	unknown	hybridization analysis	neonatal	unknown	37.50%
Hearps 2012	Australia	unknown	unknown	FISH analysis	lifespan	100% Male	100.00%
Herndier-Brandstetter 2008	Austria	9	unknown	FISH analysis	lifespan	unknown	71.43%
Hiam 2020	**Australia**	**93**	**unknown**	**qPCR analysis**	**unknown / other**	**100% Male**	**100.00%**
Hills 2009	Canada	1210	unknown	FISH analysis	neonatal	unknown	87.50%
Hills 2009	**Canada**	**2 to 5**	**unknown**	**hybridization analysis**	**neonatal**	**unknown**	**100.00%**
Hiraishi 2018	Japan	128	unknown	FISH analysis	lifespan	49.22% Female, 50.78% Male	87.50%
Hoare 2010	United Kingdom	41	unknown	FISH analysis	young adult	48.78% Female, 51.22% Male	71.43%
Hoffmann 2009	United Kingdom	129	unknown	FISH analysis	young adult	47.29% Female, 52.71% Male	100.00%
Huang 2016	The Netherlands	44	unknown	FISH analysis	adult	54.55% Female, 45.45% Male	42.86%
Hug 2003	Germany	20	unknown	hybridization analysis	adult	unknown	62.50%
Imam 2012	Canada	39	63.63% W, 22.73% A, 13.64% AIAN	qPCR analysis	young adult	43.18% Female, 56.82% Male	100.00%
Ishikawa 2012	Japan	65	unknown	hybridization analysis	lifespan	43.08% Female, 56.92% Male	87.50%
Jebaraj 2019	Germany	20	unknown	qPCR analysis	unknown / other	unknown	75.00%
Jones 2014	United Kingdom	7	unknown	qPCR analysis	neonatal	unknown	100.00%
Jones 2019	**United States**	**67**	**56.70% B, 29.80% W,**	**qPCR analysis**	**neonatal**	**43.30% Female, 56.70% Male**	**100.00%**

			13.50% O				
Jorgensen 2013	Denmark	12	unknown	FISH analysis	lifespan	100% Male	75.00%
Kapoor 2009	United States	unknown	unknown	FISH analysis	unknown / other	unknown	85.71%
Karabatsiakis 2014	Germany	50	unknown	FISH analysis	adult	100% Female	62.50%
Karlsson 2008	Sweden	19	unknown	qPCR analysis	lifespan	unknown	100.00%
Kawano 2011	United States	19	unknown	qPCR analysis	lifespan	unknown	87.50%
Kimura 2010	**United States**	**8**	**12.50% B, 75.00% WL, 25.00% NL, 12.50% O**	**hybridization analysis**	**neonatal**	**25.00% Female, 75.00% Male**	**42.86%**
Kolyada 2016	**Ukraine**	**34**	**unknown**	**qPCR analysis**	**adult**	**unknown**	**37.50%**
Kurosaka 2006	Japan	17	unknown	FISH analysis	lifespan	94.12% Female, 5.88% Male	42.86%
Lakota 2019	United States	192	unknown	FISH analysis	unknown / other	unknown	42.86%
Lakowa 2015	Germany	97	100% W	qPCR analysis	lifespan	52.58% Female, 47.42% Male	100.00%
Laurentino 2020	**Germany**	**179**	**unknown**	**qPCR analysis**	**lifespan**	**100% Male**	**87.50%**
Laye 2012	Denmark	8	unknown	qPCR analysis	young adult	12.50% Female, 87.50% Male	87.50%
Lee 2020	South Korea	21	unknown	FISH analysis	unknown / other	57.14% Female, 42.86% Male	100.00%
Lin 2010	**United States**	**33 to 60**	**4.60% B, 83.10% W, 1.50% L, 98.50% NL, 10.80% A, 10.80% AIAN, 10.80% NHOPI**	**qPCR analysis**	**adult**	**47.20% Female, 52.80% Male**	**100.00%**
Lin 2015	**United States**	**143**	**unknown**	**hybridization analysis**	**lifespan**	**100% Female**	**100.00%**
Lin 2016	**United States**	**39**	**3.30% B, 77.60%**	**qPCR analysis**	**adult**	**100% Female**	**100.00%**

			W, 7.70% L, 92.30% NL, 10.80% A, 10.80% AIAN, 10.80% NHOPI				
Liu 2019	**Japan**	**55**	**unknown**	**hybridization analysis**	**adult**	**32.73% Female, 67.27% Male**	**50.00%**
Mackey 2016	Denmark	29	unknown	hybridization analysis	young adult	100% Male	75.00%
Mahady 2020	**United States**	**23**	**unknown**	**qPCR analysis**	**adult**	**47.83% Female, 52.17% Male**	**100.00%**
Maini 1999	United Kingdom	8	unknown	hybridization analysis	young adult	unknown	37.50%
Mamdani 2015	United States	10	unknown	qPCR analysis	young adult	30.00% Female, 70.00% Male	62.50%
Mariani 2003	**Italy**	**47**	**unknown**	**FISH analysis**	**unknown / other**	**unknown**	**57.14%**
Martens 2000	Germany	1	unknown	FISH analysis	young adult	100% Female	71.43%
Martens 2002	**Germany**	**3**	**unknown**	**FISH analysis**	**young adult**	**unknown**	**25.00%**
Martens 2016	**Belgium**	**743**	**unknown**	**qPCR analysis**	**young adult**	**100% Female**	**100.00%**
Martens 2017	Belgium	641	88.50% W, 11.50% O	qPCR analysis	neonatal	49.61% Female, 50.39% Male	87.50%
Martens 2020	Belgium	1258	unknown	qPCR analysis	neonatal	48.65% Female, 51.35% Male	100.00%
Mason 2018	United States	162	12.40% B, 61.50% W, 12.40% L, 8.70% API, 5.00% O	qPCR analysis	young adult	100% Female	100.00%
Mathioudakis 2000	United States	17	unknown	FISH analysis	young adult	70.59% Female, 29.41% Male	42.86%

Mathur 2013	United States	32	unknown	FISH analysis	adult	40.63% Female, 59.37% Male	25.00%
Matsuda 2015	Japan	183	unknown	FISH analysis	unknown / other	unknown	42.86%
Meijers 2012	The Netherlands	120	unknown	FISH analysis	lifespan	56.67% Female, 43.33% Male	42.86%
Meijers 2015	The Netherlands	22	unknown	FISH analysis	adult	unknown	25.00%
Melk 2000	Canada	24	unknown	hybridization analysis	lifespan	45.83% Female, 54.17% Male	50.00%
Menon 2012	United States	18	50.00% B, 50.00% W	qPCR analysis	neonatal	unknown	100.00%
Najarro 2015	United States	22	unknown	qPCR analysis	adult	36.36% Female, 63.64% Male	37.50%
Nakamura 2002	**Japan**	**21 to 41**	**unknown**	**hybridization analysis**	**lifespan**	**44.44% Female, 55.56% Male**	**50.00%**
Nakamura 2007	**Japan**	**72**	**unknown**	**hybridization analysis**	**lifespan**	**64.52% Female, 35.48% Male**	**100.00%**
Nemtsova 2020	Ukraine	22	unknown	qPCR analysis	adult	60.26% Female, 39.74% Male	50.00%
Neuber 2003	Germany	22	unknown	FISH analysis	young adult	56.95% Female, 43.05% Male	85.71%
Ngom 2011	Gambia	60	unknown	qPCR analysis	young adult	100% Male	87.50%
Novakovic 2016	Australia	7	unknown	qPCR analysis	neonatal	71.43% Female, 28.57% Male	87.50%
O'Callaghan 2011	Australia	91	unknown	qPCR analysis	lifespan	52.75% Female, 47.25% Male	100.00%
Okuda 2000	United States	51	27.45% B, 72.55% W	hybridization analysis	lifespan	54.90% Female, 45.10% Male	50.00%

Okuda 2002	**United States**	**23 to 158**	**51.20% B, 28.60% W, 17.30% L, 2.90% O**	**hybridization analysis**	**neonatal**	**49.40% Female, 50.60% Male**	**100.00%**
Olivieri 2013	Italy	40	unknown	qPCR analysis	adult	55.00% Female, 45.00% Male	87.50%
Ong 2018	Singapore	3	unknown	FISH analysis	lifespan	unknown	71.43%
Ouyang 2007	Canada	7	unknown	FISH analysis	lifespan	28.57% Female, 71.43% Male	75.00%
Palmer 1997	United States	6	unknown	hybridization analysis	young adult	unknown	25.00%
Panczyszyn 2020	Poland	11	unknown	qPCR analysis	young adult	100% Female	75.00%
Panossian 2003	United States	15	unknown	hybridization analysis	adult	26.67% Female, 73.33% Male	87.50%
Pavanello 2020	Italy	34	unknown	qPCR analysis	lifespan	22.86% Female, 77.14% Male	100.00%
Pavlaki 2012	Greece	68	unknown	qPCR analysis	lifespan	86.76% Female, 13.24% Male	87.50%
Prather 2015	United States	87	62.80% W, 37.20% O	qPCR analysis	young adult	82.76% Female, 17.24% Male	87.50%
Qian 2016	China	4	unknown	qPCR analysis	adult	52.00% Female, 48.00% Male	87.50%
Rej 2020	**United States**	**15 to 19**	**unknown**	**qPCR analysis**	**young adult**	**unknown**	**37.50%**
Renault 2002	France and Sweden	10	unknown	hybridization analysis	lifespan	100% Male	87.50%
Richardson 2000	United States	35	unknown	hybridization analysis	unknown / other	unknown	50.00%
Riddell 2015	United Kingdom and Singapore	11	unknown	FISH analysis	lifespan	unknown	85.71%
Rigolin 2004	Italy	55	unknown	FISH analysis	adult	43.64% Female, 56.36% Male	71.43%

Risques 2008	United States	14	unknown	qPCR analysis	young adult	61.02% Female, 38.98% Male	100.00%
Robertson 2000	United Kingdom	70	unknown	hybridization analysis	lifespan	unknown	87.50%
Robertson 2001	United Kingdom	36	unknown	hybridization analysis	lifespan	unknown	87.50%
Rufer 1998	Canada	10	unknown	FISH analysis	unknown / other	unknown	87.50%
Rufer 1999	Canada	472	unknown	FISH analysis	lifespan	unknown	85.71%
Rufer 2001	Switzerland	4	unknown	FISH analysis	unknown / other	unknown	71.43%
Sakoff 2002	**Australia**	**14 to 17**	**unknown**	**hybridization analysis**	**lifespan**	**36.84% Female, 63.16% Male**	**87.50%**
Sampson 2006	United Kingdom	28	100% W	FISH analysis	adult	100% Male	85.71%
Sanderson 2017	United Kingdom	22	unknown	qPCR analysis	lifespan	40.91% Female, 59.09% Male	87.50%
Sarzotti-Kelsoe 2011	United States	1	unknown	FISH analysis	young adult	unknown	25.00%
Shao 2007	United States	65	3.07% B, 87.69% W, 7.69% L, 1.55% A	FISH analysis	adult	40.00% Female, 60.00% Male	85.71%
Shlush 2011	Israel	15	unknown	FISH analysis	young adult	100% Male	100.00%
Sillanpaa 2017	**Finland**	**20**	**unknown**	**qPCR analysis**	**adult**	**100% Female**	**75.00%**
Silva 2016	Brazil	15	unknown	FISH analysis	adult	unknown	37.50%
Simpson 2010	United States	9	unknown	qPCR analysis	young adult	100% Male	50.00%
Son 2000	United States	121	unknown	FISH analysis	lifespan	unknown	85.71%
Son 2003	United States	53	unknown	FISH analysis	lifespan	unknown	100.00%
Sondergaard 2002	Denmark	8	unknown	hybridization analysis	young adult	12.50% Female, 87.50% Male	87.50%
Spyridopoulos 2008	**Germany**	**67**	**unknown**	**FISH analysis**	**young adult**	**50.00% Female, 50.00% Male**	**100.00%**
Spyridopoulos 2009	Germany	17	unknown	FISH analysis	lifespan	100% Male	100.00%
Srinivasa 2014	United States	91	unknown	qPCR analysis	unknown / other	unknown	100.00%

Stout 2017	**United States**	**24**	**62.50% W, 25.00% L, 8.30% A, 4.20% O**	**qPCR analysis**	**young adult**	**50.00% Female, 50.00% Male**	**100.00%**
Szebeni 2014	United States	14	unknown	qPCR analysis	adult	7.14% Female, 92.86% Male	50.00%
Tahamtan 2019	**Iran**	**38**	**unknown**	**qPCR analysis**	**adult**	**100% Male**	**62.50%**
Takahashi 2004	Japan	7	100% A	hybridization analysis	young adult	62.00% Female, 38.00% Male	87.50%
Takubo 2002	**Japan**	**137 to 168**	**unknown**	**hybridization analysis**	**lifespan**	**50.00% Female, 50.00% Male**	**87.50%**
Takubo 2010	Japan	33	unknown	FISH analysis	lifespan	42.42% Female, 57.58% Male	71.43%
Tchakmakjian 2004	**United States**	**108**	**47.90% B, 52.10% W**	**hybridization analysis**	**lifespan**	**100% Male**	**100.00%**
Tefferi 2015	United States	16	unknown	qPCR analysis	lifespan	unknown	75.00%
Terasaki 2002	Japan	35	unknown	hybridization analysis	lifespan	unknown	87.50%
Thibeault 2006	United States	12	unknown	qPCR analysis	adult	25.00% Female, 75.00% Male	87.50%
Thomas 2008	Australia	56	unknown	qPCR analysis	lifespan	53.57% Female, 46.43% Male	87.50%
Tomita 2018	**Japan**	**25**	**unknown**	**FISH analysis**	**unknown / other**	**36.00% Female, 64.00% Male**	**62.50%**
Tucker 2000	United Kingdom	8	unknown	hybridization analysis	adult	unknown	87.50%
Tucker 2004	United Kingdom	11	unknown	hybridization analysis	lifespan	45.45% Female, 54.55% Male	87.50%
Vahter 2020	**Argentina**	**98**	**unknown**	**qPCR analysis**	**neonatal**	**100% Female**	**100.00%**
Valentijn 2015	**United Kingdom**	**53**	**unknown**	**qPCR analysis**	**young adult**	**100% Female**	**100.00%**

Valls-Bautista 2015	**Spain**	**40**	**unknown**	**hybridization analysis**	**adult**	**37.50% Female, 62.50% Male**	**100.00%**
van de Berg 2010	**The Netherlands**	**10**	**unknown**	**FISH analysis**	**adult**	**unknown**	**100.00%**
van Mierlo 2017	The Netherlands	11	unknown	qPCR analysis	adult	54.55% Female, 45.45% Male	87.50%
Van Ziffle 2003	**Canada**	**2 to 6**	**unknown**	**FISH analysis**	**lifespan**	**42.86% Female, 57.14% Male**	**100.00%**
Vazirpanah 2017	The Netherlands	11	unknown	qPCR analysis	young adult	9.09% Female, 90.91% Male	50.00%
Vecoli 2017	**Italy**	**112**	**unknown**	**qPCR analysis**	**young adult**	**100% Male**	**62.50%**
Verma 2012	United Kingdom	73	unknown	FISH analysis	lifespan	47.95% Female, 52.05% Male	100.00%
Verma 2017	Germany	6	unknown	qPCR analysis	young adult	unknown	75.00%
Walters 2014	United States	12	38.09% B, 28.57% W, 33.34% O	hybridization analysis	young adult	100% Male	87.50%
Wang 2018	China	25	unknown	FISH analysis	lifespan	50.00% Female, 50.00% Male	85.71%
Weng 1995	United States	30	unknown	hybridization analysis	lifespan	unknown	75.00%
Werner 2015	Germany	66	unknown	FISH analysis	lifespan	unknown	85.71%
Werner 2019	Germany	35	unknown	FISH analysis	adult	65.71% Female, 34.29% Male	50.00%
Widmann 2007	Germany	40	100% W	FISH analysis	adult	unknown	87.50%
Wilson 2008	**United Kingdom**	**32**	**unknown**	**qPCR analysis**	**adult**	**41.67% Female, 58.33% Male**	**75.00%**
Wolf 2006	Austria	17	unknown	FISH analysis	adult	unknown	87.50%
Wolthers 1996	The Netherlands	9	unknown	hybridization analysis	unknown / other	100% Male	62.50%
Wong 2011	The Netherlands	49	unknown	qPCR analysis	lifespan	40.82% Female,	87.50%

						59.18% Male	
Wu 2000	Denmark	30	unknown	hybridization analysis	lifespan	50.00% Female, 50.00% Male	37.50%
Wu 2001	Denmark	11	unknown	hybridization analysis	lifespan	45.45% Female, 54.55% Male	50.00%
Wu 2003	United States	7	unknown	hybridization analysis	young adult	unknown	75.00%
Wu 2007	**Taiwan**	**26**	**100% A**	**hybridization analysis**	**lifespan**	**unknown**	**100.00%**
Yamada 1995	Japan	5	unknown	hybridization analysis	lifespan	unknown	75.00%
Youngren 1998	**United States**	**2 to 10**	**unknown**	**hybridization analysis**	**neonatal**	**unknown**	**100.00%**
Zanet 2013	**Canada**	**12**	**unknown**	**qPCR analysis**	**young adult**	**unknown**	**100.00%**
Ziegler 2017	Germany	311	unknown	FISH analysis	lifespan	17.40% Female, 82.60% Male	87.50%
Zinkova 2017	Czech Republic	36	unknown	qPCR analysis	young adult	52.78% Female, 47.22% Male	62.50%
Zole 2019	Latvia	53	unknown	hybridization analysis	lifespan	71.70% Female, 28.30% Male	75.00%

Table 1 Legend: Bolded studies are included in the meta-analysis. Study participants applies for studies in the qualitative review only, while correlation sample size applies to studies included in the meta-analysis. B = Black, W = White, WL = White Latino, L = Latino, NL = Non-Latino, A = Asian, API = Asian or Pacific Islander, AIAN = American Indian / Alaska Native, NAOPI = Native Hawaiian / Other Pacific Islander, O = Other; unknown / other = not reported or not within the other age group categories, neonatal = under 1, young adult = range or mean of age does not include 50, adult = range or mean of age includes 50, lifespan = range of ages is greater than 30 years.

Table 2. Characteristic Differences Between Studies in the Qualitative Review versus Studies in the Meta-Analysis

Characteristic	Studies in the Qualitative Review Only	Studies Included in the Meta-Analysis
Number of Studies Included	165	55
Country of Study	32 countries	20 countries

Study Sample Size	98.18% reported; 80.00% small, 13.94% medium, 4.24% large	100.00% reported; 67.27% small, 27.27% medium, 5.46% large
Race & Ethnicity	11.51% reported; 5 categories reported	32.73% reported; 7 categories reported
TL Measurement Methodology	45.45% FISH analyses, 24.85% hybridization analyses, 29.70% qPCR analyses	14.55% FISH analyses, 36.36% hybridization analyses, 47.27% qPCR analyses, 1.82% Luminex analyses
Age Group	91.52% reported; 6.06% neonatal, 24.24% young adult, 18.79% adult, 42.42% lifespan	98.18% reported; 12.73% neonatal, 21.82% young adult, 27.27% adult, 27.27% lifespan
Sex of Participants	62.42% reported	76.36% reported
Quality Rating	75.73% average, range 12.50% to 100.00%	84.97% average, range 25.00% to 100.00%

Table 3. Subgroup Analysis of Correlations by Sample Size Grouping

Subgroup Analysis	estimate	p-value	95% LB	95% UB
Large n = 17	0.45	<0.01	0.30	0.60
Medium n = 146	0.49	<0.01	0.31	0.67
Small n = 253	0.75	<0.01	0.59	0.91

Table 4. Subgroup Analysis of Correlations by TL Measurement Methodology

Subgroup Analysis	estimate	p-value	95% LB	95% UB
FISH analyses n = 20	0.77	<0.01	0.39	1.14
hybridization analyses n = 113	0.86	<0.01	0.63	1.10
PCR analyses n = 98	0.51	<0.01	0.37	0.65
Luminex analyses n = 186	0.25	<0.01	0.23	0.27

Table 5. Sensitivity Analysis of Correlations by Blood versus Non-Blood Sample Sources

Sensitivity Analyses	estimate	p-value	95% LB	95% UB
Non-blood and Non-blood n = 274	0.74	<0.01	0.51	0.98
Non-blood and Blood	0.47	<0.01	0.31	0.63

n = 67				
Blood and Blood n = 76	0.93	<0.01	0.74	1.12

Table 6. Sensitivity Analysis of Correlations by Peripherally Available versus Surgically Obtainable Sample Sources

Sensitivity Analyses	estimate	p-value	95% LB	95% UB
Surgically Obtainable and Surgically Obtainable n = 214	0.65	<0.01	0.41	0.88
Peripherally Available and Surgically Obtainable n = 100	0.47	<0.01	0.26	0.68
Peripherally Available and Peripherally Available n = 103	0.76	<0.01	0.59	0.93

References:

Assink, M., Wibbelink, C.J.M., 2016. Fitting three-level meta-analytic models in R: A step-by-step tutorial. TQMP 12, 154–174. https://doi.org/10.20982/tqmp.12.3.p154

Blackburn, E.H., Gall, J.G., 1978. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. Journal of Molecular Biology 120, 33–53. https://doi.org/10.1016/0022-2836(78)90294-2

Branchi, I., Poggini, S., Capuron, L., Benedetti, F., Poletti, S., Tamouza, R., Drexhage, H.A., Penninx, B.W.J.H., Pariante, C.M., 2021. Brain-immune crosstalk in the treatment of major depressive disorder. European Neuropsychopharmacology 45, 89–107. https://doi.org/10.1016/j.euroneuro.2020.11.016

De Meyer, T., Nawrot, T., Bekaert, S., De Buyzere, M.L., Rietzschel, E.R., Andrés, V., 2018. Telomere Length as Cardiovascular Aging Biomarker: JACC Review Topic of the Week. J Am Coll Cardiol 72, 805–813. https://doi.org/10.1016/j.jacc.2018.06.014

Demanelis, K., Jasmine, F., Chen, L.S., Chernoff, M., Tong, L., Delgado, D., Zhang, C., Shinkle, J., Sabarinathan, M., Lin, H., Ramirez, E., Oliva, M., Kim-Hellmuth, S., Stranger, B.E., Lai, T.-P., Aviv, A., Ardlie, K.G., Aguet, F., Ahsan, H., GTEx Consortium, Doherty, J.A., Kibriya, M.G., Pierce, B.L., Aguet, F., Anand, S., Ardlie, K.G., Gabriel, S., Getz, G.A., Graubert, A., Hadley, K., Handsaker, R.E., Huang, K.H., Kashin, S., Li, Xiao, MacArthur, D.G., Meier, S.R., Nedzel, J.L., Nguyen, D.T., Segrè, A.V., Todres, E., Balliu, B., Barbeira, A.N., Battle, A., Bonazzola, R., Brown, A., Brown, C.D., Castel, S.E., Conrad, D.F., Cotter, D.J., Cox, N., Das, S., de Goede, O.M., Dermitzakis, E.T., Einson, J., Engelhardt, B.E., Eskin, E., Eulalio, T.Y., Ferraro, N.M., Flynn, E.D., Fresard, L., Gamazon, E.R., Garrido-Martín, D., Gay, N.R., Gloudemans, M.J., Guigó, R., Hame, A.R., He, Y., Hoffman, P.J., Hormozdiari, F., Hou, L., Im, H.K., Jo, B., Kasela, S., Kellis, M., Kim-Hellmuth, S., Kwong, A., Lappalainen, T., Li, Xin,

Liang, Y., Mangul, S., Mohammadi, P., Montgomery, S.B., Muñoz-Aguirre, M., Nachun, D.C., Nobel, A.B., Oliva, M., Park, YoSon, Park, Yongjin, Parsana, P., Rao, A.S., Reverter, F., Rouhana, J.M., Sabatti, C., Saha, A., Stephens, M., Stranger, B.E., Strober, B.J., Teran, N.A., Viñuela, A., Wang, G., Wen, X., Wright, F., Wucher, V., Zou, Y., Ferreira, P.G., Li, G., Melé, M., Yeger-Lotem, E., Barcus, M.E., Bradbury, D., Krubit, T., McLean, J.A., Qi, L., Robinson, K., Roche, N.V., Smith, A.M., Sobin, L., Tabor, D.E., Undale, A., Bridge, J., Brigham, L.E., Foster, B.A., Gillard, B.M., Hasz, R., Hunter, M., Johns, C., Johnson, M., Karasik, E., Kopen, G., Leinweber, W.F., McDonald, A., Moser, M.T., Myer, K., Ramsey, K.D., Roe, B., Shad, S., Thomas, J.A., Walters, G., Washington, M., Wheeler, J., Jewell, S.D., Rohrer, D.C., Valley, D.R., Davis, D.A., Mash, D.C., Branton, P.A., Barker, L.K., Gardiner, H.M., Mosavel, M., Siminoff, L.A., Flicek, P., Haeussler, M., Juettemann, T., Kent, W.J., Lee, C.M., Powell, C.C., Rosenbloom, K.R., Ruffier, M., Sheppard, D., Taylor, K., Trevanion, S.J., Zerbino, D.R., Abell, N.S., Akey, J., Chen, L., Demanelis, K., Doherty, J.A., Feinberg, A.P., Hansen, K.D., Hickey, P.F., Jasmine, F., Jiang, L., Kaul, R., Kibriya, M.G., Li, J.B., Li, Q., Lin, S., Linder, S.E., Pierce, B.L., Rizzardi, L.F., Skol, A.D., Smith, K.S., Snyder, M., Stamatoyannopoulos, J., Tang, H., Wang, M., Carithers, L.J., Guan, P., Koester, S.E., Little, A.R., Moore, H.M., Nierras, C.R., Rao, A.K., Vaught, J.B., Volpi, S., 2020. Determinants of telomere length across human tissues. Science 369, eaaz6876. https://doi.org/10.1126/science.aaz6876

Dorado-Correa, A.M., Zollinger, S.A., Heidinger, B., Brumm, H., 2018. Timing matters: traffic noise accelerates telomere loss rate differently across developmental stages. Frontiers in Zoology 15, 29. https://doi.org/10.1186/s12983-018-0275-8

Garcia-Rizo, C., Fernandez-Egea, E., Miller, B.J., Oliveira, C., Justicia, A., Griffith, J.K., Heaphy, C.M., Bernardo, M., Kirkpatrick, B., 2013. Abnormal glucose tolerance, white blood cell count, and telomere length in newly diagnosed, antidepressant-naïve patients with depression. Brain, Behavior, and Immunity 28, 49–53. https://doi.org/10.1016/j.bbi.2012.11.009

Gardner, M., Bann, D., Wiley, L., Cooper, R., Hardy, R., Nitsch, D., Martin-Ruiz, C., Shiels, P., Sayer, A.A., Barbieri, M., Bekaert, S., Bischoff, C., Brooks-Wilson, A., Chen, W., Cooper, C., Christensen, K., De Meyer, T., Deary, I., Der, G., Roux, A.D., Fitzpatrick, A., Hajat, A., Halaschek-Wiener, J., Harris, S., Hunt, S.C., Jagger, C., Jeon, H.-S., Kaplan, R., Kimura, M., Lansdorp, P., Li, C., Maeda, T., Mangino, M., Nawrot, T.S., Nilsson, P., Nordfjall, K., Paolisso, G., Ren, F., Riabowol, K., Robertson, T., Roos, G., Staessen, J.A., Spector, T., Tang, N., Unryn, B., van der Harst, P., Woo, J., Xing, C., Yadegarfar, M.E., Park, J.Y., Young, N., Kuh, D., von Zglinicki, T., Ben-Shlomo, Y., 2014. Gender and telomere length: Systematic review and meta-analysis. Experimental Gerontology 51, 15–27. https://doi.org/10.1016/j.exger.2013.12.004

Haussmann, M.F., Salomons, H.M., Verhulst, S., 2011. Telomere measurement tools: Telometric produces biased estimates of telomere length. Heredity 107, 371–371. https://doi.org/10.1038/hdy.2011.10

Haycock, P.C., Heydon, E.E., Kaptoge, S., Butterworth, A.S., Thompson, A., Willeit, P., 2014. Leucocyte telomere length and risk of cardiovascular disease: systematic review and meta-analysis. BMJ 349, g4227. https://doi.org/10.1136/bmj.g4227

Horn, T., Robertson, B.C., Gemmell, N.J., 2010. The use of telomere length in ecology and evolutionary biology. Heredity 105, 497–506. https://doi.org/10.1038/hdy.2010.113

Hu, R., Hua, X., Jiang, Q., 2019. Associations of telomere length in risk and recurrence of prostate cancer: A meta-analysis. Andrologia 51. https://doi.org/10.1111/and.13304

Khosravaniardakani, S., Bokov, D.O., Mahmudiono, T., Hashemi, S.S., Nikrad, N., Rabieemotmaen, S., Abbasalizad-Farhangi, M., 2022. Obesity Accelerates Leukocyte Telomere

Length Shortening in Apparently Healthy Adults: A Meta-Analysis. Front. Nutr. 9, 812846. https://doi.org/10.3389/fnut.2022.812846
Lee, C.-H., Giuliani, F., 2019. The Role of Inflammation in Depression and Fatigue. Frontiers in Immunology 10.
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
Lindrose, A.R., McLester-Davis, L.W.Y., Tristano, R.I., Kataria, L., Gadalla, S.M., Eisenberg, D.T.A., Verhulst, S., Drury, S., 2021. Method comparison studies of telomere length measurement using qPCR approaches: A critical appraisal of the literature. PLOS ONE 16, e0245582. https://doi.org/10.1371/journal.pone.0245582
Morinha, F., Magalhães, P., Blanco, G., 2020. Standard guidelines for the publication of telomere qPCR results in evolutionary ecology. Mol Ecol Resour 20. https://doi.org/10.1111/1755-0998.13152
Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L., Wu, J.R., 1988. A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. Proc. Natl. Acad. Sci. U.S.A. 85, 6622–6626. https://doi.org/10.1073/pnas.85.18.6622
Muller, H.J., 1938. The remaking of chromosomes. The Collecting Net; Woods Hole 13, 181–198.
Mundstock, E., Sarria, E.E., Zatti, H., Mattos Louzada, F., Kich Grun, L., Herbert Jones, M., Guma, F.T.C.R., Mazzola in Memoriam, J., Epifanio, M., Stein, R.T., Barbé-Tuana, F.M., Mattiello, R., 2015. Effect of obesity on telomere length: Systematic review and meta-analysis: Effect of Obesity on Telomeres. Obesity 23, 2165–2174. https://doi.org/10.1002/oby.21183
Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ n71. https://doi.org/10.1136/bmj.n71
Ridout, K.K., Ridout, S.J., Price, L.H., Sen, S., Tyrka, A.R., 2016. Depression and telomere length: A meta-analysis. Journal of Affective Disorders 191, 237–247. https://doi.org/10.1016/j.jad.2015.11.052
Rollings, N., Friesen, C.R., Whittington, C.M., Johansson, R., Shine, R., Olsson, M., 2019. Sex-And tissue-specific differences in telomere length in a reptile. Ecol Evol 9, 6211–6219. https://doi.org/10.1002/ece3.5164
Röver, C., Knapp, G., Friede, T., 2015. Hartung-Knapp-Sidik-Jonkman approach and its modification for random-effects meta-analysis with few studies. BMC Medical Research Methodology 15, 99. https://doi.org/10.1186/s12874-015-0091-1
Schneider, C.V., Schneider, K.M., Teumer, A., Rudolph, K.L., Hartmann, D., Rader, D.J., Strnad, P., 2022. Association of Telomere Length With Risk of Disease and Mortality. JAMA Intern Med 182, 291–300. https://doi.org/10.1001/jamainternmed.2021.7804
Steenstrup, T., Kark, J.D., Verhulst, S., Thinggaard, M., Hjelmborg, J.V.B., Dalgård, C., Kyvik, K.O., Christiansen, L., Mangino, M., Spector, T.D., Petersen, I., Kimura, M., Benetos, A., Labat, C., Sinnreich, R., Hwang, S.-J., Levy, D., Hunt, S.C., Fitzpatrick, A.L., Chen, W., Berenson, G.S., Barbieri, M., Paolisso, G., Gadalla, S.M., Savage, S.A., Christensen, K., Yashin, A.I.,

Arbeev, K.G., Aviv, A., 2017. Telomeres and the natural lifespan limit in humans. Aging (Albany NY) 9, 1130–1142. https://doi.org/10.18632/aging.101216
Teyssier, J.-R., Ragot, S., Chauvet-Gélinier, J.-C., Trojak, B., Bonin, B., 2011. Expression of oxidative stress-response genes is not activated in the prefrontal cortex of patients with depressive disorder. Psychiatry Research 186, 244–247. https://doi.org/10.1016/j.psychres.2010.07.030
van de Berg, P.J.E.J., Griffiths, S.J., Yong, S.-L., Macaulay, R., Bemelman, F.J., Jackson, S., Henson, S.M., ten Berge, I.J.M., Akbar, A.N., van Lier, R.A.W., 2010. Cytomegalovirus Infection Reduces Telomere Length of the Circulating T Cell Pool. J.I. 184, 3417–3423. https://doi.org/10.4049/jimmunol.0903442
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
Zhan, Y., Hägg, S., 2019. Telomere length and cardiovascular disease risk. Curr Opin Cardiol 34, 270–274. https://doi.org/10.1097/HCO.0000000000000613

Appendix A. PubMed Search Strategy

PubMed Search Strategy
Terms:

1 (tissues OR tissue) (tw)
2 (Epithelium OR epithelial) (tw)
3 "Epithelial Cells" (mesh)
4 Leukocytes (tw)
5 (Bone OR bones) (tw)
6 (Endothelium OR endothelial) (tw)
7 (Epidermis OR Skin) (tw)
8 (membrane OR membranes) (tw)
9 (Muscle OR muscles) (tw)
10 ("Lymph nodes" OR Lymphoid) (tw)
11 blood (tw)
12 liver (tw)
13 (Cortex OR "Cerebral Cortex") (mesh)
14 (neurons OR neuron) (tw)
15 heart (tw)
16 (lungs OR lung) (tw)
17 (Intestines OR intestine) (tw)
18 thyroid (tw)
19 hair (tw)
20 (brain OR brains) (tw)
21 ("Spinal Cord" (mesh) OR "spinal cord") (tw)
22 eyes (tw)
23 ears (tw)
24 mouth (tw)
25 larynx (tw)

26 tonsils (tw)
27 plasma (tw)
28 serum (tw)
29 Granulocytes (tw)
30 buccal (tw)
31 saliva (tw)
32 Foreskin (tw)
33 (Testicles OR testicle OR testes OR testis) (tw)
34 (breast OR breasts) (tw)
35 (ovary OR ovaries) (tw)
36 ("Umbilical Cord" (mesh) OR "Umbilical Cord" OR "Umbilical Cords") (tw)
37 placenta (tw)
38 Chorion (tw)
39 Coronary (tw)
40 (artery OR arteries) (tw)
41 (vein OR veins) (tw)
42 medulla (tw)
43 colon (tw)
44 (Intestine OR intestines) (tw)
45 fat (tw)
46 (Fibroblasts OR fibroblast) (tw)
47 1-47 (OR)
48 NOT (animals (mesh) NOT humans (mesh))
49 "telomere length" (tw)
50 48 AND 49
51 AND (human (filter)) AND (1988:2022 (pdat)) AND (english (Filter))

Appendix B. Example Data Abstraction Spreadsheet

	Article Title	Authors	Publication Year	Sample Size	Racial Demographics	Sex Demographics	Age Demographics	Sample Collection Conditions	Sample Storage Conditions	Measurement Methodology	Types of Tissues	Statistical Method of Correlation(s)	Reported Correlation(s)
Example 1:	Differences in placental telomere length suggest a link between racial disparities in birth outcomes and cellular aging	Jones, Gambala, Esteves, Wallace, Schlesinger, O' Quinn, Kidd, and Theall	2017	46 individuals	Majority Black Americans (34), and 12 White Americans, with 26 male infants and 20 female infants	All tissues collected from pregnant females' placentas	Mean of 27, range of 18 to 41 years	Dissected, washed with 1M PBS, flash frozen with liquid nitrogen (at -80°C) and extracted with QIAamp DNA mini kit	DNA was stored at -35°C	Monochrome multiplex quantitative real-time polymerase chain reaction	Placental amnion, cord, villus, and chorion	Spearman rank correlation coefficients	Amnion and cord: 0.640, amnion and villus: 0.526, amnion and chorion: 0.613. Cord and villus: 0.665, cord and chorion: 0.543. Chorion and villus: 0.540.

Appendix C. Telomere Research Network Telomere Length Data Quality Assurance

Checklist

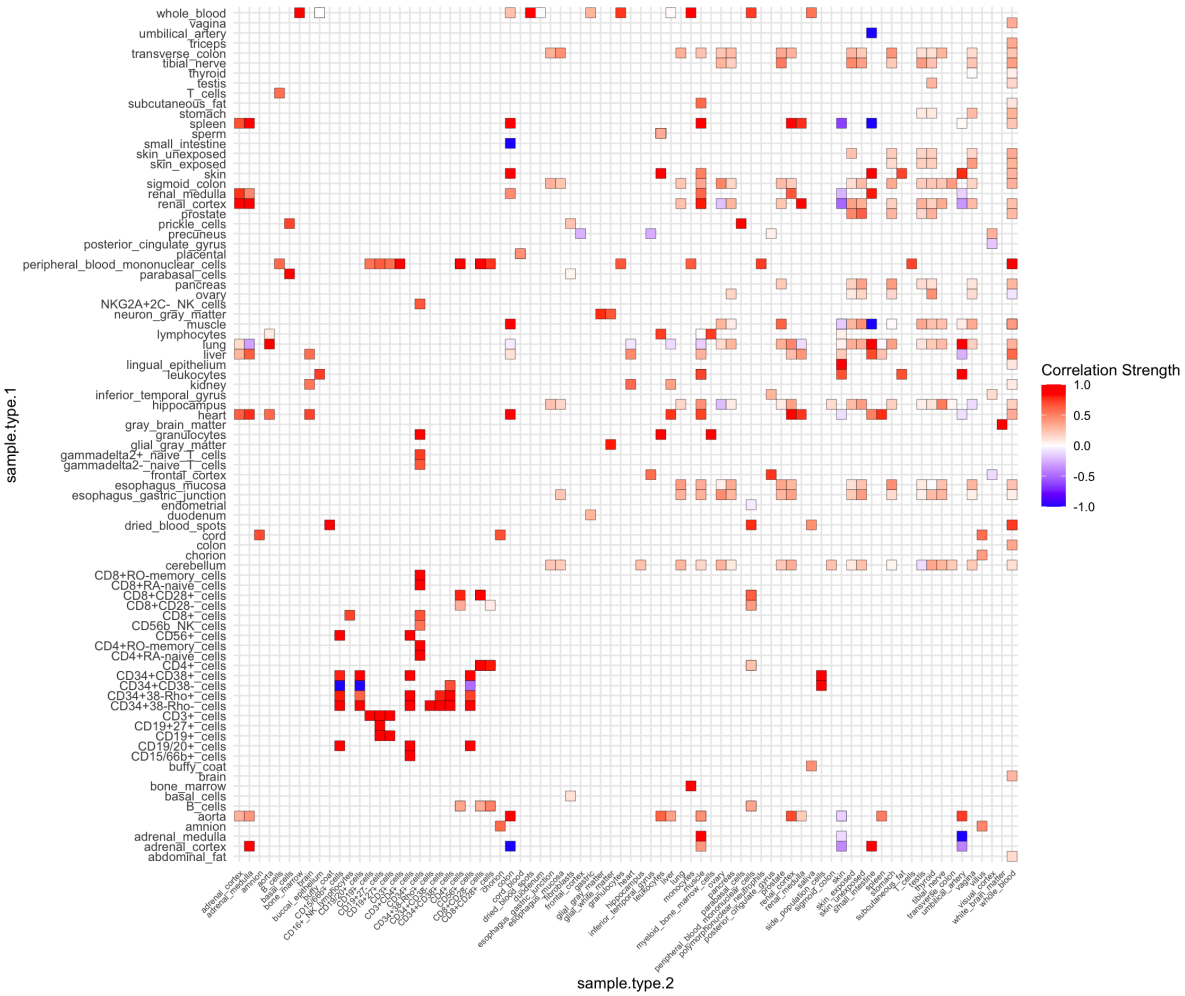
Study Criteria	Yes	No	Other (CD = cannot determine; NA = not applicable; NR = not reported)
Was the sample type identified?			
Were the sample storage temperature, time before extraction, or storage buffer stated?			
Was the DNA extraction method, storage conditions, or DNA quality / integrity stated?			
Was the assay protocol outlined or reference?			
Were sample replicates described?			
Was the source of the calibrator / control sample identified?			
Were quality control steps for telomere length data stated?			
Were indications of assay precision identified (repeatability / intra-assay variation, reproducibility / inter-assay variation, intraclass correlation coefficients, etc.)?			

Appendix D. List of Tissue Types Included in the Meta-Analysis

#	Sample / Tissue Identifier	Categories for Sensitivity Analyses	
1	Abdominal Fat	non-blood	surgically obtainable sample
2	Adrenal Cortex	non-blood	surgically obtainable sample
3	Adrenal Medulla	non-blood	surgically obtainable sample
4	Amnion	non-blood	surgically obtainable sample
5	Aorta	non-blood	surgically obtainable sample
6	B Cells	blood	peripherally available sample
7	Basal Cells	non-blood	peripherally available sample
8	Bone Marrow	non-blood	surgically obtainable sample
9	Brain	non-blood	surgically obtainable sample
10	Buccal Epithelium	non-blood	peripherally available sample
11	Buffy Coat	blood	peripherally available sample
12	CD15/66b+ Cells	blood	peripherally available sample
13	CD16+ NK Lymphocytes	blood	peripherally available sample
14	CD19/20+ Cells	blood	peripherally available sample
15	CD19+ Cells	blood	peripherally available sample
16	CD19+27- Cells	blood	peripherally available sample
17	CD19+27+ Cells	blood	peripherally available sample
18	CD3+ Cells	blood	peripherally available sample
19	CD3+CD4+ Cells	blood	peripherally available sample
20	CD34+ Cells	blood	peripherally available sample

21	CD34+38-Rho- Cells	blood	peripherally available sample
22	CD34+38-Rho+ Cells	blood	peripherally available sample
23	CD34+CD38- Cells	blood	peripherally available sample
24	CD34+CD38+ Cells	blood	peripherally available sample
25	CD4+ Cells	blood	peripherally available sample
26	CD4+RA- Naïve Cells	blood	peripherally available sample
27	CD4+RO- Memory Cells	blood	peripherally available sample
28	CD56+ Cells	blood	peripherally available sample
29	CD56b NK Cells	blood	peripherally available sample
30	CD8+ Cells	blood	peripherally available sample
31	CD8+CD28- Cells	blood	peripherally available sample
32	CD8+CD28+ Cells	blood	peripherally available sample
33	CD8+RA- Naïve Cells	blood	peripherally available sample
34	CD8+RO- Memory Cells	blood	peripherally available sample
35	Cerebellum	non-blood	surgically obtainable sample
36	Chorion	non-blood	surgically obtainable sample
37	Colon	non-blood	surgically obtainable sample
38	Cord	non-blood	surgically obtainable sample
39	Cord Blood	blood	peripherally available sample
40	Dried Blood Spots	blood	peripherally available sample
41	Duodenum	non-blood	surgically obtainable sample
42	Endometrial	non-blood	surgically obtainable sample
43	Esophagus Gastric Junction	non-blood	surgically obtainable sample
44	Esophagus Mucosa	non-blood	surgically obtainable sample
45	Fibroblasts	non-blood	peripherally available sample
46	Frontal Cortex	non-blood	surgically obtainable sample
47	Gamma Delta2- Naïve T Cells	blood	peripherally available sample
48	Gamma Delta2+ Naïve T Cells	blood	peripherally available sample
49	Gastric	non-blood	surgically obtainable sample
50	Glial Gray Matter	non-blood	surgically obtainable sample
51	Glial White Matter	non-blood	surgically obtainable sample
52	Granulocytes	blood	peripherally available sample
53	Gray Brain Matter	non-blood	surgically obtainable sample
54	Heart	non-blood	surgically obtainable sample
55	Hippocampus	non-blood	surgically obtainable sample
56	Inferior Temporal Gyrus	non-blood	surgically obtainable sample
57	Kidney	non-blood	surgically obtainable sample
58	Leukocytes	blood	peripherally available sample
59	Lingual Epithelium	non-blood	peripherally available sample
60	Liver	non-blood	surgically obtainable sample
61	Lung	non-blood	surgically obtainable sample
62	Lymphocytes	blood	peripherally available sample
63	Monocytes	blood	peripherally available sample
64	Muscle	non-blood	surgically obtainable sample
65	Myeloid Bone Marrow Cells	non-blood	surgically obtainable sample
66	Neuron Gray Matter	non-blood	surgically obtainable sample
67	NKG2A+2C- NK Cells	blood	peripherally available sample
68	Ovary	non-blood	surgically obtainable sample
69	Pancreas	non-blood	surgically obtainable sample

70	Parabasal Cells	non-blood	peripherally available sample
71	Peripheral Blood Mononuclear Cells	blood	peripherally available sample
72	Placental	non-blood	surgically obtainable sample
73	Polymorphonuclear Neutrophils	blood	peripherally available sample
74	Posterior Cingulate Gyrus	non-blood	surgically obtainable sample
75	Precuneus	non-blood	surgically obtainable sample
76	Prickle Cells	non-blood	peripherally available sample
77	Prostate	non-blood	surgically obtainable sample
78	Renal Cortex	non-blood	surgically obtainable sample
79	Renal Medulla	non-blood	surgically obtainable sample
80	Saliva	non-blood	peripherally available sample
81	Side Population Cells	blood	peripherally available sample
82	Sigmoid Colon	non-blood	surgically obtainable sample
83	Skin	non-blood	peripherally available sample
84	Skin Exposed	non-blood	peripherally available sample
85	Skin Unexposed	non-blood	peripherally available sample
86	Small Intestine	non-blood	surgically obtainable sample
87	Sperm	non-blood	peripherally available sample
88	Spleen	non-blood	surgically obtainable sample
89	Stomach	non-blood	surgically obtainable sample
90	Subcutaneous Fat	non-blood	surgically obtainable sample
91	T Cells	blood	peripherally available sample
92	Testis	non-blood	surgically obtainable sample
93	Thyroid	non-blood	surgically obtainable sample
94	Tibial Nerve	non-blood	surgically obtainable sample
95	Transverse Colon	non-blood	surgically obtainable sample
96	Triceps	non-blood	surgically obtainable sample
97	Umbilical Artery	non-blood	surgically obtainable sample
98	Vagina	non-blood	surgically obtainable sample
99	Villus	non-blood	surgically obtainable sample
100	Visual Cortex	non-blood	surgically obtainable sample
101	White Brain Matter	non-blood	surgically obtainable sample
102	Whole Blood	blood	peripherally available sample



Supplementary Materials in Heat Map Format. The correlations for each of the 102 tissue types included in the meta-analysis where red is a strong positive correlation and blue is a strong negative correlation between sample types.

CHAPTER 3

Precision MMqPCR Telomere Length Measurement Methodology

Background:

Telomeres are ribonucleoprotein structures at the ends of all eukaryotic chromosomes. The repetitive human telomeric DNA sequence of 5'-TTAGGG-3' is highly conserved across eukaryotic species. The telomere critically protects the integrity of linear DNA and genetic information by preserving chromosome stability. Progressive shortening of telomeres occurs in dividing cells as a result of incomplete lagging-strand DNA synthesis, DNA damage, and other factors (Hemann et al., 2001; Xu et al., 2013). Measuring the length of telomeres has been of increasing scientific interest as studies have reported telomere length (TL) is a biomarker of aging and age-related diseases across the human lifespan.

Meta-associations of TL with overall mortality, environmental exposures, and health outcomes, including cancer, cardiovascular disease, diabetes have been reported (Haycock et al., 2014; Mundstock et al., 2015; Ridout et al., 2016; Astuti et al., 2017; Wang et al., 2018; Hu et al., 2019). These meta-associations derive from studies utilizing one of over two dozen different TL measurement methodologies. Selecting the best TL measurement method for a research study is crucial step to ensuring accurate results, as each method possess its own advantages and disadvantages compared to other methods. The quantitative polymerase chain reaction (qPCR) based method of TL measurement is a widely accepted method for epidemiological studies described in meta-analyses. The qPCR method of TL measurement is less costly than other methods, requires a lower amount of DNA than other methods, and produces more TL measurements per assay than other methods. However, the qPCR method reports an indirect measurement of TL per DNA sample, as it is based on signals from both a telomere sequence

amplicon and single copy gene amplicon. The qPCR method reports the ratio of telomere (T) to single copy gene (S) signals, which is a relative measure within a cohort or experiment unless the T/S ratio is transformed into a Z-score for comparison between studies.

The qPCR method was improved when Richard Cawthon minimized the variation of T/S ratios, increased the throughput, and reduced costs with creation of the monochrome multiplex qPCR (MMqPCR) assay (Cawthon, 2009). The MMqPCR assay possesses the benefits of the qPCR assay, with additional advantages of measuring T and S signals within the same qPCR plate using one reporting fluorophore with results being produced within one business day (Figures 1 and 2). The MMqPCR method for measuring TL is highly reproducible, with expected better precision than qPCR due to the single well design, and higher intraclass correlation coefficients (ICCs) on average due to the telomere and single copy genes being measured within the same well on the same plate. This method is also more efficient, requiring 50% less DNA per sample relative to qPCR, which itself requires less DNA than most TL methods (Axelrad et al., 2013). Furthermore, this multiplex method lowers costs and enhances throughput by processing double the number of samples per assay in the same amount of time as the singleplex qPCR assay.

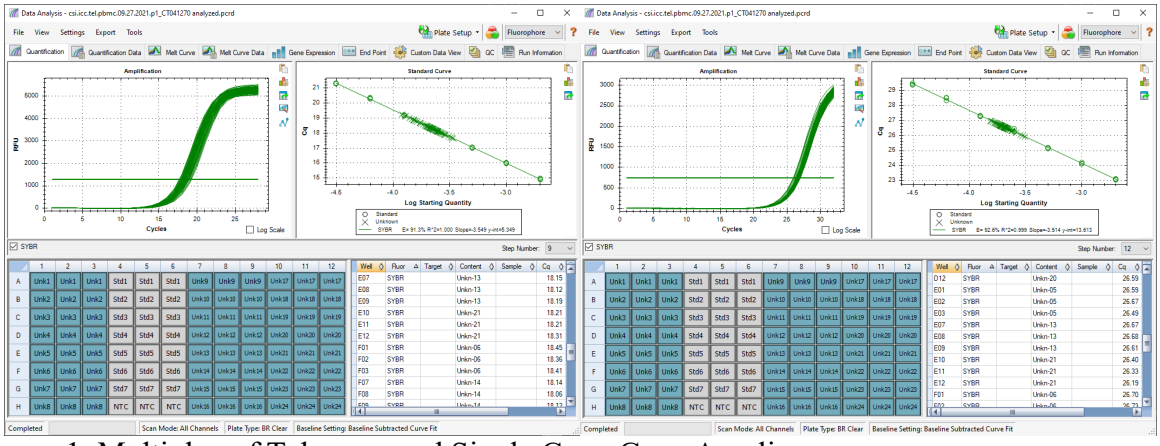
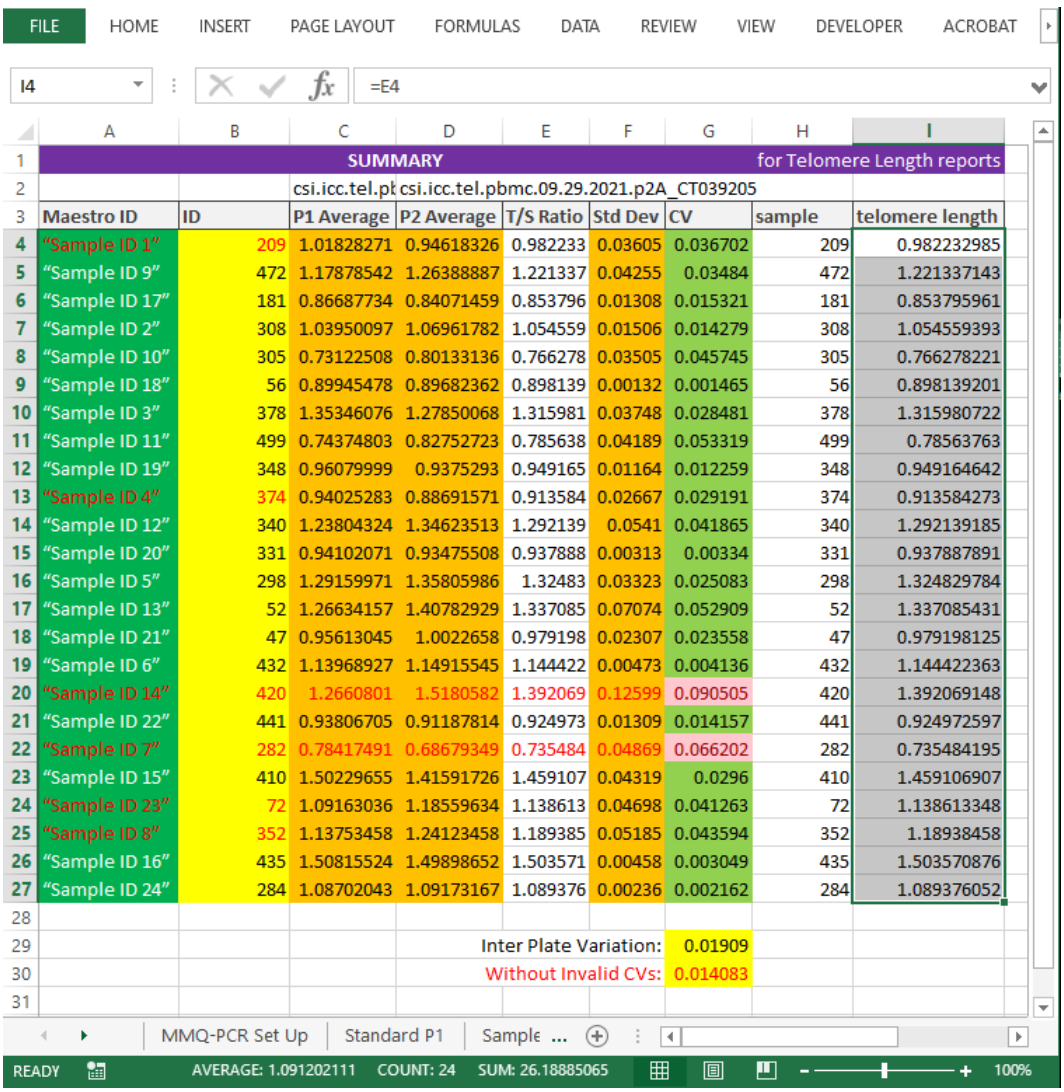


Figure 1. Multiplex of Telomere and Single Copy Gene Amplicons.



FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW DEVELOPER ACROBAT

I4 =E4

	A	B	C	D	E	F	G	H	I
1	SUMMARY								for Telomere Length reports
2			csi.icc.tel.pb	csi.icc.tel.pbmc.09.29.2021.p2A_CT039205					
3	Maestro ID	ID	P1 Average	P2 Average	T/S Ratio	Std Dev	CV	sample	telomere length
4	"Sample ID 1"	209	1.01828271	0.94618326	0.982233	0.03605	0.036702	209	0.982232985
5	"Sample ID 9"	472	1.17878542	1.26388887	1.221337	0.04255	0.03484	472	1.221337143
6	"Sample ID 17"	181	0.86687734	0.84071459	0.853796	0.01308	0.015321	181	0.853795961
7	"Sample ID 2"	308	1.03950097	1.06961782	1.054559	0.01506	0.014279	308	1.054559393
8	"Sample ID 10"	305	0.73122508	0.80133136	0.766278	0.03505	0.045745	305	0.766278221
9	"Sample ID 18"	56	0.89945478	0.89682362	0.898139	0.00132	0.001465	56	0.898139201
10	"Sample ID 3"	378	1.35346076	1.27850068	1.315981	0.03748	0.028481	378	1.315980722
11	"Sample ID 11"	499	0.74374803	0.82752723	0.785638	0.04189	0.053319	499	0.78563763
12	"Sample ID 19"	348	0.96079999	0.9375293	0.949165	0.01164	0.012259	348	0.949164642
13	"Sample ID 4"	374	0.94025283	0.88691571	0.913584	0.02667	0.029191	374	0.913584273
14	"Sample ID 12"	340	1.23804324	1.34623513	1.292139	0.0541	0.041865	340	1.292139185
15	"Sample ID 20"	331	0.94102071	0.93475508	0.937888	0.00313	0.00334	331	0.937887891
16	"Sample ID 5"	298	1.29159971	1.35805986	1.32483	0.03323	0.025083	298	1.324829784
17	"Sample ID 13"	52	1.26634157	1.40782929	1.337085	0.07074	0.052909	52	1.337085431
18	"Sample ID 21"	47	0.95613045	1.0022658	0.979198	0.02307	0.023558	47	0.979198125
19	"Sample ID 6"	432	1.13968927	1.14915545	1.144422	0.00473	0.004136	432	1.144422363
20	"Sample ID 14"	420	1.2660801	1.5180582	1.392069	0.12599	0.090505	420	1.392069148
21	"Sample ID 22"	441	0.93806705	0.91187814	0.924973	0.01309	0.014157	441	0.924972597
22	"Sample ID 7"	282	0.78417491	0.68679349	0.735484	0.04869	0.066202	282	0.735484195
23	"Sample ID 15"	410	1.50229655	1.41591726	1.459107	0.04319	0.0296	410	1.459106907
24	"Sample ID 23"	72	1.09163036	1.18559634	1.138613	0.04698	0.041263	72	1.138613348
25	"Sample ID 8"	352	1.13753458	1.24123458	1.189385	0.05185	0.043594	352	1.18938458
26	"Sample ID 16"	435	1.50815524	1.49898652	1.503571	0.00458	0.003049	435	1.503570876
27	"Sample ID 24"	284	1.08702043	1.09173167	1.089376	0.00236	0.002162	284	1.089376052
28									
29						Inter Plate Variation:	0.01909		
30						Without Invalid CVs:	0.014083		
31									

MMQ-PCR Set Up | Standard P1 | Sample ...

READY AVERAGE: 1.091202111 COUNT: 24 SUM: 26.18885065 100%

Figure 2. Example of Telomere Length Data Produced in the MMqPCR Assay.

Given the advantages of MMqPCR TL measurement, this method is well-suited for population-level studies of TL associations with exposures, health outcomes, and biological process. However, many researchers find commencing the MMqPCR method in a laboratory setting to be demanding given the challenges of setting up the protocol. Here, we describe the TL measurement methodology utilized in our laboratory. Taking advantage of the MMqPCR's repeatability, efficiency, and cost-effectiveness and we describe the detailed steps for conducting this method to increase assay precision and accuracy of the MMqPCR TL measurement methodology for TL researchers.

This article further outlines steps for cleaning data and obtaining a high ICC from MMqPCR TL measurement using duplicate DNA extractions from the same source samples to demonstrate the repeatability of TL measurements generated with this method. Importantly, this article identifies quality control steps conducted that increase comparative information on TL associations across studies. This repeatability, combined with the efficiency and cost-effectiveness of this method, make the MMqPCR TL measurement described here opportune for epidemiological TL research.

Protocol:

1. Stock Reagent, Storage Conditions, and Reagent Preparation

1.1 All stock reagents and materials used in this protocol are listed in Table 1. [Place Table 1 here.]

Table 1. Table of Specific Materials and Equipment.

Material / Equipment	**Company**	**Catalog Number**	**Comments / Description**	**Storage Temperature and Conditions**
PCR Hood	USA Scientific	4263-2588	Nucleic Acid Workstation with HEPA Filtration, AirClean Systems Combination PCR Workstation	Room temperature
HEPA Filter	USA Scientific	Replacement Filters	High-Efficiency Particulate Air Filter for AirClean Workstations	Room temperature
UV Light	USA Scientific	4288-2540	UV Light Bulb for Workstations	Room temperature
Pre-Hood Filter	USA Scientific	4235-3724	Prefilter for AirClean Systems Workstations	Room temperature
Ethanol	Thermo Fisher	T038181000	Absolute Ethanol, 200 Proof, Molecular Biology Grade	Room temperature
TE Buffer	Thermo Fisher	J75893.AE	1x TE Buffer, Molecular Biology Grade, Ultrapure, 100mL	Room temperature
Sodium Acetate	Thermo Fisher	J63560.EQE	3M NaOAc pH 5.2	Room temperature
Ice Plates	Millipore Sigma	Z606634	Eppendorf PCR Cooler for 96 Well plates	-20oC

Syringes	Thermo Fisher	148232A	BD Luer-Lok Disposable Syringes without Needles, 10mL	Room temperature
Syringe Filters	Thermo Fisher	NC1766920	Tisch Scientific Nylon Syringe Filters, 5um, 25mm, Luer-Lok / Luer Slip, Sterile	Room temperature
DTT	Research Products International	D11000-5.0	DL-Dithiothreitol, Cleland's Reagent, 5 grams	-20oC
0.5mL Tubes	USA Scientific	1605-0099	Seal-Rite 0.5mL Microcentrifuge Tubes, Sterile	Room temperature
1.5mL Tubes	USA Scientific	1615-5599	Seal-Rite 1.5mL Microcentrifuge Tubes, Sterile	Room temperature
5mL Tubes	Thermo Fisher	3391276	Argos Technologies Microcentrifuge Tubes – 5mL	Room temperature
15mL Tubes	Thermo Fisher	14-959-53A	Corning 352196 Falcon 15mL Conical Centrifuge Tubes	Room temperature
PCR Strips	Thermo Fisher	AB0776	Low Profile Tubes and Flat Caps, Strips of 8	Room temperature
Kimwipes	Thermo Fisher	06666A	Kimberly-Clark Professional Kimtech Science Kimwipes Delicate Task Wipers, 1-Ply	Room temperature
Loading Trough	Thermo Fisher	14387069	Thermo Scientific Matrix Reagent Reservoirs	Room temperature
Plate Sealing Film	Bio-Rad	MSB1001	Microseal "B" PCR Plate Sealing Film, Adhesive, Optical	Room temperature
96 Well Plate	Bio-Rad	HSP9601	Hard-Shell 96-Well PCR Plates, Low Profile, Thin Wall, Skirted, White / Clear	Room temperature
Weigh Boat	Thermo Fisher	01-549-752	Fisherbrand Sterile Hexagonal Weighing Boat, 10mL	Room temperature
Stainless Steel Spatula	Thermo Fisher	3990240	Bel-Art SP Scienceware Stainless-Steel Sampling Spoon and Spatula	Room temperature
Scale	Thermo Fisher	01-922-329	OHAUS 30430060 PR Series Analytical Balance, 62g Capacity	Room temperature
Popspin	Genesee Scientific	31-500B	Poseidon 31-500B Mini Centrifuge, Blue Lid	Room temperature
Plate Spinner	Thermo Fisher	14-100-141	Fisherbrand Mini Plate Spinner Centrifuge, 230 V	Room temperature
Vortex	Thermo Fisher	14-955-151	Fisherbrand Mini Vortex Mixer, 115 V, 50/60 Hz	Room temperature

Sharpies	Sharpie	2151734	Brush Twin Permanent Markers, Black	Room temperature
Aluminum Foil	Office Depot	3489072	Reynolds Wrap Standard Aluminum Foil Roll, 12" x 75', Silver	Room temperature
Scissors	Office Depot	458612	Office Depot Brand Scissors, 8”, Straight, Black, Pack of 2	Room temperature
Pipette Tips (Single Channel)	USA Scientific	1181-3850 1180-1850 1111-0880 1111-2890	10µL Graduated TipOne RPT Filter Tips 20µL Beveled TipOne RPT Filter Tips 200µL Natural TipOne Pipette Tips in Racks 1000µL Natural Graduated TipOne Pipette Tips in Racks	Room temperature
Pipettors (Single Channel)	Thermo Fisher	F144802G F123600 F123601 F123602	Gilson Pipetman Classic Pipets, 1 to 10µL Gilson Pipetman Classic Pipets, 2 to 20µL Gilson Pipetman Classic Pipets, 20 to 200µL Gilson Pipetman Classic Pipets, 200 to 1000µL	Room temperature
Pipette Tips (Multichannel)	Ranin	17005860	Pipette Tips SR LTS 20µL F 960A/5, 20µL Maximum	Room temperature
Pipettors (Multichannel)	Ranin	17013802 17013803	Pipet-Lite Multi Pipette L8-10XLS, 0.5 to 10µL Pipet-Lite Multi Pipette L8-20LS+, 2 to 20µL	Room temperature
Big Tube Rack	Thermo Fisher	344817	Fisherbrand 4-Way Tube Rack	Room temperature
PCR Tube Rack	Thermo Fisher	344820	Fisherbrand 96-Well PCR Tube Rack	Room temperature
Small Tube Rack	Thermo Fisher	21-402-17	Thermo Fisher 8601 Reversible Microtube Racks with Lid	Room temperature
SYBR Green	Thermo Fisher	S7563	SYBR Green I Nucleic Acid Gel Stain – 10,000X Concentrate in DMSO	-20oC, away from light
AmpliTaq Gold Kit – Polymerase and Buffer	Thermo Fisher	4311806	AmpliTaq Gold DNA Polymerase with Gold Buffer and $MgCl_2$ ($MgCl_2$ in this kit is not used), 10X Gold Buffer, 2.5U AmpliTaq Gold Polymerase	-20oC
1M $MgCl_2$	Thermo Fisher	50152107	Biotang Inc 1M $MgCl_2$ 1M Magnesium Chloride Solution, Prepared in 18.2 Megohms Water and	4oC

			Filtered through 0.22 Micron Filter	
25mM dNTPs	New England BioLabs	N0446S	Deoxynucleotide Solution Set	-20oC
Betaine	Thermo Fisher	AAJ77507AB	Betaine, 5M Solution, Molecular Biology Grade, Ultrapure, 10mL	-20oC
PCR Grade H_2O	Thermo Fisher	AM9937	Nuclease-Free Water (not DEPC-Treated)	Room temperature
1x Gold Buffer	In House	Not Applicable	10X Gold Buffer diluted with PCR Grade H_2O	-20oC
100mM DTT	In House	Not Applicable	Made with stock DTT, diluted sodium acetate, and PCR Grade H_2O	-20oC
Telomere Forward Primer	Integrated DNA Technologies	Custom	See separate table	-20oC when hydrolized
Telomere Reverse Primer	Integrated DNA Technologies	Custom	See separate table	-20oC when hydrolized
Single Copy Gene Forward Primer	Integrated DNA Technologies	Custom	See separate table	-20oC when hydrolized
Single Copy Gene Reverse Primer	Integrated DNA Technologies	Custom	See separate table	-20oC when hydrolized
CFX96 Optical Reaction Module for Real-Time PCR Systems with Starter Package	Bio-Rad	1845096	96-well optical module for real-time PCR	Room temperature
CFX Maestro Software	Bio-Rad	12004110	Software for real-time PCR plate setup, data collection, statistics, and graphing of results	Room temperature
Microsoft Excel	Microsoft	Not Applicable	Microsoft 365 package, Excel software application	Room temperature
R Software	The R Project for Statistical Computing	Not Applicable	R version 4.2.2	Room temperature

1.2 Store 1X TE in 5mL tube aliquots at room temperature for up to two years.

1.3 Store SYBR Green in 3µL aliquots in individual PCR strip tubes at -20 °C for up to two years away from light; ensure complete thawing before aliquoting and wrap aliquots in aluminum foil after using a part of the aliquot in master mix preparation.

1.4 Store AmpliTaq Gold DNA Polymerase in original tubes and 10X Gold Buffer in 1.5mL tubes in 660µL aliquots at -20 °C for up to one year; ensure complete thawing before aliquoting.

1.4.1 Make 1X Gold Buffer by adding 9.9µL of 10X Gold Buffer to 89.1µL PCR grade H2O in 0.5mL tubes; label the tubes and store at -20 °C for up to one year.

1.5 Store Betaine in 5mL tubes in 1,280µL aliquots at -20 °C away from light for up to two years; ensure complete thawing before aliquoting.

1.6 Store MgCl2 in 0.5mL tubes in 70µL aliquots at 4 °C away from light for up to one year.

1.7 Store telomere and single copy gene oligonucleotide lyophilized primers at room temperature. Make telomere and single copy gene oligonucleotide primers by ordering the custom standard DNA oligos from Integrated DNA Technologies as defined in Table 2, and place them in the PCR hood (here, the single copy gene is Albumin; if choosing a different single copy gene, the master mix amounts may need to be adjusted to ensure acceptable PCR efficiency).

Table 2. Telomere and Single Copy Gene Oligonucleotide Primer Sequences.

According to Cawthon 2002:	**Custom Ordered Sequence:**
	Primer specifications include: quantity of XX
Telomere Forward Primer (telg)	5' – ACA CTA AGG TTT GGG TTT GGG TTT GGG TTT GGG TTA GTG T – 3'
Telomere Reverse Primer (telc)	5' – TGT TAG GTA TCC CTA TCC CTA TCC CTA TCC CTA TCC CTA ACA – 3'
Single Copy Gene Forward Primer (albd2)	5' – GCG GGC CCG CGT GGC GGA GCG AGG CCG GAA AAG CAT GGT CGC CTG T – 3'
Single Copy Gene Reverse Primer (albu2)	5' – GCC TCG CTC CGG GAG CGC CGC GCG GCC AAA TGC TGC ACA GAA TCC TTG – 3'

1.7.1 Obtain 1X TE buffer and PCR tube strips and place in PCR hood.

1.7.2 Check each individual oligo tube for nmol concentration ([nmol]= alpha), then vortex each oligo tube for 10 seconds, place tubes in popspin for 5 seconds, and observe where the dehydrated oligo is in the tube to ensure the oligo is at the bottom of the tube before adding 1X TE.

1.7.3 Rehydrate each oligo tube by adding 1X TE equal to 10 times the alpha (e.g., if the nmol concentration on the tube reads 24.6nmol, then add 246µL 1X TE for up to one year in the oligo tube) to create a 100µM solution of each primer.

1.7.4 Label PCR strip tubes with “telg”, “telc”, “albd2” or “albu2” and month and year and fill with the respective oligo.

1.7.5 For “telg” and “telc”, aliquot 16µL volumes then close tubes and cut between PCR tubes and closed tops with scissors.

1.7.6 For “albd2” and “albu2”, aliquot 11µL volumes then close tubes and cut between PCR tubes and closed tops with scissors.

1.7.7 Store all oligos at -20 °C.

1.8 Obtain the dNTPs, stored at -20 °C for up to one year, a 5mL tube, and 0.5mL tubes and place in hood.

1.8.1 Vortex each dNTP tube for 10 seconds after thawing and popspin for 5 seconds.

1.8.2 In the 5mL tube, add equal parts of each of the four dNTPs, approximately 225µL from each tube, then vortex the 5mL tube well and wait for the dNTP mix to be at the bottom of the tube fore aliquoting in 210µl amounts in 0.5mL tubes and store at -20 °C for up to one year.

1.9 Obtain sodium acetate, DTT, PCR grade H2O, two 15mL tubes, 0.5mL tubes, a spatula, weigh boat, a loading trough, a 10mL syringe, and a syringe filter and place items in PCR hood (CAUTION: DTT is a hazardous reagent, it is harmful if swallowed, causes skin irritation, and causes serious eye damage; wash thoroughly after handling, do not eat, drink, or smoke when using DTT, wear protective gloves, eye protection, and face protection when handling, avoid breathing DTT vapors, and avoid prolonged or repeated exposure).

1.9.1 Make 10mL 0.01M sodium acetate by adding 33.33µl 3M sodium acetate and 9,967µl PCR grade H2O into a 15mL tube and vortex well.

1.9.2 Measure out 0.1545g DTT in a weigh boat on the scale using the stainless-steel spatula, then scoop this into the 15mL tube of 0.01M sodium acetate and wash the weigh boat with the

solution using the pipettes, vortexing the tube once all the DTT is in the solution until the powder dissolves fully.

1.9.3 Pour the 15mL tube into the loading trough, suck the entire solution into a plastic syringe from a corner of the loading trough.

1.9.4 Attach the syringe filter to the end of the filled syringe and slowly drip the solution through a clean 15mL tube by lightly pushing the syringe plunger down and solution dripping through the filter, ensuring the solution is dripped into the 15mL tube.

1.9.5 Aliquot a reasonable amount of this solution in 200µL volumes into 0.5mL tubes, label tubes with date and store at -20 °C for up to two months.

2. Genomic DNA Extraction and Sample Preparation

2.1 Perform a genomic DNA extraction for samples in accordance with manufacturer's guidelines using kits or established methods within the laboratory; here, DNA from peripheral blood mononuclear cells (PBMCs) were extracted using the QIAmp DNA Blood Mini Kit (QIAGEN; Hilden, Germany) per manufacturer guidelines.

2.1.1 Check the DNA sample quality by spectrophotometer and Qubit assay, using the High Sensitivity or Broad Range dsDNA Qubit kit according to manufacturer guidelines (Thermo Fisher; Waltham, Massachusetts); samples with unacceptable 260/280 and 260/230 ratios and concentrations of dsDNA below Qubit detection should not be analyzed for telomere length (Wilfinger et al., 1997; Boesenberg-Smith et al., 2012).

2.1.2 Dilute the samples with PCR grade H2O to be at the appropriate amount for MMqPCR as required; where the total sample amount is 15 divided by a (diluted if necessary) Qubit concentration of double-stranded DNA, and the water amount added to tubes in a PCR strip is 75µL minus the total sample amount; a sample MMqPCR template for identifying the correct

dilution factor is available on the Telomere Research Network webpage (MMqPCR Set Up Sheet, Column Q).

2.1.3 Put samples into PCR strips with Samples A1-8 in PCR strip A, repeat for PCR strips B and C, as detailed in the MMqPCR Set Up Sheet and depicted in Table 5 with appropriate sample amounts (MMqPCR Set Up Sheet, Column S) and PCR grade H2O amounts (MMqPCR Set Up Sheet, Column T); place these completed sample PCR strips into a PCR strip rack until ready to prepare the plates; samples in PCR strips can be stored overnight at 4 °C to be plated the next day.

2.2 Perform a genomic DNA extraction on a control sample of the same sample type as the samples being assayed for a given project, using the same DNA extraction procedure that was used for that project. For example, if the samples to be analyzed are all blood, create a control sample by extracting DNA from control blood. Here, a control PBMC DNA sample was created using the same DNA extraction protocol utilized for the sample DNA extractions.

2.2.1 Check the control DNA sample quality by spectrophotometer and Qubit assay, using the High Sensitivity or Broad Range dsDNA Qubit kit according to manufacturer guidelines (Thermo Fisher; Waltham, Massachusetts); the control sample should not have unacceptable 260/280 and 260/230 ratios and concentrations of dsDNA below Qubit detection (Wilfinger et al., 1997; Boesenberg-Smith et al., 2012).

2.2.2 One control DNA sample stock should be utilized for samples from the same cohort or research question, to ensure comparability between telomere length assays; aliquot control DNA at a concentration of 2ng/μL in 150μL amounts in 0.5mL tubes, labeled with the sample type and date and stored at -20 °C for up to 5 years.

2.2.3 Prepare a control standard curve by taking an individual control DNA sample, and after thawing if needed, vortex for 30 seconds, popspin for 5 seconds, and aspirate the control into the first tube in a PCR tube strip.

2.2.4 Pipette 70µL of PCR grade H2O into the second through eighth tubes of the PCR tube strip. Then aspirate 70µL of the control DNA sample in the first tube and move this to the second PCR tube, wait 30 seconds, and then resuspend the solution in the second PCR tube, aspirate 70µL of this and move to the third PCR tube. Repeat this for tubes three though six for a 2-fold serial dilution of 7 standards, then place this completed control standard PCR strip into the PCR strip rack until ready to prepare the plates; the control standard serial dilution can be stored overnight at 4 °C to be plated the next day.

3. MMqPCR Master Mix Preparation

3.1 Gather the reagent aliquots to make the master mix for the MMqPCR assay as listed in Table 3. [Insert Table 3 here.]

Table 3. Final Volumes and Concentrations of Reagents.

Reagent	**Volume per Well**	**Concentration in Master Mix**	**Volume of Reagent Aliquots**
telg oligo	0.05626µL	0.9µM	16µL
telc oligo	0.05626µL	0.9µM	16µL
albd2 oligo	0.0375µL	0.6µM	11µL
albu2 oligo	0.0375µL	0.6µM	11µL
SYBR Green	0.1875µL	0.75X	3µL
Gold Buffer (1X and 10X)	2.5µL	1X	99µL (1X) and 660µL (10X)
dNTPs	0.8µL	0.8mM	210µL
$MgCl_2$	0.25µL	10mM	70µL
DTT	0.75µL	3mM	200µL
Betaine	5µL	1M	1,280µL
AmpliTaq Gold	0.5µL	2.5U/µL	200µL
PCR Grade H_2O	4.2625µL	Not applicable	4,000µL

3.1.1 Obtain a Betaine aliquot, PCR grade H2O, a 10X Gold Buffer aliquot, a 1X Gold Buffer aliquot, a SYBR Green aliquot, a MgCl2 aliquot, a DTT aliquot, a dNTP aliquot, and aliquots of each of the four oligos and place in the PCR hood.

3.1.2 After each aliquot is thawed at room temperature, vortexed and popspinned, take 1,235.2µL of PCR grade H2O and place it into the 5mL tube holding the Betaine (now the master mix tube); take 640µL of the 10X Gold Buffer and add it to the master mix; take 64µL of the MgCl2 and add it to the master mix; take 192µL of the DTT and add it to the master mix; take 204.8µL of the dNTPs and add it to the master mix; take 14.4µL of each Telomere primer and slowly add each to the master mix; and take 9.6µL of each Single Copy Gene primer and slowly add each to the master mix.

3.1.3 Take 1µL of the SYBR Green aliquot and add it to the 1X Gold Buffer aliquot; vortex this for 10 seconds then popspin; add 48µL of this SYBR Green / 1X Gold Buffer aliquot to the master mix.

3.1.4 Take the AmpliTaq Gold DNA Polymerase out of -20°C storage and take 128µL of this and slowly add it to the master mix, then vortex the 5mL master mix tube for 30 seconds and set back into big tube rack; move AmpliTaq Gold DNA Polymerase immediate back to -20 °C storage.

4. 96-Well Plate Preparation

4.1 Get the ice plate holders from the freezer and place in PCR hood.

4.2 Get two 96-well plates and place in PCR hood inside ice plate holders, which will run simultaneously on two different thermocyclers.

4.2.1 Label each plate with technician initials, plate number (1 or 2), and date on the bottom edge of each plate, see Figure 3.

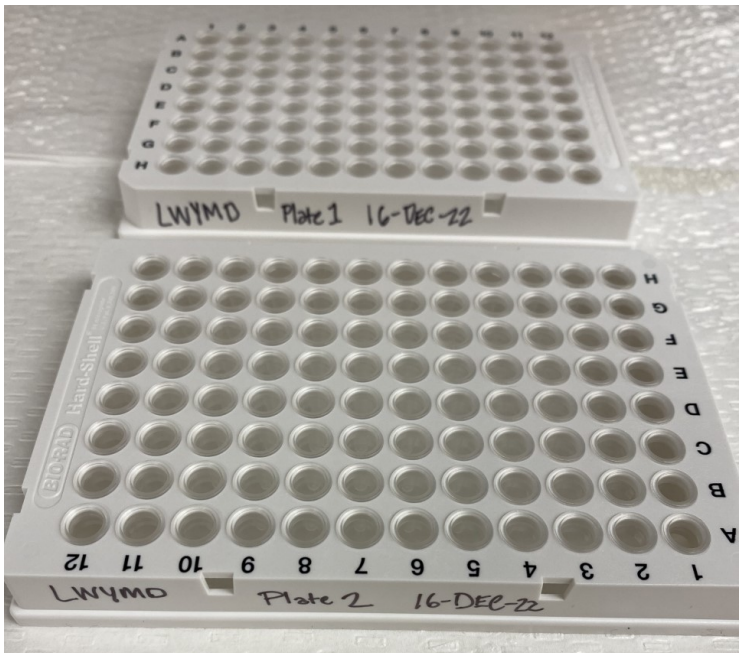


Figure 3. Example Labelling for P1 and P2 Plates.

4.3 Get two Ranin pipette tip boxes and remove plastic packaging before placing boxes into the PCR hood.

4.4 Get a loading trough and place in PCR hood.

4.4.1 Pour the vortexed master mix into the loading trough.

4.4.2 Use a 200µL DNA pipette tip to push remaining master mix to the bottom of the tube then aspirate into the loading trough.

4.5 Get the 2 to 20µL DNA multichannel pipette, set the pipette to 15µL, open a Ranin pipette tip box and load tips onto multichannel, pushing at the base of each pipette tip to ensure a tight seal.

4.5.1 Make sure that the tips do not have any filter within the region where the master mix will be aspirated; if there is a piece of filter then replace the tip.

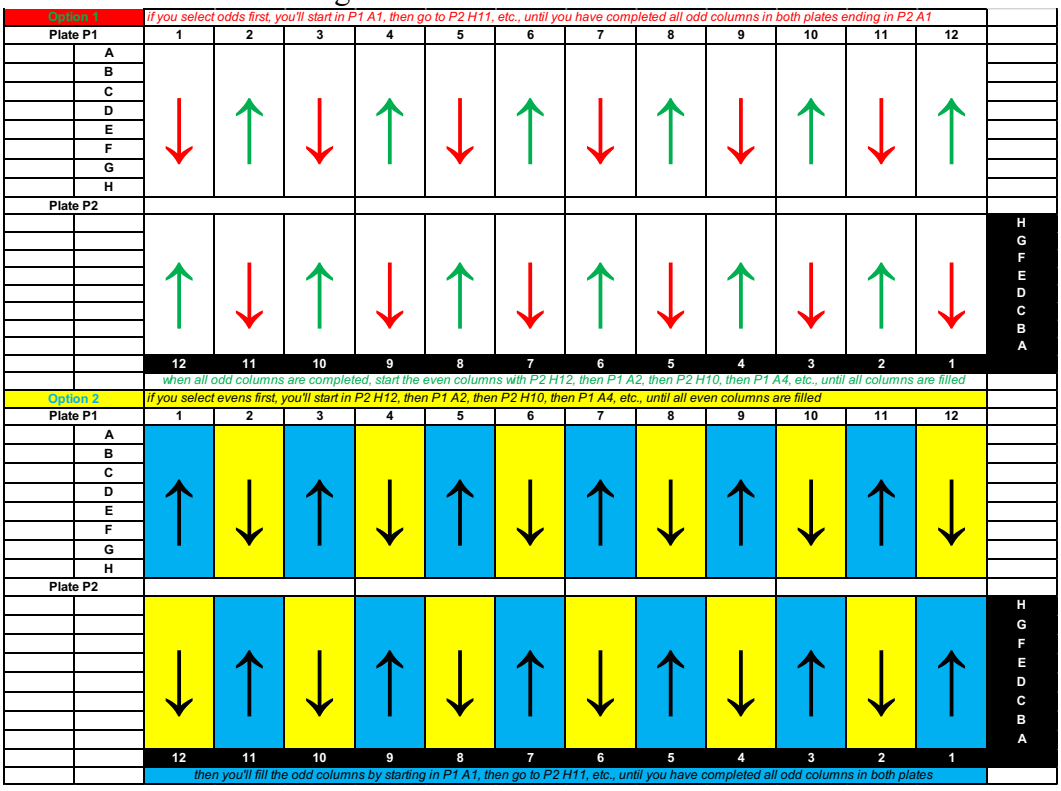
4.6 Swirl the loading trough before aspirating 15µL into the multichannel pipette tips, making sure the tips fill at the same time with the same volume in each tip.

4.6.1 Expel the master mix into each column of both plates in the following scheme:

4.6.2 Turn the P1 plate 180 degrees so the numbers of the columns are closest to the technician.

4.6.3 Choose odds or even numbers, and start with the column with the odd or even number from left to right, expelling the solution into one plate that has the first number in the chosen series, then moving to the other plate with the number that has the next odd or even number (whichever was chosen to fill first, i.e. if the technician choses odds and started with 1 on the P2 plate then fill 11 on the P1 plate); after the technician has filled all of the wells of the series (odd or even) do the other series (odd or even) of columns in the same fashion. See Table 4 for a flow chart of this process. [Insert Table 4 here]

Table 4. Process for Filling Plates.



4.6.4 When expelling master mix into a column, dispense a third at the bottom of the well, then pull the pipette up the well while pressing the tip against the wall of the well, then expel the rest halfway up the well making sure there are no bubbles created, and lastly flick the tips out of the well.

4.6.5 Remove the tips and get new pipette tips for each column of both plates using 1 Ranin pipette tip box in total, then throw out the loading trough when all 96-wells in both plates are filled with 15µL of master mix.

4.7 Once the master mix has been dispensed, vortex then popspin for 5 seconds the four closed PCR strips containing the samples and standard curve.

4.7.1 Line up the PCR strips in the PCR tube rack in the order they will be place in the plate, here the first 3 columns are samples (PCR strip A), columns 4-6 are the standard (PCR strip SC), columns 7-9 are samples (PCR strip B), and columns 10-12 are samples (PCR strip C) as seen in Table 5. [Insert Table 5 here.]

Table 5. Organization of Samples and Control Standard in Plates.

Example Organization	*PCR strip A*			*Control Standard*			*PCR strip B*			*PCR strip C*		
	1	2	3	4	5	6	7	8	9	10	11	12
A	*Sample A1*			*Standard at 2.0*			*Sample B1*			*Sample C1*		
B	*Sample A2*			*Standard at 1.0*			*Sample B2*			*Sample C2*		
C	*Sample A3*			*Standard at 0.5*			*Sample B3*			*Sample C3*		
D	*Sample A4*			*Standard at 0.25*			*Sample B4*			*Sample C4*		
E	*Sample A5*			*Standard at 0.125*			*Sample B5*			*Sample C5*		
F	*Sample A6*			*Standard at 0.0625*			*Sample B6*			*Sample C6*		
G	*Sample A7*			*Standard at 0.03125*			*Sample B7*			*Sample C7*		
H	*Sample A8*			*Negative Control*			*Sample B8*			*Sample C8*		

4.8 Turn both plates 180 degrees so the numbers of the columns are flipped in relation to the technician for each plate (if the technician saw the label for P2 previously, now they will see the label for P1 on the plate edges).

4.9 Get the 0.5 to 10µL DNA multichannel pipette, set the pipette to 10µL, open a Ranin pipette tip box and load tips onto multichannel, pushing at the base of each pipette tip to ensure a tight seal.

4.9.1 Make sure that the tips do not have any filter within the region where the samples or control standard will be aspirated; if there is a piece of filter then replace the tip.

4.9.2 Aspirate and dispense 10µL of the samples and standard curve solutions using the multichannel pipette, starting with the leftmost PCR strip in the PCR strip rack.

4.9.3 Open the PCR strip and fill the first 3 columns of each plate by alternating the columns in a similar fashion as the master mix was placed in the plate.

4.9.4 Remove the tips and get new pipette tips for each column of both plates using 1 Ranin pipette tip box in total, then throw out the PCR strips when all 96-wells in both plates are filled with the appropriate samples or control standard.

4.10 Once the plates are filled with master mix and samples or control standard, remove the 96-well plates from the ice plates and obtain 2 plate sealing films.

4.10.1 Tap the plates gently on the bench and cover with sealing films, using fingernails to seal the edges and Kimwipes to run over each well to ensure a good seal with the film.

4.10.2 Mix the plates by swirling the plates on the hood top for 15 seconds, then place the sealed plates into the plate spinner for 2 minutes with the well openings facing the center of the plate spinner.

4.10.3 Once the 2 minutes in the plate spinner has passed, place the plates in a thermocycler with the numbers of the columns at the top of each plate in legible order, wiping a Kimwipe on the top of each plate before closing the thermocycler top.

5. MMqPCR Thermocylcing

5.1 While the plates are spinning, turn on the computer and the thermocyclers.

5.1.1 Open the thermocycler software; here the CFX Maestro software is used.

5.2 Create the telomere length multiplex protocol in accordance with Cawthon's original multiplex thermocycling protocol (Cawthon, 2009) as depicted in Figure 4:

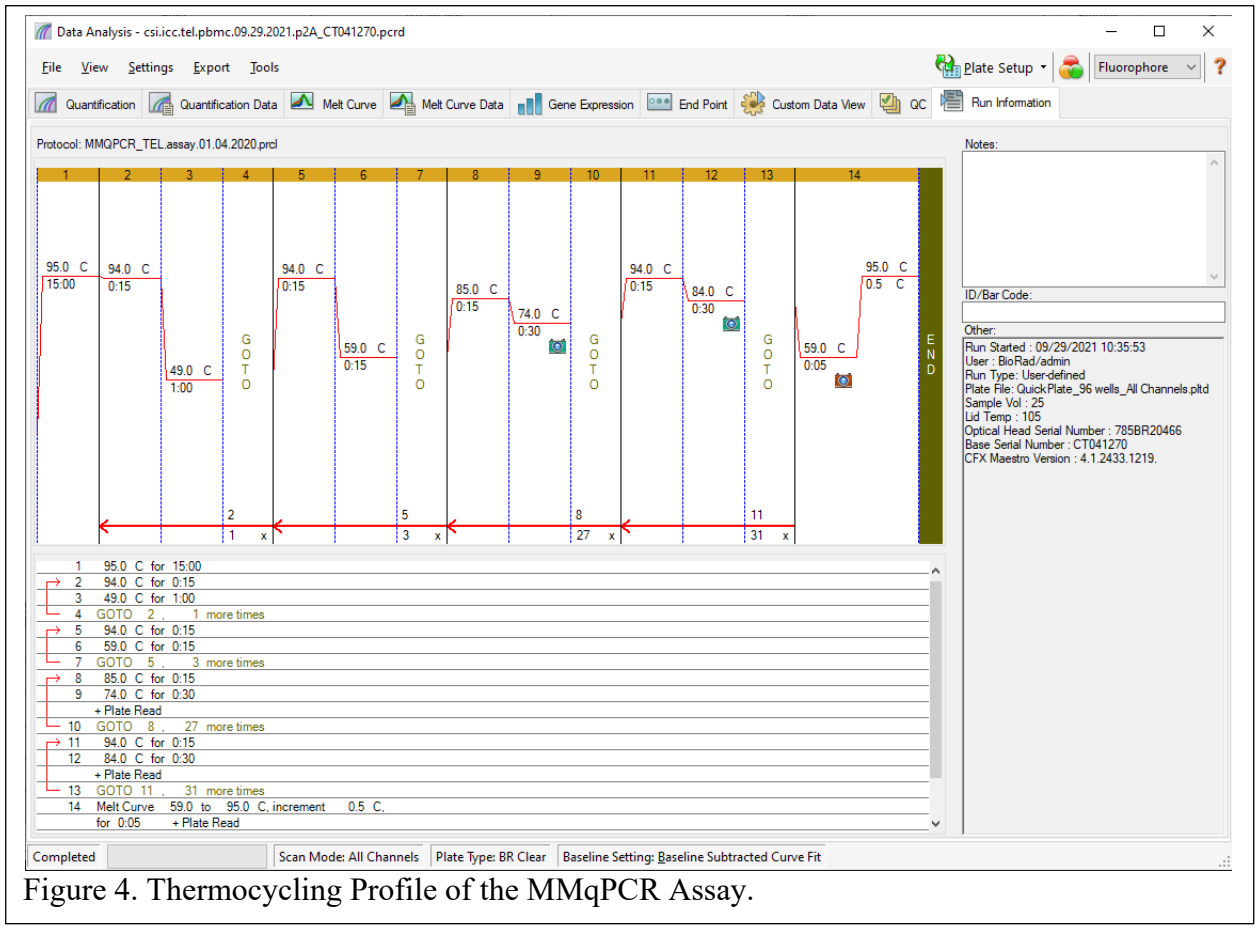


Figure 4. Thermocycling Profile of the MMqPCR Assay.

5.2.1 Add an incubation step to activate the AmpliTaq Gold DNA Polymerase, 95 oC for 15 minutes.

5.2.2 To avoid primer-dimer binding, run two cycles of 94 oC for 15 seconds, 49 oC for 1 minute, then three cycles of 94 oC for 15 seconds, 59 oC for 15 seconds.

5.2.4 For telomere amplification, twenty-seven cycles of 85 oC for 15 seconds, 74 oC for 30 seconds, then signal acquisition.

5.2.5 For single copy gene amplification, thirty-one cycles of 94 oC for 15 seconds, 84 oC for 30 seconds, then signal acquisition.

5.2.6 As a best practice, include a melt curve from 59 oC to 95 oC at 5 seconds intervals for each increasing degree in the thermal cycling protocol.

5.3 Click “Start Run” for both thermocyclers.

5.4 Lastly, name the analysis file including the following: cohort or experiment name, type of tissue, plate number (respective to which plate is in each thermocycler), date, and machine number of the thermocycler for that plate.

6. Telomere Length Data Analysis

6.1 Once the thermocycling is complete, analyze the data in the following manner to produce telomere length values for the samples ran:

6.2 In the software (here CFX Maestro is used), select the "Plate Setup" feature to identify where the standard control serial dilution is located on each plate.

6.2.1 Highlight all wells and select "Select Fluorophores" then check the box labeled "SYBR" and uncheck all other boxes, then click "OK."

6.2.2 While all wells are still highlighted, under the word "Load", check the box next to "SYBR". Now all the wells should have "SYBR" written on them.

6.2.3 Next, highlight the three non-template control wells at the bottom of the standard control serial dilution and select "NTC" from the sample type menu on the right. The three wells should be yellow and be called NTC.

6.2.4 Highlight the twenty-one wells of the standard control serial dilution and select "Standard" from the sample type menu. These well should be green. While these wells are still highlighted, select "Replicate Series," select "3" from the "Replicate Size" drop down menu, select "horizontal" then select "Apply." These wells should be labeled in sets of three from Std-1 though Std7.

6.2.5 While the standard is still highlighted select "Dilution Series," in the "Dilution Factor" field, type "2" then enter the dilution starting concentration and directionality according to the plate number as follows:

6.2.6 For the P1 plate, enter “2.00E-3” in the “Starting Concentration” field, and check the box for “Decreasing” then select “Apply.” The values in the twenty-one wells should have the concentration values written in each well ranging from 2.00E-03 to 3.13E-05 top to bottom.

6.2.7 For the P2 plate, enter “3.13E-5” in the “Starting Concentration” field, and check the box for “Increasing” then select “Apply.” The values in the twenty-one wells should have the concentration values written in each well ranging from 3.13E-05 to 2.00E-03 top to bottom.

6.3 For the samples, highlight the columns for PCR strip A, select “Unknown” from the sample type menu, then select “Replicate Series,” select “3” from the “Replicate Size” drop down menu, then select “Horizontal” and select “Apply.” The wells for these columns should be blue and labeled in by row in sets of 3 from Unk-1 to Unk-8.

6.3.1 Repeat step 6.3 for columns for PCR strip B, and they should be labeled Unk-9 through Unk-16.

6.3.2 Repeat step 6.3 for columns for PCR strip C, and they should be labeled Unk-17 through Unk-24.

6.4 Select “OK” at the bottom right of the “Plate Editor” window.

6.5 For quality control of the PCR assay, ensure that curves in the “Quantification” tab are appropriate, e.g., no inverted amplification curves, for both “Step 9” and “Step 12” which are the telomere and single copy gene amplicons respectively and are accessed by the drop down menu on the middle right side of the software window.

6.5.1 Ensure that the PCR efficiencies reported on both “Step 9” and “Step 12” are between 90% and 110% and these two efficiencies do not differ from each other more than 10%, i.e. 92.4% for “Step 9” and 98% for “Step 12” is appropriate but 92.4% for “Step 9” and 108% for “Step 12” is not appropriate, and that the NTCs do not have amplification.

6.5.2 These quality control steps should be met before analyzing data for individual samples for both P1 and P2 plates.

6.6 Following the sample MMqPCR template from the Telomere Research Network website, select the twenty-one wells in the P1 analysis file that correspond to the standard control serial dilution. Ensure “Step 9” is selected. Highlight these “Cq” values in the lower right corner of the software window, right click and select “Copy.” Paste these values into the Standard 1 Sheet, column B. Select “Step 12” then right click and select “Copy” for these “Cq” values. Paste these values into the Standard 1 Sheet, column I.

6.6.1 Identify the slope value in “Step 9” and type this into cell C4 in the Standard 1 Sheet. Identify the slope value in “Step 12” and type this into cell J4 in the Standard 1 Sheet.

6.6.2 Ensure that the coefficient of variation (CV) for each dilution triplicate is less than 0.1. To correct this, up to three individual wells of the twenty-one standard control serial dilution wells may be excluded from the analysis while the above quality control steps are met, step 6.6 is repeated with each exclusion executed, and the sample readings fall within the standard points in the “Quantification” graphs for “Step 9” and “Step 12.”

6.6.3 Repeat steps 6.6 – 6.6.2 for the P2 analysis file and corresponding Standard 2 Sheet.

6.6.4 When both Standard 1 and Standard 2 Sheets are completed with acceptable CVs, individual sample data may be collected from the analysis files. Ensure quality control adjustments are reported in the P1 v P2 Sheet column L, e.g., in cell L22 number how many wells were excluded in the P1 analysis file.

6.7 Select the seventy-two wells in the P1 analysis file “Quantification” tab, ensuring “Step 9” is selected. Highlight these “Cq” and “SQ” values in the lower right corner of the software window, right click and select “Copy.” Paste these values into the Samples P1 Sheet, starting in

cell D3. Select “Step 12” then right click and select “Copy” for these “Cq” and “SQ” values, the paste these values into the Samples P1 Sheet, starting in cell F3.

6.7.1 Repeat step 6.7 for the samples in the P2 analysis file and corresponding Samples P2 Sheet.

6.7.2 With the automatically calculated telomere to single copy gene (T/S) ratios (column H), average T/S ratios (column I), standard deviations (column J), and CVs (column K) in Samples P1 and Samples P2 Sheets, ensure that the CVs for each sample triplicate are less than 0.1. To correct any CVs greater than 0.1, one T/S ratio from the six measurements per sample may be deleted, e.g., one T/S ratio from either the Samples P1 Sheet or the Samples P2 Sheet may be deleted per individual sample. Only samples passing this quality control may be included in final telomere length results, otherwise the sample should be repeated and deleted from P1 v P2 and ICC Data Sheets for this run.

6.7.3 For quality control, check the individual sample interplate CVs in the P1 v P2 Sheet, column G. These should be less than 0.05; to correct any CVs greater than 0.05, perform the exclusion of one T/S ration from the six measurements per sample across Sheets, as described in 6.7.2, e.g., if one of the six has already been excluded and the interplate CV is too high, then this sample should be repeated and data from this run should be removed from final telomere length results.

6.7.4 For quality control, check the intraplate variation at the bottom of Samples P1 and Samples P2 Sheets, which should be less than 0.05 (5%).

6.7.5 For quality control, check the overall interplate variation in P1 v P2 sheet, cell G29, which should be less than 0.06 (6%). Ensure that samples that did not pass quality control are removed from the interplate CV calculation in cell G30.

6.7.6 Take final TL data for each sample that passed quality control (P1 v P2 Sheet, column I) and place in the separate overall cohort or experiment data file; similarly, take the data for intraclass correlation coefficient (ICC) calculations for each sample that passed quality control (ICC Data Sheet, columns E through K) and place in a separate data file for ICC calculation. Lastly, take final assay run quality control data from P1 v P2 Sheet (columns J through L), and place in a separate quality control data file for the overall cohort or experiment.

6.8 To calculate ICCs for the overall cohort or experiment, follow directions listed on the Telomere Research Network website, Resources for Study Design & Analysis tab, "Calculating repeatability of TL measures using ICC" resource.

6.9. Lastly, to improve comparability between TL studies, transform the final TL data for each sample in the separate overall cohort or experiment data file into Z-scores as follows:

6.9.1 Calculate the mean of all the TLs in the data file; calculate the absolute value of the standard deviation of the TL data file.

6.9.2 Subtract the mean value from each sample TL and place this value into a new column; divide this new value by the standard deviation and place this Z-score into a new column. Utilize these Z-scored TLs for future data analyses.

7. MMqPCR Data Reporting

7.1 When using the TL data created from the MMqPCR assay to answer any research questions, report that the MMqPCR assay was conducted and the following items in alignment with the Telomere Research Network Minimum Reporting Guidelines (Lindrose and Drury, 2020):

7.1.1 State the sample type utilized in the cohort or experiment, as extracted and prepared in Step 2. Include storage conditions (temperature, duration, buffer, etc.) of the samples.

7.1.2 Identify the DNA extraction method used. Include DNA storage conditions, like freeze-thaw cycles. Also include the number of DNA samples assessed for quality and integrity, the average of the DNA quality and integrity measurements, and method of documentation (e.g., spectrophotometer and Qubit assay) as identified in Step 2.1.1., as well as the number of samples not assayed due to poor sample quality.

7.1.3 Identify the standard control sample source and the concentrations of the serial dilutions as defined in Step 2 and Table 5.

7.1.3 Identify the machine used for MMqPCR, here it is the CFX96 Touch Real-Time PCR Detection System (Bio-Rad Laboratories; Hercules, California).

7.1.4 Define the sources of the master mix reagents as outlined in Table 1 and report the final reaction volume of 25µL.

7.1.5 Define the sequences and names of the telomere and singe copy gene primers and the final primer concentrations in the reaction mixture as outlined in Tables 2 and 3. Report the average efficiencies of both amplicons as collected in the separate quality control data file for the overall cohort or experiment.

7.1.6 Report the thermocycling protocol as defined in Step 5.

7.2 Report the methods within the sample MMqPCR template excel for calculating CVs between sample replicates, interplate CVs, intraplate CV, and sample T/S ratios. Report all CVs.

7.2.1 Report the ICC for the overall cohort or experiment as conducted in Step 6.8; identify the level of independence of replicates as by plate, e.g., P1 versus P2. Calculate the average number of replicates per sample from the separate ICC sheet and report this as well.

7.2.2 Report the mean, standard deviation, range, and Z-scores of TLs in the overall cohort or experiment using the separate overall cohort or experiment data file as described in Steps 6.7.6 and 6.9.2.

7.2.3 Lastly, report the number of samples that did not pass quality control and the amount that were repeated on the MMqPCR assay or excluded from final analysis.

7.3 Cite this protocol and other Telomere Research Network guidelines and resources where appropriate.

Representative Results:

The results presented here offer an example of highly repeatable TL measurements obtained by following the above protocol. Venous blood samples from consented adults were collected in Becton, Dickinson and Company Vacutainer Orange Top Rapid Serum Separator tubes and stored at -20 °C for less than twenty-four hours prior to separation into PBMCs and DNA extraction. For these results, PBMCs from thirty-six individuals were separated into two aliquots prior to DNA extraction, creating two DNA samples per participant. Samples were assigned random numbers so that the MMqPCR technician was blinded from sample linkage prior to DNA quantification and quality checks and MMqPCR TL analysis. All DNA samples were checked for quality via spectrophotometer and Qubit assay; the average DNA metrics are presented below [Insert (Table 6) here]. The DNA samples underwent up to three freeze-thaw cycles prior to MMqPCR TL analysis.

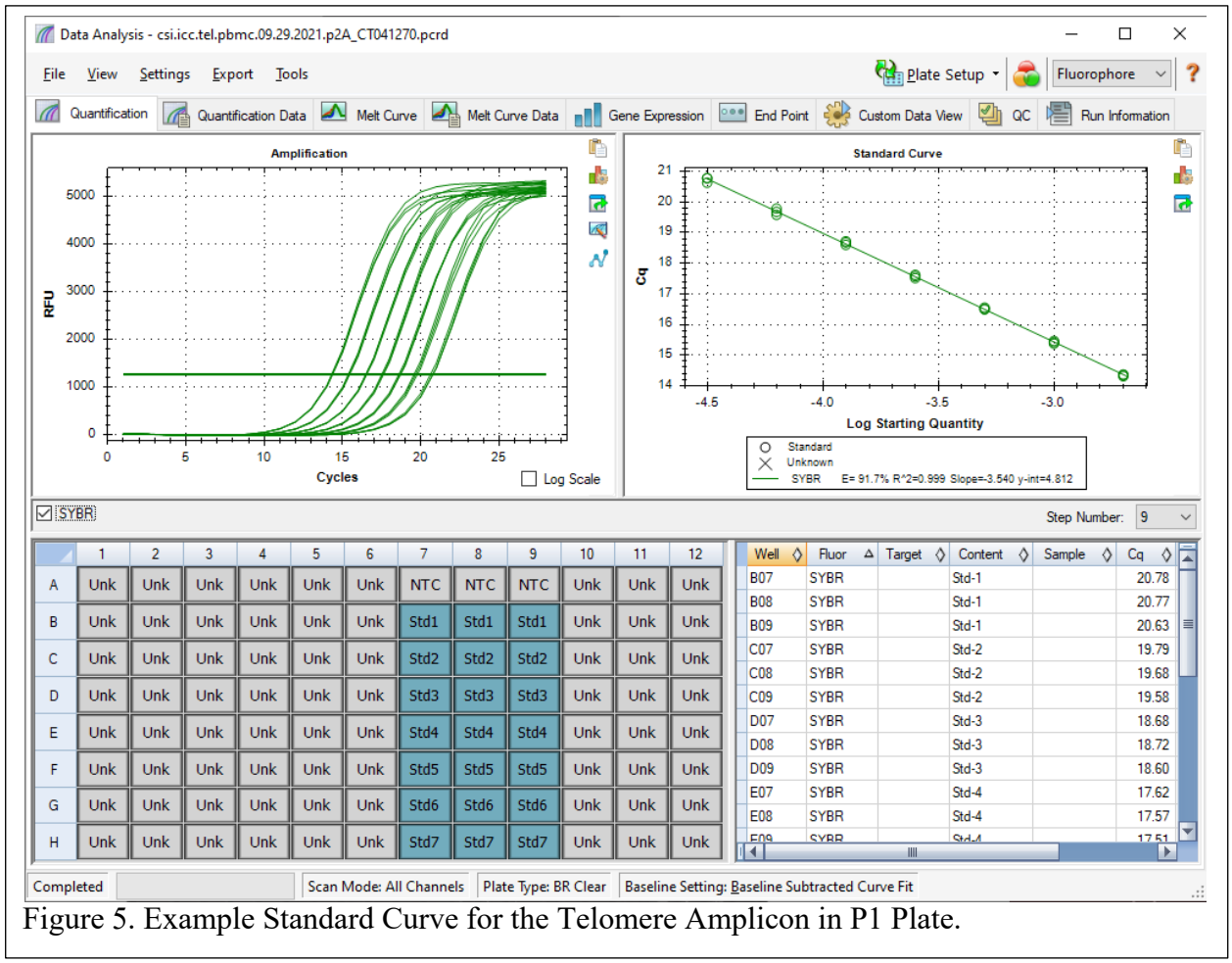
The independent standard curves created for both step 9 (telomere amplicon) and step 12 (single copy gene amplicon, albumin here) are presented below, where the sequences and concentrations are as outlined in in Tables 2 and 3 [insert (Figures 5 and 6) here]. The standard curve was created using a DNA sample from pooled extractions of PBMCs from one individual,

Table 6. Spectrophotometer and Qubit DNA Quality Metrics for Sample Results.

PBMC Sample Number & Actual ID	Blinded ID	Extraction Date	Initial Volume (uL)	Sample Amount (uL)	260 Raw	280 Raw	230 Raw	320 Raw	260	280	230	260/280	260/230	total [DNA]	Qubit ds[DNA]
1	298	7/14/2021		100	0.1595	0.106	0.118	0.044	0.1105	0.059	0.044	1.8775	2.524	110.3	120
CSI-0136-01	378	7/14/2021	650.00	100	0.1825	0.1225	0.134	0.0565	0.1205	0.0635	0.047	1.885	2.542	120.1	140
2	472	7/14/2021		100	0.136	0.093	0.107	0.043	0.0865	0.0475	0.033	1.822	2.6105	86.55	99.3
CSI-0137-01	352	7/14/2021	400.00	100	0.1425	0.096	0.108	0.043	0.096	0.051	0.038	1.88	2.5225	95.7	114
3	499	7/14/2021		100	0.1645	0.1305	0.1395	0.088	0.0765	0.045	0.0255	1.715	2.967	76.6	94.7
CSI-0138-01	305	7/14/2021	600.00	100	0.124	0.086	0.101	0.041	0.0785	0.042	0.03	1.868	2.582	78.35	98
4	72	7/15/2021		100	0.081	0.065	0.0855	0.046	0.0285	0.015	0.0095	1.939	2.985	28.5	34.2
CSI-0124-01	284	7/15/2021	400.00	100	0.0915	0.076	0.095	0.057	0.029	0.016	0.01	1.8	2.854	29.25	34.1
5	432	7/15/2021		100	0.144	0.099	0.115	0.0475	0.091	0.048	0.0375	1.9055	2.4085	90.9	117
CSI-0124-02	340	7/15/2021	400.00	100	0.126	0.089	0.106	0.046	0.074	0.04	0.03	1.867	2.4715	74.4	94.7
6	420	7/15/2021		100	0.105	0.0755	0.0925	0.042	0.058	0.031	0.022	1.8915	2.6885	58.35	70
CSI-0125-01	52	7/15/2021	500.00	100	0.105	0.0755	0.0955	0.042	0.059	0.0315	0.026	1.8788	2.279	58.8	71.3
7	209	7/15/2021		100	0.1825	0.1185	0.127	0.045	0.1315	0.0705	0.0535	1.875	2.462	131.7	161
CSI-0130-01	308	7/15/2021	450.00	100	0.172	0.111	0.121	0.042	0.1255	0.066	0.05	1.884	2.5065	125.2	145
8	331	7/16/2021		100	0.099	0.074	0.094	0.047	0.047	0.025	0.017	1.888	2.7015	47	51.6
CSI-0127-01	56	7/16/2021	300.00	100	0.1175	0.1025	0.1135	0.07	0.043	0.0295	0.013	1.4375	3.339	42.9	54
9	348	7/16/2021		100	0.0955	0.072	0.092	0.046	0.045	0.024	0.016	1.907	2.739	45.2	53.6
CSI-0127-02	181	7/16/2021	450.00	100	0.114	0.0955	0.106	0.062	0.047	0.0305	0.0155	1.556	3.0525	47.3	61.9
10	282	7/16/2021		100	0.1515	0.106	0.118	0.055	0.091	0.0485	0.034	1.8775	2.6515	90.95	115
CSI-0129-01	374	7/16/2021	400.00	100	0.155	0.104	0.12	0.046	0.104	0.055	0.043	1.8845	2.403	104.05	135
11	410	7/16/2021		100	0.1005	0.072	0.089	0.041	0.0555	0.0295	0.021	1.884	2.638	55.4	66.7
CSI-0131-01	435	7/16/2021	300.00	100	0.102	0.074	0.093	0.042	0.0555	0.029	0.022	1.8925	2.5375	55.45	64.9
12	441	7/29/2021		100	0.1055	0.076	0.098	0.042	0.0585	0.0315	0.026	1.8585	2.2315	58.35	48.4
CSI-0005-03	47	7/29/2021	350.00	100	0.1025	0.0735	0.093	0.0415	0.056	0.0295	0.021	1.8905	2.6565	56.05	49
13	437	7/29/2021		100	0.0705	0.063	0.087	0.0565	0.0065	0.004	0.0035	1.6695	1.93	6.7	8.8
CSI-0006-02	152	7/29/2021	100.00	100	0.056	0.05	0.0745	0.042	0.009	0.006	0.0015	1.57	6.362	9.5	8.8
14	14	7/29/2021		100	0.088	0.068	0.0865	0.044	0.044	0.022	0.017	1.9705	2.5905	43.65	35.7
CSI-0010-02	300	7/29/2021	400.00	100	0.105	0.0815	0.1	0.0545	0.048	0.024	0.016	1.9065	2.883	45.75	37.6
15	59	7/29/2021		100	0.1055	0.079	0.0975	0.0485	0.052	0.0275	0.019	1.865	2.7275	51.75	43.1
CSI-0015-03	488	7/29/2021	400.00	100	0.1	0.072	0.089	0.04	0.053	0.029	0.024	1.8015	2.22	53.05	47.1
16	314	8/16/2021		100	0.1155	0.084	0.106	0.047	0.0625	0.033	0.0285	1.875	2.1885	62.25	65.5
CSI-0016-03	310	8/16/2021	300.00	100	0.131	0.095	0.117	0.055	0.07	0.037	0.031	1.87	2.2675	69.95	68.7
17	227	8/16/2021		100	0.0965	0.072	0.0945	0.045	0.047	0.025	0.0195	1.8715	2.3885	46.7	51.2
CSI-0017-03	427	8/16/2021	300.00	100	0.1065	0.079	0.109	0.049	0.0525	0.028	0.0305	1.891	1.733	52.85	53.3
18	67	8/16/2021		100	0.1785	0.115	0.128	0.044	0.1295	0.0685	0.055	1.887	2.366	129.55	150
CSI-0018-03	228	8/16/2021	800.00	100	0.2195	0.141	0.15	0.052	0.1605	0.085	0.066	1.8825	2.4245	160.5	174
19	274	8/16/2021		100	0.1105	0.0775	0.096	0.041	0.0655	0.0345	0.025	1.8925	2.624	65.2	71.3
CSI-0019-02	24	8/16/2021	300.00	100	0.1695	0.111	0.123	0.046	0.119	0.063	0.048	1.8915	2.478	119.05	137
20	458	8/16/2021		100	0.1105	0.082	0.1	0.048	0.0575	0.031	0.023	1.875	2.511	57.75	63
CSI-0021-03	494	8/16/2021	400.00	100	0.1085	0.079	0.098	0.045	0.0585	0.031	0.024	1.8735	2.4715	58.45	64.3
21	187	8/16/2021		100	0.068	0.056	0.08	0.0425	0.021	0.011	0.0085	1.8375	2.4945	20.95	26.8
CSI-0022-03	242	8/16/2021	200.00	100	0.0685	0.057	0.081	0.044	0.0195	0.01	0.0075	1.866	2.6015	19.5	25.8
22	401	8/16/2021		100	0.087	0.066	0.085	0.0405	0.0425	0.0225	0.0155	1.872	2.763	42.4	46.1
CSI-0023-03	547	8/16/2021	400.00	100	0.0905	0.068	0.089	0.042	0.043	0.0235	0.016	1.8175	2.6425	42.8	45.5
23	43	8/16/2021		100	0.0955	0.07	0.0945	0.042	0.0495	0.026	0.023	1.8765	2.1425	49.35	57.4
CSI-0031-02	268	8/16/2021	100.00	100	0.103	0.075	0.094	0.043	0.056	0.029	0.023	1.946	2.4435	55.95	62.5
24	407	8/17/2021		100	0.117	0.084	0.101	0.0445	0.067	0.036	0.023	1.8595	2.869	66.85	74
CSI-0031-03	345	8/17/2021	500.00	100	0.1315	0.095	0.112	0.053	0.0715	0.038	0.0255	1.8765	2.808	71.6	80.7
25	61	8/17/2021		100	0.163	0.108	0.121	0.045	0.1125	0.06	0.043	1.878	2.591	112.2	130
CSI-0032-02	22	8/17/2021	200.00	100	0.1775	0.1165	0.1275	0.047	0.1255	0.067	0.047	1.873	2.691	125.4	142
26	134	8/17/2021		100	0.09	0.068	0.087	0.042	0.044	0.026	0.013	1.726	3.355	44.1	48.5
CSI-0032-03	500	8/17/2021	300.00	100	0.106	0.082	0.099	0.052	0.05	0.026	0.017	1.929	2.933	50.15	53.5
27	95	8/17/2021		100	0.058	0.051	0.076	0.042	0.01	0.006	0.0035	1.5665	2.666	9.95	13
CSI-0035-02	418	8/17/2021	50.00	100	0.058	0.051	0.0745	0.042	0.011	0.007	0.003	1.657	4.1365	11.1	13.6
28	463	8/17/2021		100	0.117	0.0855	0.1055	0.048	0.0625	0.0345	0.029	1.817	2.1295	62.5	69.3
CSI-0039-03	280	8/17/2021	550.00	100	0.12	0.085	0.106	0.043	0.071	0.038	0.0295	1.861	2.381	70.7	79.3
29	125	8/17/2021		100	0.112	0.079	0.096	0.041	0.065	0.036	0.024	1.8165	2.6595	64.75	70
CSI-0049-02	354	8/17/2021	500.00	100	0.119	0.084	0.101	0.044	0.069	0.038	0.025	1.8175	2.781	69.25	75.3
30	124	8/18/2021		100	0.1365	0.094	0.109	0.045	0.085	0.045	0.029	1.8745	2.9565	85	96
CSI-0050-02	490	8/18/2021	450.00	100	0.1465	0.101	0.115	0.05	0.0875	0.047	0.042	1.8695	2.091	87.4	98
31	468	8/18/2021		100	0.117	0.084	0.101	0.046	0.072	0.0275	0.0325	2.623	2.2015	72	65.7
CSI-0107-01	459	8/18/2021	250.00	100	0.125	0.0895	0.1065	0.049	0.072	0.0415	0.033	1.736	2.1825	71.8	79.3
32	40	8/18/2021		100	0.094	0.07	0.089	0.043	0.0455	0.024	0.014	1.9015	3.1565	45.45	50.3
CSI-0108-01	102	8/18/2021	200.00	100	0.1095	0.086	0.1185	0.058	0.0475	0.025	0.029	1.877	1.622	47.1	41.4
33	476	8/18/2021		100	0.1115	0.079	0.093	0.041	0.065	0.0355	0.022	1.824	2.957	64.75	74.7
CSI-0109-01	402	8/18/2021	200.00	100	0.1125	0.079	0.095	0.041	0.066	0.035	0.024	1.8805	2.785	66	72
34	171	8/18/2021		100	0.094	0.072	0.09	0.046	0.042	0.023	0.014	1.816	2.9545	41.95	44.8
CSI-0118-01	471	8/18/2021	100.00	100	0.0945	0.071	0.088	0.043	0.045	0.024	0.015	1.862	3.061	45.15	49.3
35	129	8/18/2021		100	0.0845	0.0645	0.093	0.042	0.038	0.02	0.02	1.8735	1.916	37.75	39.1
CSI-0122-01	192	8/18/2021	200.00	100	0.092	0.07	0.087	0.044	0.044	0.022	0.015	2.0165	2.9135	43.55	45.6
36	13	8/18/2021		100	0.07	0.058	0.0775	0.042	0.0235	0.013	0.0045	1.835	5.2655	23.4	25.9
CSI-0055-03	423	8/18/2021	100.00	100	0.074	0.059	0.08	0.042	0.027	0.015	0.008	1.823	3.393	26.8	28.7

with a seven-point serial dilution from 0.0313ng to 2ng of DNA, and all reagents in the MMqPCR analysis are as described in Table 1 with a final reaction volume of 25µL. The average telomere efficiency was 94.40% and the average albumin efficiency was 100.74%, with an average of 0.75 points removed from the standard curves, which meets the quality control criteria of this protocol. The assays were run on the CFX96 Touch Real-Time PCR Detection System which created this below example melt curve that shows the individual amplicon products being produced at distinct temperatures [insert Figure 7 here]. Identifying two peaks in the melt curve ensures that the PCR was correctly executed in this multiplex assay. The two peaks created in the melt curve at the end of the thermal cycling show individual products, the

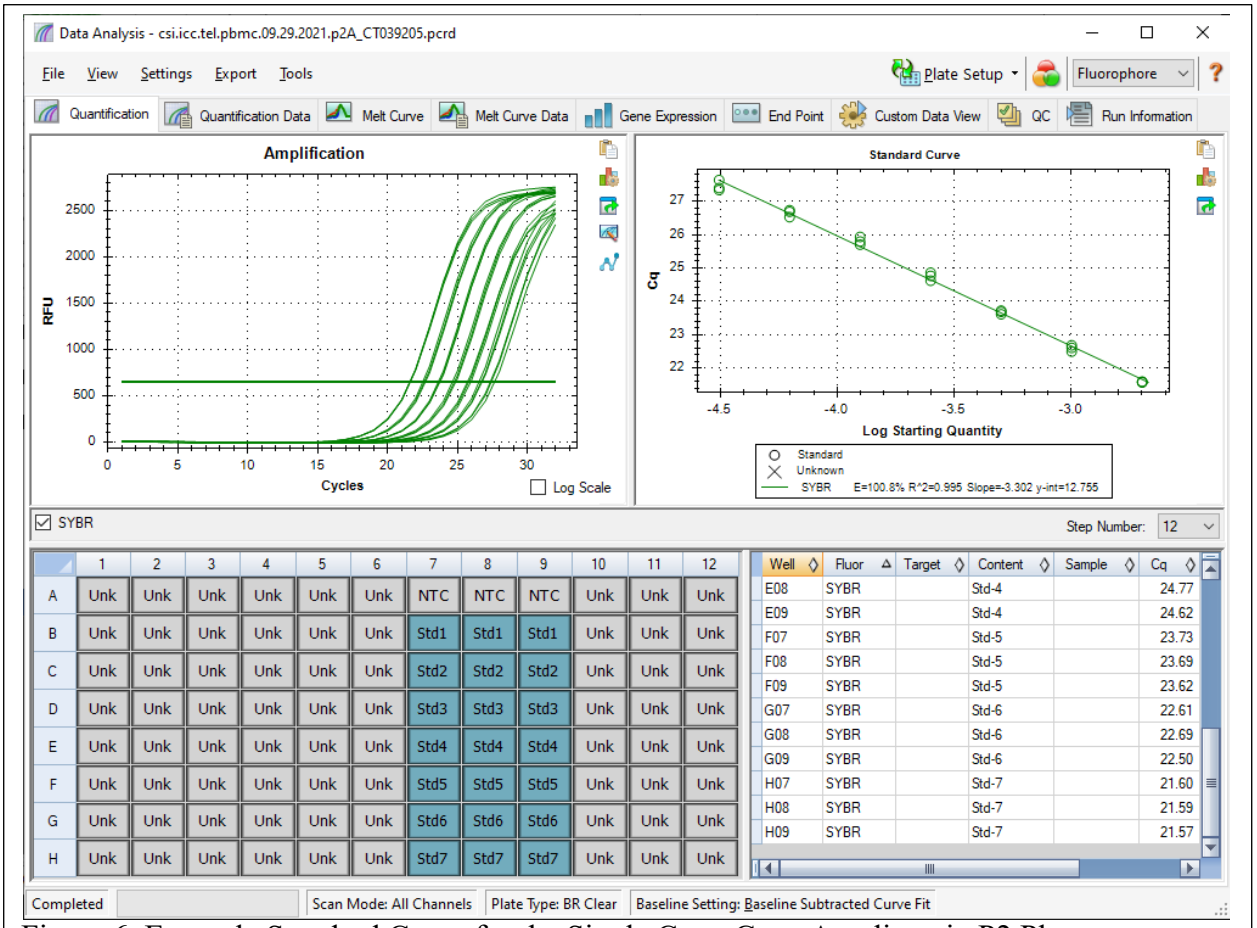
first peak at 85 oC represents the telomere amplicon while the peak at 94 oC represents the albumin amplicon.



	1	2	3	4	5	6	7	8	9	10	11	12
A	Unk	Unk	Unk	Unk	Unk	Unk	NTC	NTC	NTC	Unk	Unk	Unk
B	Unk	Unk	Unk	Unk	Unk	Unk	Std1	Std1	Std1	Unk	Unk	Unk
C	Unk	Unk	Unk	Unk	Unk	Unk	Std2	Std2	Std2	Unk	Unk	Unk
D	Unk	Unk	Unk	Unk	Unk	Unk	Std3	Std3	Std3	Unk	Unk	Unk
E	Unk	Unk	Unk	Unk	Unk	Unk	Std4	Std4	Std4	Unk	Unk	Unk
F	Unk	Unk	Unk	Unk	Unk	Unk	Std5	Std5	Std5	Unk	Unk	Unk
G	Unk	Unk	Unk	Unk	Unk	Unk	Std6	Std6	Std6	Unk	Unk	Unk
H	Unk	Unk	Unk	Unk	Unk	Unk	Std7	Std7	Std7	Unk	Unk	Unk

Well	Fluor	Target	Content	Sample	Cq
B07	SYBR		Std-1		20.78
B08	SYBR		Std-1		20.77
B09	SYBR		Std-1		20.63
C07	SYBR		Std-2		19.79
C08	SYBR		Std-2		19.68
C09	SYBR		Std-2		19.58
D07	SYBR		Std-3		18.68
D08	SYBR		Std-3		18.72
D09	SYBR		Std-3		18.60
E07	SYBR		Std-4		17.62
E08	SYBR		Std-4		17.57

Figure 5. Example Standard Curve for the Telomere Amplicon in P1 Plate.



	1	2	3	4	5	6	7	8	9	10	11	12
A	Unk	Unk	Unk	Unk	Unk	Unk	NTC	NTC	NTC	Unk	Unk	Unk
B	Unk	Unk	Unk	Unk	Unk	Unk	Std1	Std1	Std1	Unk	Unk	Unk
C	Unk	Unk	Unk	Unk	Unk	Unk	Std2	Std2	Std2	Unk	Unk	Unk
D	Unk	Unk	Unk	Unk	Unk	Unk	Std3	Std3	Std3	Unk	Unk	Unk
E	Unk	Unk	Unk	Unk	Unk	Unk	Std4	Std4	Std4	Unk	Unk	Unk
F	Unk	Unk	Unk	Unk	Unk	Unk	Std5	Std5	Std5	Unk	Unk	Unk
G	Unk	Unk	Unk	Unk	Unk	Unk	Std6	Std6	Std6	Unk	Unk	Unk
H	Unk	Unk	Unk	Unk	Unk	Unk	Std7	Std7	Std7	Unk	Unk	Unk

Well	Fluor	Target	Content	Sample	Cq
E08	SYBR		Std-4		24.77
E09	SYBR		Std-4		24.62
F07	SYBR		Std-5		23.73
F08	SYBR		Std-5		23.69
F09	SYBR		Std-5		23.62
G07	SYBR		Std-6		22.61
G08	SYBR		Std-6		22.69
G09	SYBR		Std-6		22.50
H07	SYBR		Std-7		21.60
H08	SYBR		Std-7		21.59
H09	SYBR		Std-7		21.57

Figure 6. Example Standard Curve for the Single Copy Gene Amplicon in P2 Plate.

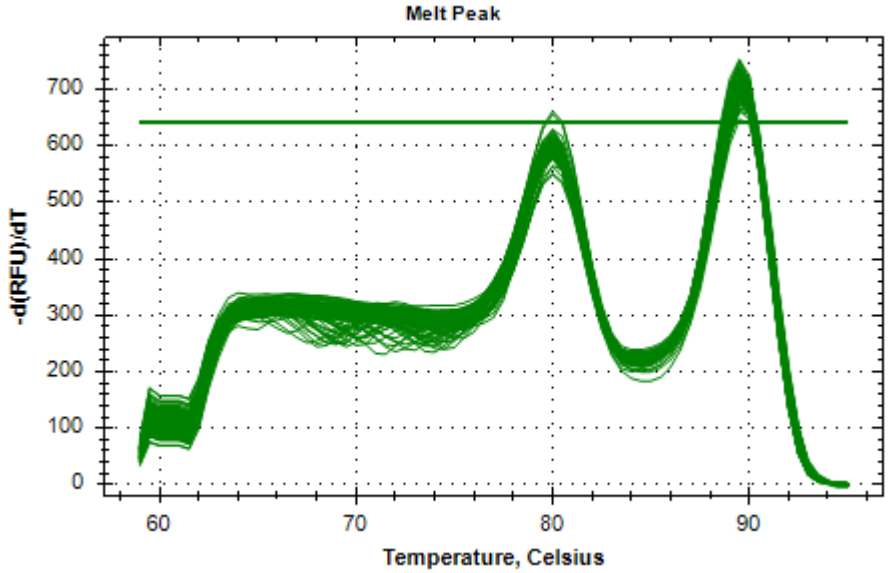
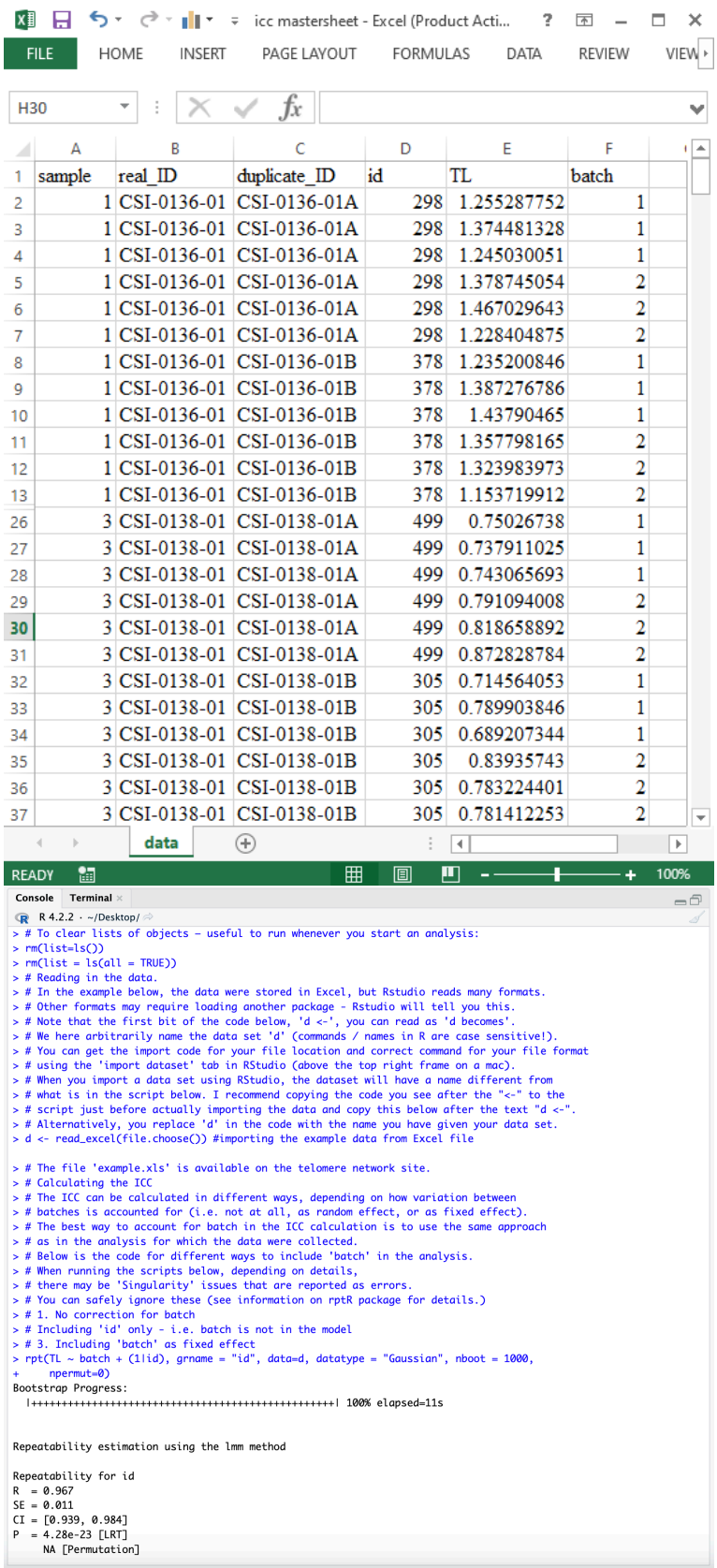


Figure 7. Example Melt Curve with Individual Peaks.

Approximately 9 samples were repeated on the assay with all DNA samples passing quality control, an average interplate CV of 3.22% and an average intraplate CV of 1.98%. The average TLs and ICC for samples from both extractions with a batch effect of individual plates were calculated and presented below (Table 7 and Figures 2 and 8). The mean was 1.05, the standard deviation was 0.24, the range was 0.59 to 1.51, and Z-scores are listed in Table 7 per sample. Average TLs for both DNA extractions were well correlated within individuals and the ICC was high across DNA extractions (Figures 8 and 9). Average TLs across DNA extractions had a significant Pearson correlation coefficient of 0.967 with a p-value less than 0.001. The ICC of 0.967 [CI: 0.939, 0.984] indicates the high reproducibility of TL results from the same biological sample.

Table 7. Mean T/S Ratios and Z-scored Telomere Length for Sample Results.

PBMC Sample Number & Actual ID	Blinded ID	Average T/S Ratio	Batch (plate date)	Z-Scored TL
1	298	1.32482978383103	1	1.171940815110040
CSI-0136-01	378	1.31598072188481	1	1.134377843448000
2	472	1.19269963682172	4	0.611067696809703
CSI-0137-01	352	1.23646086924910	4	0.796827717878023
3	499	0.78563763036647	1	-1.116850890051980
CSI-0138-01	305	0.76627822118413	1	-1.199028746069850
4	72	1.13924497065764	4	0.384160469520845
CSI-0124-01	284	1.08690003038346	4	0.161963866070778
5	432	1.14442236297730	1	0.406137741481225
CSI-0124-02	340	1.29213918524268	1	1.033174012950470
6	420	1.34978015292748	4	1.277851476943220
CSI-0125-01	52	1.27805544289262	4	0.973390589802854
7	209	1.09947106747839	4	0.215326077006699
CSI-0130-01	308	0.98158593560291	4	-0.285079038803779
8	331	0.93788789109457	1	-0.470570836412168
CSI-0127-01	56	0.89813920076806	1	-0.639298204059880
9	348	0.94916464212225	1	-0.422702680231696
CSI-0127-02	181	0.85379596136197	1	-0.827528759494591
10	282	0.86040779160071	4	-0.799462508698608
CSI-0129-01	374	0.89598320526685	4	-0.648450089130568
11	410	1.45910690724685	1	1.741927533745310
CSI-0131-01	435	1.50357087553639	1	1.930670565604650
12	441	0.92497259718564	1	-0.525394367093292
CSI-0005-03	47	0.97919812485863	1	-0.295214945622690
13	437	1.29064754029847	2	1.026842198673500
CSI-0006-02	152	1.22525943340468	2	0.749279263617574
14	14	1.04586231106607	2	-0.012235240813246
CSI-0010-02	300	1.00701895305223	2	-0.177119605415842
15	59	1.30993007647374	2	1.108693740037490
CSI-0015-03	488	1.39837652082493	2	1.484135939782670
16	314	1.16733772798847	4	0.503410109080270
CSI-0016-03	310	1.18406114901023	4	0.574398581502591
17	227	1.20754944350479	2	0.674102950830144
CSI-0017-03	427	1.15900941927628	2	0.468057658611567
18	67	0.64924145835314	2	-1.695832659564100
CSI-0018-03	228	0.58980739491533	2	-1.948121552357610
19	274	0.92309371056511	2	-0.533369965559247
CSI-0019-02	24	0.95582436407984	2	-0.394433136104860
20	458	0.90172686243838	2	-0.624069105826179
CSI-0021-03	494	0.89732143597661	2	-0.642769495805160
21	187	0.86768477913814	4	-0.768572762608618
CSI-0022-03	242	0.73467248235489	4	-1.333190477473140
22	401	1.04986958098321	2	0.004775032959763
CSI-0023-03	547	0.92423227400975	2	-0.528536930524954
23	43	0.78458437454612	2	-1.121321806717840
CSI-0031-02	268	0.76600065718457	2	-1.200206964587590
24	407	0.75755060813164	2	-1.236076184969020
CSI-0031-03	345	0.77186349728122	2	-1.175320067225140
25	61	1.41874867476499	3	1.570612748940010
CSI-0032-02	22	1.38616940881843	3	1.432318537535740
26	134	1.21709895161064	3	0.714639213831400
CSI-0032-03	500	1.24325288802095	3	0.825658842601251
27	95	0.90526273917431	3	-0.609059827052953
CSI-0035-02	418	0.90726607685668	3	-0.600555952060384
28	463	0.89717555985310	3	-0.643388718580043
CSI-0039-03	280	0.91156155434411	3	-0.582322279371025
29	125	0.82094530051955	4	-0.966975002475136
CSI-0049-02	354	0.84466460101840	4	-0.866290046499433
30	124	1.45622843256179	3	1.729708830375170
CSI-0050-02	490	1.45111821145193	3	1.708016690360440
31	468	0.93404241358913	3	-0.486894325098518
CSI-0107-01	459	0.91788296230929	3	-0.555488828582341
32	40	1.13670587672656	3	0.373382387735425
CSI-0108-01	102	1.21651217654587	3	0.712148439638137
33	476	0.83559487194366	3	-0.904789717796475
CSI-0109-01	402	0.82117967037248	3	-0.965980136782308
34	171	1.49415288024327	3	1.890692555228140
CSI-0118-01	471	1.50483793699360	3	1.936049055878590
35	129	0.79855660232104	3	-1.062011746605880
CSI-0122-01	192	0.80931895919997	3	-1.016327118233230
36	13	0.82920308819828	4	-0.931921903462805
CSI-0055-03	423	0.92789797665863	4	-0.512976559769515



	A	B	C	D	E	F
1	sample	real_ID	duplicate_ID	id	TL	batch
2	1	CSI-0136-01	CSI-0136-01A	298	1.255287752	1
3	1	CSI-0136-01	CSI-0136-01A	298	1.374481328	1
4	1	CSI-0136-01	CSI-0136-01A	298	1.245030051	1
5	1	CSI-0136-01	CSI-0136-01A	298	1.378745054	2
6	1	CSI-0136-01	CSI-0136-01A	298	1.467029643	2
7	1	CSI-0136-01	CSI-0136-01A	298	1.228404875	2
8	1	CSI-0136-01	CSI-0136-01B	378	1.235200846	1
9	1	CSI-0136-01	CSI-0136-01B	378	1.387276786	1
10	1	CSI-0136-01	CSI-0136-01B	378	1.43790465	1
11	1	CSI-0136-01	CSI-0136-01B	378	1.357798165	2
12	1	CSI-0136-01	CSI-0136-01B	378	1.323983973	2
13	1	CSI-0136-01	CSI-0136-01B	378	1.153719912	2
26	3	CSI-0138-01	CSI-0138-01A	499	0.75026738	1
27	3	CSI-0138-01	CSI-0138-01A	499	0.737911025	1
28	3	CSI-0138-01	CSI-0138-01A	499	0.743065693	1
29	3	CSI-0138-01	CSI-0138-01A	499	0.791094008	2
30	3	CSI-0138-01	CSI-0138-01A	499	0.818658892	2
31	3	CSI-0138-01	CSI-0138-01A	499	0.872828784	2
32	3	CSI-0138-01	CSI-0138-01B	305	0.714564053	1
33	3	CSI-0138-01	CSI-0138-01B	305	0.789903846	1
34	3	CSI-0138-01	CSI-0138-01B	305	0.689207344	1
35	3	CSI-0138-01	CSI-0138-01B	305	0.83935743	2
36	3	CSI-0138-01	CSI-0138-01B	305	0.783224401	2
37	3	CSI-0138-01	CSI-0138-01B	305	0.781412253	2

Figure 8. Example Intraclass Correlation Coefficient Sheet and Calculation.

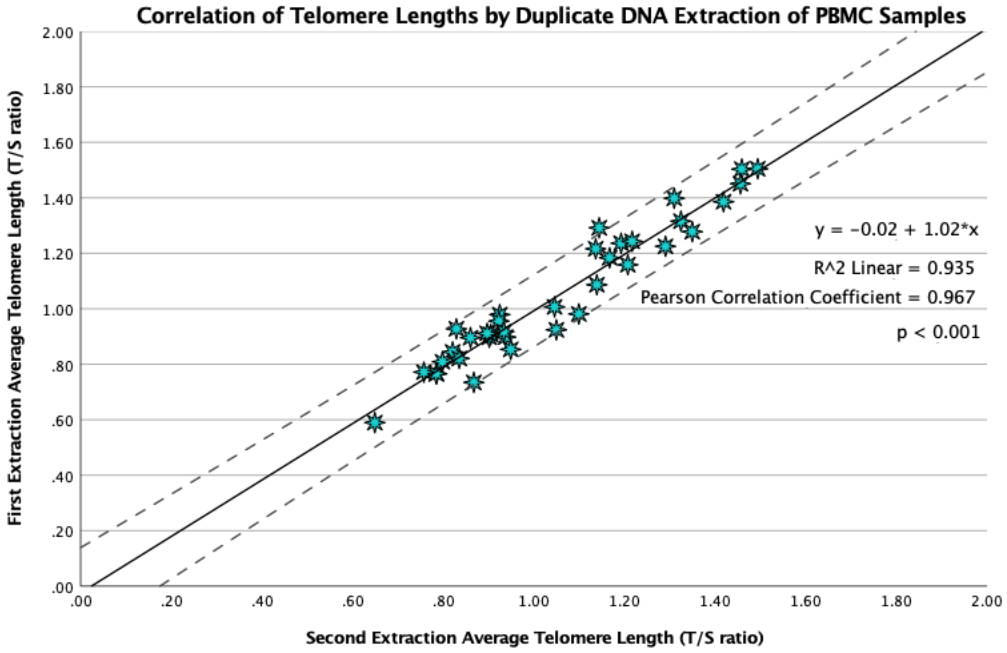


Figure 9. Correlation of Telomere Lengths Across DNA Extractions for Sample Results. Each cyan star represents one sample (n = 36) that was extracted twice.

Discussion:

The MMqPCR TL measurement provides repeatable TL findings, as seen in the results from duplicate DNA extractions for PBMC samples. Furthermore, this method maximizes efficiency with both the telomere and single copy gene amplicons on the same PCR plate, and minimizes costs compared to other methods. However, there are limitations to the method. For example, given that two thermocyclers are utilized in this protocol this method requires large purchases in advance if access to core facilities is not feasible. The throughput of this assay decreases if only one thermocycler is available, similar to the throughput of singleplex PCR, but with enhanced precision and reproducibility with have duplicate triplicates for sample reads. While the minimal amount of DNA needed for the MMqPCR assay is advantageous, the assay is limited to the quality of DNA used for control standards and samples analyzed. Measuring the quality of all DNA samples by purity of nucleic acids with 260/280 ratio between 1.7 to 2.0 and lack of unwanted organic compounds with 260/230 ratio between 2.0 and 2.2. via

spectrophotometer, as well as accurate assessment of double-stranded DNA via Qubit assay are critical steps in this protocol to obtaining repeatable TL data.

Other critical steps in this protocol include the creation of reagents and appropriate storage times. For example, if both the telomere and single copy gene have low efficiencies, the DTT aliquot is a likely cause, and new aliquots of DTT should be made according to this protocol. Conversely, if only one of the telomere or single copy gene amplicons has a low efficiency, consider making a new aliquot of this primer. If this does not increase the efficiencies, then the dNTP aliquot may be the issue and new aliquots of dNTPs should be made from stock dNTPs not older than one calendar year. An important clause of troubleshooting this protocol is to only adjust one reagent at a time to determine the cause of plates not passing the quality control criteria.

A critical quality control step is to ensure that all six no template control wells across plates do not amplify above the threshold fluorescence level. Changing out PCR grade H20 stocks regularly and aliquoting PCR grade H20 sub-stocks for each pair of plates run will minimize sources of contamination and no template control amplification. Practicing sterile technique when in the PCR hood and decontaminating with ultra-violet light after each use of the PCR hood will aid in passing this quality control step as well. Two other critical quality control steps are the coefficient of variation cutoffs for sample triplicates and between plate sample variation. If the coefficient of variations for samples are above the maximum values, vortexing samples, their dilutions, and the PCR strips more vigorously at their respective vortexing steps is recommended when rerunning the samples.

While particular to DNA quality and technician experience, the MMqPCR method is best situated for large population-based studies investigating TL because it requires small amounts of

DNA, is more efficient than singleplex PCR, and is more affordable in reagents and technician time than other methods. For decades, the most popular method of TL measurement was terminal restriction fragment (TRF) analysis using Southern hybridization. However, the MMqPCR assay described here provides repeatable results with less DNA and time than TRF. The MMqPCR method has been utilized in epidemiological studies of 300 or more participants to define the relationship of TL with environment, disease pathology, and genetic biomarkers to date (Maasen et al., 2020; Zöchmeister et al., 2018; Hsieh et al., 2020). These reports and this protocol show the reliability of the MMqPCR method for collecting TL in population-based studies.

References:

Lindrose, A., Drury, S. Minimum Reporting Recommendations for PCR-based Telomere Length Measurement. doi: 10.31219/osf.io/9pzst. Open Science Framework. (2020).

Hsieh, A.Y.Y. et al. Inverse relationship between leukocyte telomere length attrition and blood mitochondrial DNA content loss over time. Aging. 12 (15), 15196–15221, doi: 10.18632/aging.103703 (2020).

Zöchmeister, C. et al. Leukocyte telomere length throughout the continuum of colorectal carcinogenesis. Oncotarget. 9 (17), 13582–13592, doi: 10.18632/oncotarget.24431 (2018).

Maasen, K. et al. Periconceptional environment predicts leukocyte telomere length in a cross-sectional study of 7–9 year old rural Gambian children. Scientific Reports. 10 (1), 9675, doi: 10.1038/s41598-020-66729-9 (2020).

Axelrad, M.D., Budagov, T., Atzmon, G. Telomere Length and Telomerase Activity; A Yin and Yang of Cell Senescence. Journal of Visualized Experiments. (75), 50246, doi: 10.3791/50246-v (2013).

Wilfinger, W.W., Mackey, K., Chomczynski, P. Effect of pH and Ionic Strength on the Spectrophotometric Assessment of Nucleic Acid Purity. BioTechniques. 22 (3), 474–481, doi: 10.2144/97223st01 (1997).

Boesenberg-Smith, K.A., Pessarakli, M.M., Wolk, D.M. Assessment of DNA Yield and Purity: an Overlooked Detail of PCR Troubleshooting. Clinical Microbiology Newsletter. 34 (1), 1–6, doi: 10.1016/j.clinmicnews.2011.12.002 (2012).

Xu, L., Li, S., Stohr, B.A. The Role of Telomere Biology in Cancer. Annual Review of Pathology: Mechanisms of Disease. 8 (1), 49–78, doi: 10.1146/annurev-pathol-020712-164030 (2013).

Mundstock, E. et al. Effect of obesity on telomere length: Systematic review and meta-analysis: Effect of Obesity on Telomeres. Obesity. 23 (11), 2165–2174, doi: 10.1002/oby.21183 (2015).

Haycock, P.C., Heydon, E.E., Kaptoge, S., Butterworth, A.S., Thompson, A., Willeit, P. Leucocyte telomere length and risk of cardiovascular disease: systematic review and meta-analysis. BMJ. 349, g4227, doi: 10.1136/bmj.g4227 (2014).

Wang, Q., Zhan, Y., Pedersen, N.L., Fang, F., Hägg, S. Telomere Length and All-Cause Mortality: A Meta-analysis. Ageing Research Reviews. 48, 11–20, doi: 10.1016/j.arr.2018.09.002 (2018).

Hemann, M.T., Strong, M.A., Hao, L.-Y., Greider, C.W. The Shortest Telomere, Not Average Telomere Length, Is Critical for Cell Viability and Chromosome Stability. Cell. 107 (1), 67–77, doi: 10.1016/S0092-8674(01)00504-9 (2001).

Astuti, Y., Wardhana, A., Watkins, J., Wulaningsih, W. Cigarette smoking and telomere length: A systematic review of 84 studies and meta-analysis. Environmental Research. 158, 480–489, doi: 10.1016/j.envres.2017.06.038 (2017).

Ridout, K.K., Ridout, S.J., Price, L.H., Sen, S., Tyrka, A.R. Depression and telomere length: A meta-analysis. Journal of Affective Disorders. 191, 237–247, doi: 10.1016/j.jad.2015.11.052 (2016).

Hu, R., Hua, X., Jiang, Q. Associations of telomere length in risk and recurrence of prostate cancer: A meta-analysis. Andrologia. 51 (7), doi: 10.1111/and.13304 (2019).

Cawthon, R.M. Telomere length measurement by a novel monochrome multiplex quantitative PCR method. Nucleic Acids Research. 37 (3), e21–e21, doi: 10.1093/nar/gkn1027 (2009).

Supplemental Video. Precision MMqPCR Telomere Length Measurement Methodology. Available via this link: https://tulane.box.com/s/42nhpmfob3chyn9fsgecvuc4xiikh4o7

CHAPTER 4

Validity, Reliability, and Transcultural Adaptations of the Bayley Scales of Infant and Toddler Development (BSID-III-NL) for Children in Suriname

Background:

The Republic of Suriname is a demographically diverse country defined by the World Bank Group as a middle-income country (MIC) located on the northeastern edge of South America. The country was named after the indigenous Surinen tribe [1]. As a consequence of multiple forced movements of slaves and hegiras of immigrants during Dutch colonization of Suriname, the population is highly diverse and includes Hindustani, Tribal (formerly Maroons), Creoles, Javanese, Indigenous (formerly Amerindian), Caucasian, and individuals of mixed descent [2]. Across ethnic groups, Suriname has high perinatal mortality of 40/1000 live and still births, where approximately 1 in 5 pregnancies end in at least one negative birth outcome, and neonatal deaths account for 50% of all perinatal deaths [3,4]. Contributing factors to negative birth outcomes in Suriname include neurotoxicant exposures during pregnancy, especially for Indigenous Surinamese [5].

In an effort to understand the etiology of poor birth outcomes and the impact on neurodevelopmental trajectories of infants and toddlers, a large epidemiologic cohort of pregnant women was initiated by the Caribbean Consortium for Research in Environmental and Occupational Health (CCREOH) [6]. CCREOH is exploring regional differences in known neurotoxicants measured in pregnant women from the capital city of Suriname, Paramaribo, where the majority of the country's population resides, the coastal city of Nickerie, which produces the majority of agricultural exports for Suriname, as well as the colloquially known Interior region that mainly consists of dense tropical rainforest [4, 6, 7, 8, 9]. Findings from the

CCREOH cohort to date report internationally unacceptable levels of neurotoxicants in maternal blood during pregnancy [8]. For example, an average of 10 times the U.S. Environmental Protection Agency's (EPA) action level for methylmercury (Hg), a known neurotoxicant, was measured in blood from pregnant Surinamese women in the Interior region [9].

The German Human-Biomonitoring research studies report that Hg levels of 5 micrograms per liter in blood or greater have negative impacts on health, and a suggested action level for intervention is 3.5 micrograms per liter [10]. Exposure to Hg places the subject at elevated risk to adverse functional outcomes in multiple organ systems across the lifespan, as Hg is a heavy metal that is toxic to neurons at the cellular, molecular, and behavioral levels [11]. Hg exposure in adults is detrimental to overall health, especially the nervous system, through various cellular mechanisms including oxidative stress [12]. Hg exposure to pregnant females affects the offspring's nervous system and behavior in early development and across the lifespan, as Hg and other neurotoxicants cross the placenta and concentrate in the fetus during pregnancy [13-16].

Two well-known studies comparable to the CCREOH cohort in size and longitudinal design that assessed the effects of Hg exposure *in utero* on infant neurodevelopment–the Seychelles Child Development Study and the Faroe Islands Study–have reported contradictory results, in part due to ethnic differences in study populations [13]. The Seychelles Child Development Study, in a primarily African-descent cohort, has not reported negative associations of Hg exposure and child neurodevelopment [17]. However, in the primarily Scandinavian-descent cohort of the Faroe Islands Study, negative associations of Hg exposure and child neurodevelopment have been reported, leading to discussions on the vulnerability to prenatal Hg exposure being enhanced or attenuated by genetic factors or cultural differences [13,18]. In adulthood, exposure to Hg *in utero* has negative associations with cognitive functions

at 22 years of age in the Faroe cohort, but not in the Seychelles cohort of the same age on identical cognitive assessments [17,18] While these island populations were chosen for study on the long-term effects of prenatal Hg exposure due to shared staple foods containing bio-available Hg, the coastal country of Suriname faces extreme exposure to elevated Hg levels from contamination through artisanal, subsistence, small-scale gold mining, especially in areas populated by minority ethnic groups like the Interior region [9].

Pregnant Surinamese women in the Interior region also had mean lead (Pb) levels above the U.S. EPA action level, and while each neurotoxicant works through different or shared molecular mechanisms to affect neuronal development, the combination effects of Hg, Pb, and other neurotoxicants on health outcomes are under-studied [19-21]. Multiple reports show individual *in utero* neurotoxicant exposure is shown to negatively affect longitudinal infant neurodevelopment, but few studies report interactive effects of multiple *in utero* neurotoxicant exposures on infant neurodevelopment [22-24]. Despite the role of prenatal neurotoxicant exposures in negative birth outcomes reported in Suriname, there is a gap in the current literature on the neurodevelopmental trajectories of Surinamese infants.

As a first step in understanding the impact of these neurotoxicant exposures *in utero* on Surinamese infant neurodevelopmental trajectories, I conducted an experiment to evaluate the content validity and psychometric properties of a well-established and internationally validated measure of overall neurodevelopment in infants and toddlers, the Bayley Scales of Infant and Toddler Development. The creation and validation of neurodevelopmental assessments has typically occurred only in high-income counties (HICs) such as the United States and the Netherlands with subsequent utilization of these same assessments in low and middle-income countries (LMICs) [1,25]. The utilization of neurodevelopmental assessments in new countries

and cultures has occurred with varying amounts of consideration for cultural adaptation. One infant assessment measure that has been utilized internationally to assess the neurodevelopmental impact of *in utero* exposures and poor birth outcomes is the Bayley Scales of Infant and Toddler Development, 3rd Edition (BSID-III) [27,28].

The BSID-III has 5 subscales that directly assess language (receptive and expressive), motor (gross and fine), and cognitive function [26,28]. Age-specific start points were established in each subscale to ensure that 95% of participants in that group completed three consecutive items [29]. The BSID-III is individually administered and applicable to infants from 1-month to 42-months of age with existing normative data on 1,700 infants in 17 age groups from the United States [28-30].The adapted Dutch version, the BSID-III-NL, has separate normative data [31].Despite widespread use of the BSID-III in low, middle, and high-income countries, adaptation and validation of the scale has been inconsistent and rarely completed in accordance with the International Test Commission's (ITC) recommendations.

In response to the use of educational and psychological assessments, like the BSID-III, being applied to international samples, the ITC created guidelines for adaptation in new cultural contexts [32].These guidelines outline recommendations for adaptation of assessment to other cultures and include recommendations related to pre-condition (decisions prior to test adaptation), test development, confirmation, administration, score scales and interpretation, and documentation steps for proper adaptation [32]. Insufficient consideration and cultural adaptation of the BSID-III, or inadequate assessment of content validity and reliability, similar to other neurodevelopmental assessments, can result in apparent developmental differences across cultures that are the result of differences in the early environment, rather than true differences in neurodevelopmental trajectories and risk. One example of this was found in a study of Australian

toddlers where the authors reported significant differences in motor development scores of Australian toddlers compared to United States norms. In discussion of how significant differences occurred, the authors note that these differences were most likely the result of cultural differences in the physical context, i.e. the more frequent presence of bouncing apparatuses in Australia, rather than indicative of developmental risk [33]. Beyond initial cultural adaptation of the test content, the ITC also recommends validity and reliability analyses before the implementation of assessments for clinical and intervention purposes [32].

My study assessed the need for cultural adaptation of the BSID-III-NL through consultation with Surinamese pediatric neuropsychologists and pediatricians. In accordance with ITC guidelines, the BSID-III-NL booklets were purchased from the publisher, and Surinamese pediatricians were consulted on the constructs, imagery, culture, and language appropriateness of the BSID-III-NL before administration [26,32]. Following the recommendations of Surinamese medical professionals, only minor adaptations to the booklet imagery were made to ensure culturally appropriateness. Next, construct validity and reliability were analyzed of the adapted BSID-III-NL in a representative sample of Surinamese infants utilizing the BSID-III administration manual.

Building upon the validation of the BSID-III-NL to measure neurodevelopment of Surinamese infants, I further aim to leverage this neurodevelopmental measure and investigate the overall cumulative effects of varying levels of Hg, Pb, aluminum (Al), manganese (Mn), selenium (Se), cadmium (Cd), and tin (Sn) on Surinamese neurodevelopmental trajectories. Recent advances in statistics will allow a more detailed analysis of the CCREOH data on multiple *in utero* exposures on infant neurodevelopment than is reported in current literature [24,34]. In addition, the effects of cumulative prenatal exposure to multiple neurotoxicants are

known to negatively affect the biological development of infants, as indexed by shortening of telomeres [35].

Telomeres are the protective nucleic acid and protein cap at the end of all eukaryotic chromosomes, and telomere length (TL) is a biomarker of aging, associated with age-related disease risk where TL decreases across the lifespan and critically short TL initiates cellular death [36]. Studies have reported that greater cumulative prenatal neurotoxicant exposures are associated with shorter TL of newborns [37,38]. However, reports from the current literature are inconsistent on the direction of relationships between prenatal neurotoxicant exposures, infant neurodevelopment, and TL [18,39]. The proposed study addresses a gap in the current literature specific to Suriname by leveraging the existing CCREOH cohort database, which contains prenatal neurotoxicant concentrations of Hg, Pb, Al, Mn, Se, Cd, and Sn, and neurodevelopmental performances of Surinamese infants. The proposed research additionally contributes to the CCREOH cohort database through measurement of infant TL and investigation of the long-term effects of *in utero* neurotoxicant exposures on infant neurodevelopment and biological aging.

Methodology:

Caribbean Consortium for Research in Environmental and Occupational Health

As part of the CCREOH/Meki Tamara research program, pregnant women and their offspring were recruited from three geographic locations in Suriname as previously described, including Paramaribo, Nickerie, and the Interior region of Suriname [4]. The overall goal of this research is to assess psychosocial and environmental exposures in pregnant Surinamese women, and their relation with perinatal and neurodevelopmental outcomes in offspring. Mothers between the ages of 16 and 45 years-old were invited to participate in the study at prenatal

clinics and midwife facilities. Recruitment for this study began in December 2016 and completed in July 2019. Following informed consent, demographic information on parity, maternal age, education, income, ethnicity, marital status, prenatal occupational exposures, and dietary choices were obtained at multiple time points during pregnancy after mothers were consented. A trained research assistant conducted study interviews and questionnaires in Dutch. All participants were singleton births and brought to one of seven centers between the ages of 10 months and 30 months of age for neurodevelopmental assessment.

Participants

The raw BSID-III-NL scores from a subset of 299 participants from the larger CCREOH study were utilized for analyses of content validity, construct validity, convergent validity, and internal subscale reliability. Participants for this analysis were collected between May 2018 and July 2019 with 57.90% recruited from Paramaribo, 26.80% recruited from Nickerie, and 15.40% recruited from the Interior region. Demographic and regional characteristics of this subset do not significantly differ from the larger cohort [4]. The 299 participants of this study were evaluated

Table 1. **Demographics of Study Participants**. Percentages of females, males, participants in Paramaribo, Nickerie, or the Interior, and participant number are reported by age group with number and percentages displayed for all 299 participants. The lower age limit and upper age limit in months and days for each age group are also displayed.

	Age Group 1:	**Age Group 2:**	**Age Group 3:**	**Age Group 4:**	**Age Group 5:**	**Age Group 6:**	**Age Group 7:**	**Age Group 8:**	**Age Group 9:**	**Overall:**
Sex of Participant										
Female:	63.20%	57.90%	58.50%	37.10%	35.30%	54.30%	48.80%	38.60%	30.00%	140 (46.80%)
Male:	36.80%	42.10%	41.50%	62.90%	64.70%	45.70%	51.20%	61.40%	70.00%	159 (53.20%)
Location of Participant										
Paramaribo:	57.90%	57.90%	78.00%	65.70%	100.00%	71.70%	41.50%	25.00%	5.00%	173 (57.90%)
Nickerie:	31.60%	31.60%	22.00%	22.90%	0.00%	19.60%	41.50%	34.10%	50.00%	80 (26.80%)
Interior:	10.50%	10.50%	0.00%	11.40%	0.00%	8.70%	17.10%	40.90%	45.00%	46 (15.40%)
Number of Participants										
	19	19	41	35	34	46	41	44	20	n = 299
Group Age Limits										
Lower Age:	10 months, 16 days	13 months, 16 days	14 months, 16 days	15 months, 16 days	16 months, 0 days	16 months, 16 days	17 months, 16 days	19 months, 16 days	22 months, 16 days	
Upper Age:	13 months, 15 days	14 months, 15 days	15 months, 15 days	15 months, 31 days	16 months, 15 days	17 months, 15 days	19 months, 15 days	22 months, 15 days	26 months, 15 days	

once between the ages of 10 months and 16 days and 26 months and 15 days. Eight participants included in this subset were born premature, specifically one participant in Age Group 3, one participant in Age Group 5, one participant in Age Group 6, four participants in Age Group 8, and one participant in Age Group 9 (Table 1).

Materials and Setting

The BSID-III was translated and adapted for Dutch speaking infants in the Netherlands [14]. The BSID-III was translated into Dutch, adapted for Dutch culture, assessed for content validity, item sequence suitability, and the appropriateness of age group start points before a normed sample of 1,912 participants was created for the Netherlands [17-19]. Following this multi-stage validation process, the BSID-III-NL is considered an accurate assessment of neurodevelopment for Dutch infants. The BSID-III-NL was utilized in Suriname given that Dutch is the national language in Suriname and the historical ties between the two countries. All ten of the medical personnel and research assistants who conducted the BSID-III-NL in Suriname were trained and received feedback on administration of the BSID-III-NL by a single licensed clinical psychologist (MM). The subscales included in this study were the receptive communication which had 49 items, expressive communication which had 46 items, gross motor which had 72 items, fine motor which had 66 items, and the cognitive subscales which had 91 items. The administration of each subscale began at a specific starting point based on the age of the participant in accordance with the BSID-III manual as outlined in Supplementary Table 1. The participant was required to complete successfully the first three consecutive items from the respective start point. If the participant failed to complete the first three consecutive items, the administrator went back to a previous start point to begin task administration until the infant completed three consecutive items that initiated from a specified start point. Full credit was

given for items prior to the start point for each participant. The administration of the BSID-III-NL then continued until the participant had five consecutive incorrect responses. For this study, results from the BSID-III-NL administered between May 2018 and July 2019 in Paramaribo, Nickerie, and the Interior region of Suriname were utilized.

Statistical Analyses

A total of 96,876 correct or incorrect responses were entered into a database from 324 items for each of the 299 participants. Responses from participants were split by participant's age into 9 different age groups determined by Normative Tables available in the BSID-III Administration Manual [20]. The age groups started at 10 months and 16 days of age and went through 26 months and 15 days. The 9 age groups encompassed somewhere between 15 days to 4 months as outlined in Table 1. Descriptive analytic statistics were calculated and carried out using IBM SPSS version 26.0 and SAS version 9.4 [21,22]. K-Modes cluster analysis was conducted using R [23]. A p-value of 0.010 was used to determine statistical significance.

Construct Validity

Construct validity was assessed by convergence of age of participant and the BSID-III-NL score. This convergent validity was assessed by comparison of mean raw scores across age groups. The raw score for each subscale in each age group was compared by repeated pairwise analysis of variance with Bonferroni post-hoc corrections for multiple pairwise comparisons. Further construct validity was additionally conducted by two statistical methods. First, pairwise tetrachoric correlation coefficients indicating the strength and direction of associations between each of the 324 items with themselves were computed to create the correlation matrix for all participant responses. The correlation matrix was planned to be used for an exploratory factor analysis of the scale construct validity. Second, cluster analysis using K-Modes clustering was

conducted by age group for each subscale to maximize the similarly within clusters and maximize the dissimilarly between clusters. Extraction of number of clusters was determined by the percent of variance explained, with a minimum amount of 80% of variance explained. Two clusters of responses were produced for each age group across all subscales, where the first cluster of responses at the beginning of the subscale are correct responses to items in the subscale for each age group. Inversely, the second cluster of responses at the end of the subscale are incorrect responses to items in the subscale for each age group.

Reliability

The reliability of each subscale was estimated through calculation of Cronbach's alpha coefficient for internal consistency of items assessing the expected neurodevelopmental domain separately in each subscale. The domains of receptive communication, expressive communication, gross motor, fine motor, and cognition were assessed across all age groups.

Ethics Statement

The study was conducted according to the principles of the Declaration of Helsinki and the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use, Good Clinical Practice guidelines (ICH-GCP). The study protocol was reviewed and approved by the Central Committee on Research Involving Human Subjects, the Medical Ethical Committee of the Ministry of Health in Suriname (VG 023-14), and the Institutional Review Board of Tulane University's School of Public Health and Tropical Medicine, New Orleans, LA, USA.

Results:

Demographics of Participants

Table 1 outlines demographic information of the subset of CCREOH / Meki Tamara participants in the study (n = 299). Participants were grouped by age into 9 Age Groups based on the normative aged data in the BSID-III manual. Participants were predominantly from Paramaribo and approximately equivalent proportions of males and females. Demographics of the 299 participants used in this validation study did not differ from the larger CCREOH study cohort by percentage of males and females or region of participants.

Modifications of the BSID-III-NL for Validity in Suriname

The BSID-III-NL utilizes various toys or image materials. Two types of items in the BSID-III-NL kit were altered for administration in Suriname following multiple focus group discussions with Surinamese BSID-III-NL administrators, pediatricians, and neuropsychologists about the different environmental and cultural exposures for Surinamese infants compared to the Netherlands. The content validity was assessed through discussions with assessment administrators and pediatric neuropsychologists in Suriname.

Picture Book

Items in the picture book were changed to be more culturally relevant to Suriname. Specifically, images of dogs were replaced with images of birds, images of cats were replaced with images of monkeys, images of Western washing machines were replaced with images of Surinamese people washing clothes in a natural water source, and similar images of Western objects and actions were replaced with images of Surinamese settings depicting Surinamese people as seen in Supplementary Table 2.

Puzzles

Two-piece symmetrical puzzles of a dog and an ice cream cone were changed to images of a monkey and a boat, respectively (Supplementary Table 2). Dogs and ice cream cones are uncommon objects in the Interior region of Suriname.

Construct Validity: Exploratory Factor Analysis

Pairwise tetrachoric correlation coefficients were created from the binary results of items in the BSID-III-NL to create the correlation matrix (Supplementary Table 3). As credit was given to participants for all items before their start points in each subscale and participants were administered items until five consecutive incorrect responses occurred, pairwise correlations were unable to be computed for 28.26% of the expressive communication subscale, 38.78% of the receptive communication subscale, 53.03% of the fine motor subscale, 40.28% of the gross motor subscale, and 24.18% of the cognitive subscale because at least one of the pairs had more than two missing values. Given the percentage of missing correlations, exploratory factor analysis could not be utilized for construct validation. K-Modes cluster analysis was then used to assess the factor structure of the BSID-III-NL raw scores for infants in Suriname.

Receptive Communication Construct Validity

Convergent Validity

The maximum potential raw score in the receptive communication subscale was 49. In this sample, the mean raw score for the receptive communication subscale was 17.95 ± 3.94. When comparing mean raw scores by age group, the following significant differences were observed: Age Group 1 was significantly different than Groups 6, 7, 8, and 9; Age Group 2 was different than Groups 8 and 9; Age Groups 3, 4, 5, 6, 7, and 8 were different than Age Group 9 as indicated by blue significance bars (Figure 1a). All significant differences were present at a p-value of 0.010.

Cluster Analysis

Cluster analysis grouped responses to items into two specific clusters, one cluster for item numbers in the first cluster of correct responses in age group color, and item numbers in the second cluster for incorrect responses in gray (Figure 1b). These two clusters explained a minimum of 85.80% to a maximum of 90.90% of all variability of responses across the age groups in the receptive communication subscale. The last item grouped into the correct responses cluster for all age groups was number 14, and the last item grouped into the correct responses cluster for any age group was number 25. Descriptions of items 14 to 26 are available (Supplementary Table 4).

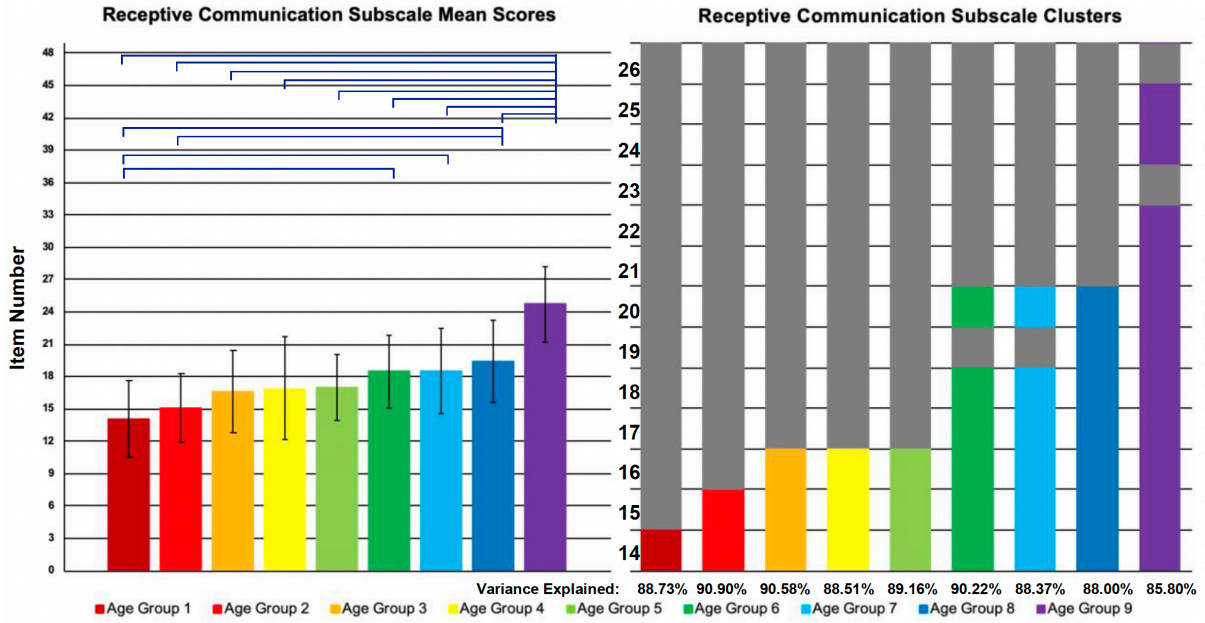


Figure 1. Mean raw score and item clusters of the receptive communication subscale increase with age. Age Group 1 is represented in dark red, Age Group 2 is represented in light red, Age Group 3 is represented in orange, Age Group 4 is represented in yellow, Age Group 5 is represented in light green, Age Group 6 is represented in dark green, Age Group 7 is represented in light blue, Age Group 8 is represented in dark blue, and Age Group 9 is represented in purple. (Left) Mean raw scores in the receptive communication subscale with one standard deviation black bar above and below the observed mean for each age group by item number with significant differences represented in blue bars for between age group comparisons. (Right) Receptive communication subscale item numbers in the first cluster of correct responses in age group color and item numbers in the second cluster of incorrect responses with the overall response variance explained below by age group.

Expressive Communication Construct Validity

Convergent Validity

The maximum potential raw score in the expressive communication subscale was 46. In this sample, the mean raw score for the expressive communication subscale was 18.96 ± 4.67. When comparing mean raw scores by age group, the following significant differences were observed: Age Group 1 was significantly different than Groups 5, 6, 7, 8, and 9; Age Group 2 was different than Groups 6, 7, 8, and 9; Age Group 3 was different than Groups 6, 8, and 9; Age Groups 4, 5, 6, 7, and 8 were different than Age Group 9 as indicated by blue significance bars (Figure 2a). All significant differences were present at a p-value of 0.010.

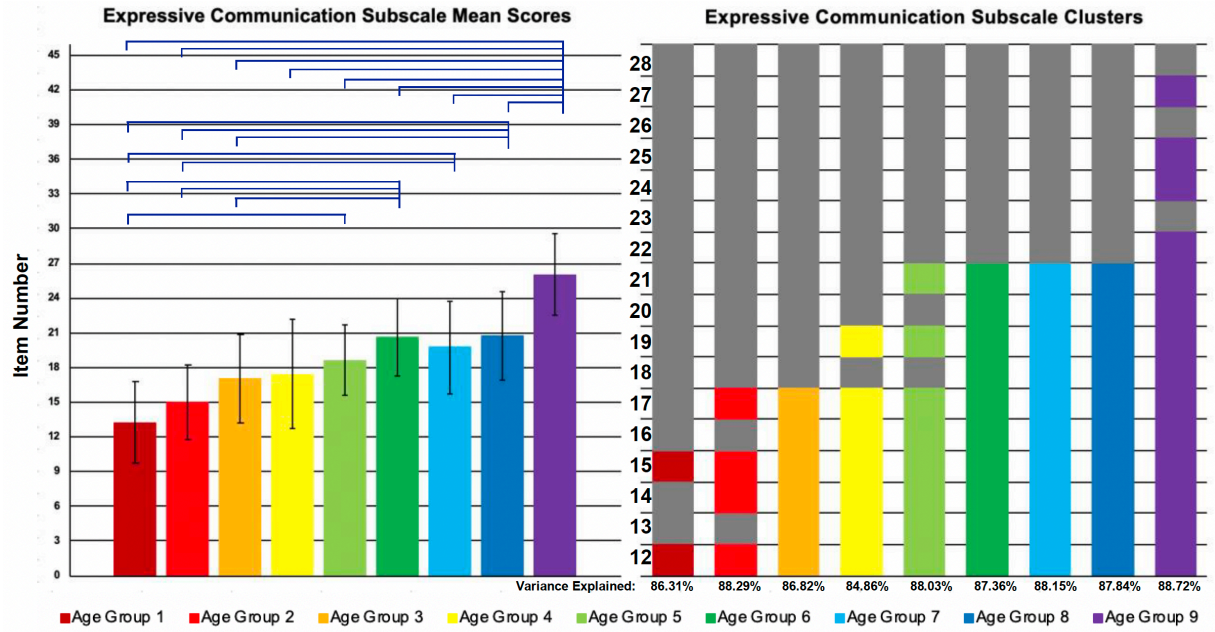


Figure 2. Mean raw score and item clusters of the expressive communication subscale increase with age. Age Group 1 is represented in dark red, Age Group 2 is represented in light red, Age Group 3 is represented in orange, Age Group 4 is represented in yellow, Age Group 5 is represented in light green, Age Group 6 is represented in dark green, Age Group 7 is represented in light blue, Age Group 8 is represented in dark blue, and Age Group 9 is represented in purple. (Left) Mean raw scores in the expressive communication subscale with one standard deviation black bar above and below the observed mean for each age group by item number with significant differences represented in blue bars for between age group comparisons. (Right) Expressive communication subscale item numbers in the first cluster of correct responses in age group color and item numbers in the second cluster of incorrect responses with the overall response variance explained below by age group.

Cluster Analysis

Cluster analysis grouped responses to items into two specific clusters, one cluster for item numbers in the first cluster of correct responses in age group color, and item numbers in the second cluster for incorrect responses in gray (Figure 2b). These two clusters explained a minimum of 84.86% to a maximum of 88.72% of all variability of responses across the age groups in the expressive communication subscale. The last item grouped into the correct responses cluster for all age groups was number 12, and the last item grouped into the correct responses cluster for any age group was number 27. Descriptions of items 12 to 28 are available (Supplementary Table 4). Item 22, which requires participants, from a picture book, to name one image of a cookie, a bottle, shoes, a car, a bird, a balloon, a bed, a kitten, a ball, a spoon, an apple, or a book from a picture book was clustered with incorrect responses for Age Groups 5, 6, 7, and 8, although the previous items were clustered with correct responses for these same age groups.

Gross Motor Construct Validity

Convergent Validity

The maximum potential raw score in the gross motor subscale was 72. In this sample, the mean raw score for the gross motor subscale was 49.33 ± 5.04. When comparing mean raw scores by age group, the following significant differences were observed: Age Group 1 was significantly different than Groups 3, 4, 5, 6, 7, 8, and 9; Age Group 2 was different than Groups 6, 8, and 9; Age Group 3 was different than Groups 8 and 9; Age Groups 4, 5, 6, and 7 were different than Age Group 9 as indicated by blue significance bars (Figure 3a). All significant differences were present at a p-value of 0.010.

Cluster Analysis

Cluster analysis grouped responses to items into two specific clusters, one cluster for item numbers in the first cluster of correct responses in age group color, and item numbers in the second cluster for incorrect responses in gray (Figure 3b). These two clusters explained a minimum of 85.61% to a maximum of 90.02% of all variability of responses across the age

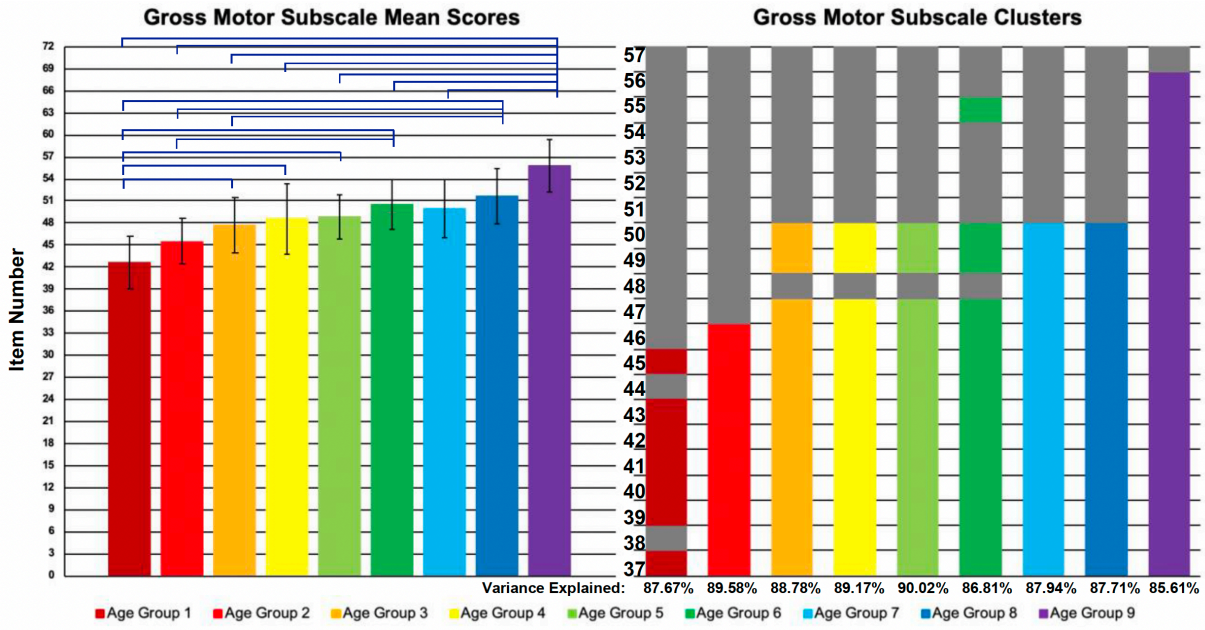


Figure 3. Mean raw score and item clusters of the gross motor subscale increase with age. Age Group 1 is represented in dark red, Age Group 2 is represented in light red, Age Group 3 is represented in orange, Age Group 4 is represented in yellow, Age Group 5 is represented in light green, Age Group 6 is represented in dark green, Age Group 7 is represented in light blue, Age Group 8 is represented in dark blue, and Age Group 9 is represented in purple. (Left) Mean raw scores in the gross motor subscale with one standard deviation black bar above and below the observed mean for each age group by item number with significant differences represented in blue bars for between age group comparisons. (Right) Gross motor subscale item numbers in the first cluster of correct responses in age group color and item numbers in the second cluster of incorrect responses with the overall response variance explained below by age group.

groups in the gross motor subscale. The last item grouped into the correct responses cluster for all age groups was number 37, and the last item grouped into the correct responses cluster for any age group was number 56. Descriptions of items 37 to 57 are available (Supplementary Table 4).

Fine Motor Construct Validity

Convergent Validity

The maximum potential raw score in the fine motor subscale was 66. In this sample, the mean raw score for the fine motor subscale was 33.77 ± 3.22. When comparing mean raw scores by age group, the following significant differences were observed: Age Group 1 was significantly different than Groups 3, 4, 5, 6, 7, 8, and 9; Age Group 2 was different than Groups 8 and 9; Age Group 3 was different than Groups 7, 8, and 9; Age Groups 4 and 5 were different than Groups 8 and 9; Age Groups 6 and 7 were different than Age Group 9 as indicated by blue significance bars (Figure 4a). All significant differences were present at a p-value of 0.010.

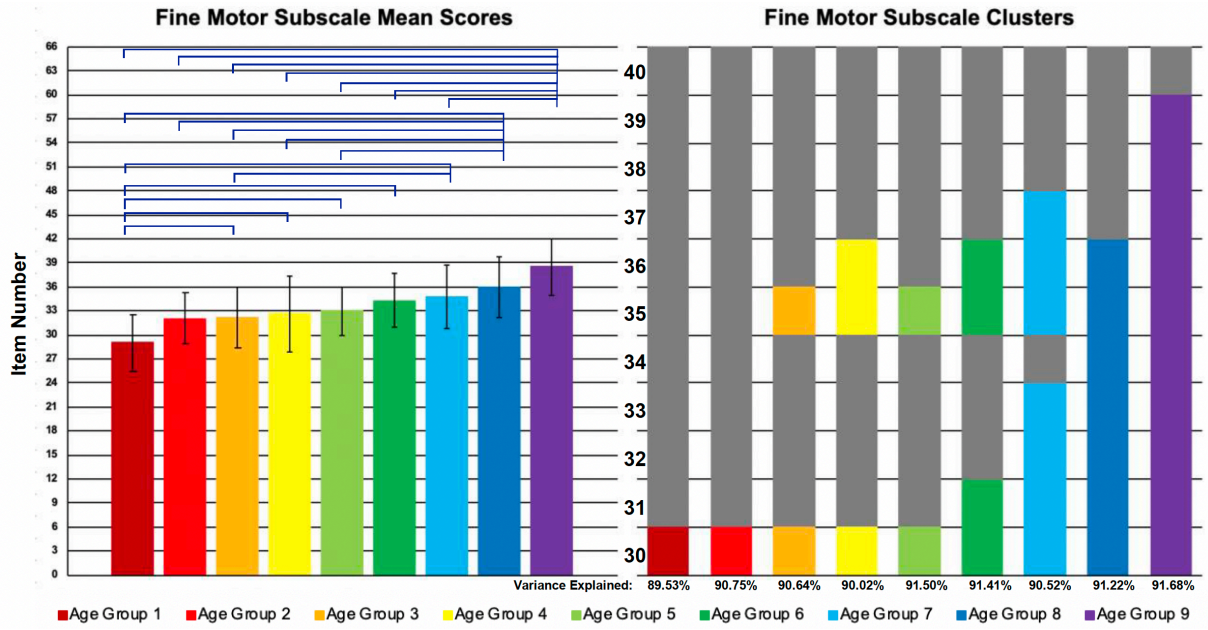


Figure 4. Mean raw score and item clusters of the fine motor subscale increase with age. Age Group 1 is represented in dark red, Age Group 2 is represented in light red, Age Group 3 is represented in orange, Age Group 4 is represented in yellow, Age Group 5 is represented in light green, Age Group 6 is represented in dark green, Age Group 7 is represented in light blue, Age Group 8 is represented in dark blue, and Age Group 9 is represented in purple. (Left) Mean raw scores in the fine motor subscale with one standard deviation black bar above and below the observed mean for each age group by item number with significant differences represented in blue bars for between age group comparisons. (Right) Fine motor subscale item numbers in the first cluster of correct responses in age group color and item numbers in the second cluster of incorrect responses with the overall response variance explained below by age group.

Cluster Analysis

Cluster analysis grouped responses to items into two specific clusters, one cluster for item numbers in the first cluster of correct responses in age group color, and item numbers in the second cluster for incorrect responses in gray (Figure 4b). These two clusters explained a minimum of 89.53% to a maximum of 91.68% of all variability of responses across the age groups in the fine motor subscale. The last item grouped into the correct responses cluster for all age groups was number 30, and the last item grouped into the correct responses cluster for any age group was number 39. Descriptions of items 30 to 40 are available (Supplementary Table 4). Item 34, which requires participants to imitate a written line on paper with a pencil by holding the pencil in a partial tripod or quadrapod thumb position was clustered with incorrect responses for Age Groups 3, 4, 5, 6, and 7. However, the items following 34, such as item 35, which requires participants to place three coins into a piggybank slot, were clustered with correct responses for these age groups.

Cognitive Construct Validity

Convergent Validity

The maximum potential raw score in the cognitive subscale was 91. In this sample, the mean raw score for the cognitive subscale was 52.88 ± 7.05. When comparing mean raw scores by age group, the following significant differences were observed: Age Group 1 was significantly different than Groups 4, 5, 6, 7, 8, and 9; Age Group 2 was different than Groups 7, 8, and 9; Age Group 3 was different than Groups 6, 7, 8, and 9; Age Groups 4 and 5 were different than Groups 8 and 9; Age Groups 6, 7, and 8 were different than Age Group 9 as indicated by blue significance bars (Figure 5a). All significant differences were present at a p-value of 0.010.

Cluster Analysis

Cluster analysis grouped responses to items into two specific clusters, one cluster for item

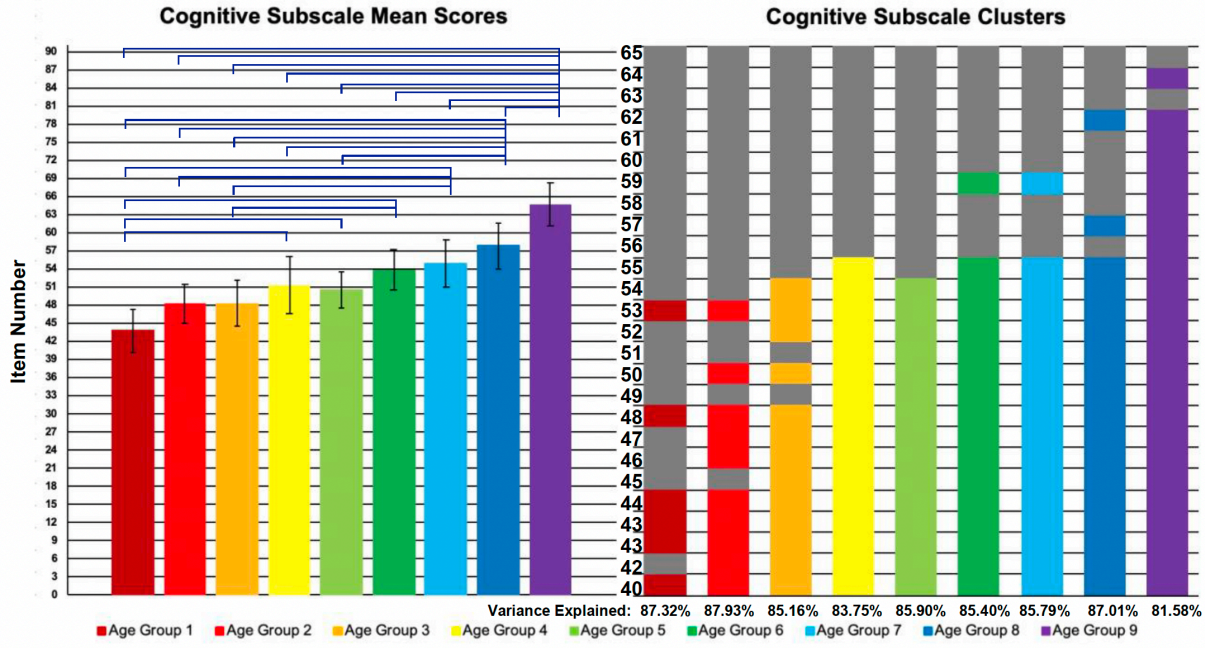


Figure 5. Mean raw score and item clusters of the cognitive subscale increase with age. Age Group 1 is represented in dark red, Age Group 2 is represented in light red, Age Group 3 is represented in orange, Age Group 4 is represented in yellow, Age Group 5 is represented in light green, Age Group 6 is represented in dark green, Age Group 7 is represented in light blue, Age Group 8 is represented in dark blue, and Age Group 9 is represented in purple. (Left) Mean raw scores in the cognitive subscale with one standard deviation black bar above and below the observed mean for each age group by item number with significant differences represented in blue bars for between age group comparisons. (Right) Cognitive subscale item numbers in the first cluster of correct responses in age group color and item numbers in the second cluster of incorrect responses with the overall response variance explained below by age group.

numbers in the first cluster of correct responses in age group color, and item numbers in the second cluster for incorrect responses in gray (Figure 5b). These two clusters explained a minimum of 81.58% to a maximum of 87.93% of all variability of responses across the age groups in the cognitive subscale. The last item grouped into the correct responses cluster for all age groups was number 40, and the last item grouped into the correct responses cluster for any age group was number 64. Descriptions of items 40 to 65 are available (Supplementary Table 4). Item 48, which requires participants to initiate relational play between objects and her- or himself, and item 53, which requires participants to initiate relational play while using objects as

they are intended to be used with others, were clustered with correct responses for all 9 Age Groups.

Reliability

For each of the Bayley subscales, the Cronbach's alpha was used to measure the internal consistency, that is, how closely related a set of items are as a group and how every item reliably measures the same construct [24]. Table 2 reports the internal consistency for each subscale across all 299 participants. The higher the Cronbach's alpha the greater the reliability of the subscales in testing the desired construct. The Cronbach's alpha for the subscales ranged between 0.77 and 0.91 with all scales indicating a greater internal consistency than the standard cut off of reliability set at 0.70 [25].

Table 1. Reliability of Each Subscale.

Cronbach's Alpha Coefficient	**Receptive Communication Subscale**	**Expressive Communication Subscale**	**Gross Motor Subscale**	**Fine Motor Subscale**	**Cognitive Subscale**
Questions Relating to Subscale	0.8590	0.8870	0.8870	0.7750	0.9060

Discussion:

The present study demonstrated the construct validity and reliability of the BSID-III-NL for Surinamese infants. The BSID-III-NL was created for the ethnically homogenous population of the Netherlands. However, although cultural and historical ties between the Netherlands and Suriname exist, early development in Suriname is characterized by different childcare practices and differential exposures to games, books, and toys compared to the Netherlands requiring adaptations to prevent cultural differences impacting scores of Surinamese infants on the BSID-III-NL. Before adaptation, the extent to which the content of items in the scale were appropriate for Surinamese infants was assessed by Surinamese pediatricians and neuropsychologists. Visual adaptations to the picture book and puzzles were similar to those reported in other studies with

respect to their study populations' cultural exposures [26,27]. Consideration of cultural differences in the early environment and evaluation of the relevance of photos and pictures in the BSID-III-NL prior to the implementation of testing likely enhanced the content validity of the BSID in this population. The findings of this study confirm the content validity of the culturally adapted BSID-III-NL for use in Suriname.

Construct validity, or the extent to which items of the BSID-III-NL measured language, motor, and cognitive functioning, was confirmed through the convergent increases of age and subscale raw scores. The BSID-III-NL captured a large proportion of significant differences between age groups, where 87 out of 180 possible pairwise comparisons between age groups across the five subscales were significantly different. These age-related increases in performance were similar to results by previous studies assessing the construct validity of culturally adapted BSID-III in different countries [26,27]. The sensitivity of the BSID-III in detecting age-related increases in performance is consistent with our findings of content validity. Without cultural adaptations, a previous study in a LMIC, failed to find significant evidence of convergent validity associated with age-related increases in performance [28]. Alternatively, with appropriate consideration of cultural differences, the significant age-related increases found across all age groups, and in each subscale, in the present study provide strong evidence of the convergent validity in this adaptation of the BSID-III-NL [27]. Taken together, these results support the utilization of the ITC's recommendations for cultural adaptation of educational and psychological assessments in new populations.

In addition to demonstrating high content and construct validity, this study reports the extent to which items in the subscales were reliably intercorrelated, or the internal consistency, in accordance with ten of the fifteen requirements described in the Guidelines for Reporting

Reliability and Agreement Studies (GRRAS) and consistent with ITC guidelines [15,29]. Overall, the internal consistency of each subscale was above 0.77. The Cronbach's alphas for this study were similar to previous Cronbach's alpha estimates for the culturally adapted BSID-III in studies conducted with Iranian, Malaysian, Ethiopian, and Vietnamese infants, and the mean estimate across all subscales is higher than those previously reported [27,30-32]. Notably, the lowest Cronbach's alpha reported for any subscale in this sample was higher than that reported for younger age groups in the original BSID-III scale construction and validation (0.71), providing further evidence of the internal consistency of the five subscales across three distinct geographical regions and the ethnically diverse Surinamese population [12].

Construct validity was also demonstrated through cluster analysis. Due to the binary nature of raw scores from participants, and the utilization of the strict procedural guidelines for administration, construct validity testing through exploratory factor analysis was not feasible. As an alternative, cluster analysis revealed that at least 81.56% of all variability, across age groups, was explained by a two-cluster model of correct and incorrect responses for all subscales. The variance explained by the two clusters supports the item sequence of the BSID-III-NL and the ability of the scale to capture developmental trajectories. Further investigation by the authors of alternative cluster analyses without items 22 in the Expressive Communication, 34 in the Fine Motor, and 48 and 53 in the Cognitive subscales (e.g. picture identification, quadrapod grasp, and relational play respectively) did not increase the amount of variance explained by the two-cluster model, making the removal or repositioning of those items unnecessary. Beyond the significant amount of construct validity, there were additional notable observations within and between subscales.

The increasing size of the correct response cluster with age groups seen in Receptive Communication, Expressive Communication, Gross Motor, and the Cognitive subscales was less pronounced in the Fine Motor subscale. Multiple items of interest identified by cluster analysis in the Fine Motor subscale were items that involved holding a pencil. Early in life, Surinamese infants, particularly in the Interior Region, are not commonly exposed to written styles of communication or pencils. While items involving the pencils were clustered in the incorrect response cluster for most participants, the youngest infants were able to complete advanced Fine Motor items. This is displayed in the end of the correct response cluster of the youngest age group, around 10 months of age, corresponding to the Fine Motor item 30, which required the infant to deliberately draw on paper.

Similar observations occurred in the Receptive Communication and Gross Motor subscales, where the end of the correct response clusters of the youngest age group corresponded to the start points of the oldest age groups. The youngest age group's correct response cluster had two items in the Cognitive subscale that were past the oldest age groups' start point. These were items 48 and 53, which assessed the infants' ability to demonstrate relational play with themselves and with BSID-III-NL objects, i.e. feeding a stuffed animal with a spoon. These results in the Cognitive subscale may reflect advanced imaginative play or performance of domestic activities in Surinamese infants.

Despite the evidence of content and construct validity in this initial study, there are limitations. First, this study evaluated only part of the complete validated age range of the BSID [12,13]. Analysis of the BSID-III-NL for other age groups in this population has not been completed, and caution should be utilized when administering this scale without further evaluation of validity and reliability. Second, although Dutch is the official language of

Suriname and the BSID-III-NL is written in Dutch, there are 25 recognized languages in Suriname [33]. Discrimination between language difference and developmental delay is a challenge in multilinguistic infants and continues to be a hurdle when assessing neurodevelopment in diverse populations. Despite the evidence in this study of high construct validity in both the Receptive and Expressive Communication domains, the BSID-III may underestimate language development in infants exposed to multiple languages even when considering cultural adaptations, and further studies are needed to assess the degree, if any, of underestimation [34,35]. Third, as only eight infants in this subset of participants were born premature, additional analysis of the validity and reliability of the BSID-III-NL in preterm infants, as well as the ability of the BSID-III-NL to capture developmental differences in preterm infants compared to full term infants as needed. Lastly, this populations sample of Surinamese infants is a subset of participants in the CCREOH study with known risk factors for developmental delay, including exposure to environmental toxicants in utero. As such, future studies are needed to address the ability of the BSID-III-NL to capture neurodevelopmental consequences of prenatal exposures and to provide guidance on the clinical utility of the BSID-III-NL, which this initial analysis was not designed to address.

In conclusion, this internationally utilized developmental measure, with minor cultural adaptations, was found to be reliable and valid for infants from 10 months to 26 months of age across three unique geographical and cultural regions in Suriname. The overall validity and reliability of the scale will be further addressed in the larger CCREOH cohort and will include additional considerations of differences by District and ethnic background in the larger population. As BSID-III-NL testing continues in Suriname, the creation of population-based norms and comparison to Dutch norms and United States norms may be feasible. Creation of

Surinamese norms will advance the clinical utility of the BSID-III-NL in Suriname. The establishment of the validity and reliability of the BSID-III-NL in Surinamese children was a critical first step in determining the neurodevelopmental impact of maternal exposure to the wide range of psychosocial stressors and environmental neurotoxicants found in Suriname.

References:

[1] World Bank. International Economics Dept. Development Data Group. World Development Indicators. World Bank; 1978.

[2] Hassankhan MS, Roopnarine L, White C, Mahase R, editors. Legacy of Slavery and Indentured Labour: Historical and Contemporary Issues in Suriname and the Caribbean. Routledge; 2016 Nov 10.

[3] Zijlmans W, Hindori-Mohangoo A. Determinants of neonatal mortality in Suriname: preliminary findings from a perinatal and infant mortality survey. Annals of Global Health. 2015;1(81):121-2.

[4] Zijlmans W, Wickliffe J, Hindori- Mohangoo A, et al. Caribbean Consortium for Research in Environmental and Occupational Health (CCREOH) Cohort Study: influences of complex environmental exposures on maternal and child health in Suriname. BMJ Open 2020;10:e034702. doi:10.1136/bmjopen-2019-034702

[5] Hawkins W, Mans D, Lichtveld MY. The Caribbean Consortium for Research in Environmental and Occupational Health (CCREOH). West Indian Medical J. Suppl., 2015;1-75.

[6] Ouboter PE, Landburg G, Satnarain GU, Starke SY, Nanden I, Simon-Friedt B, Hawkins WB, Taylor R, Lichtveld MY, Harville E, Wickliffe JK. Mercury levels in women and children from interior villages in Suriname, South America. Int. J. of Environ. Res. and Public Health. 2018 May;15(5):1007.

[7] Wickliffe, JK. Multi-element and metal concentrations in blood samples from pregnant women in the Republic of Suriname. In: Proceedings of the *International Society for Environmental Epidemiology;* August 2019; Utrecht, the Netherlands. 2019.

[8] Hambleton, Ronald K., Peter F. Merenda, and Charles D. Spielberger, eds. *Adapting educational and psychological tests for cross-cultural assessment*. Psychology Press, 2004.

[9] Bayley N. Manual for the Bayley Scales of Infant Development. Psychological Corporation; 1969.

[10] Dudani A. The Quality of Caregiver-Infant Interactions and Neurodevelopmental Outcomes of Infants Prenatally Exposed to Pesticides in Banana Growing Areas in Costa Rica.

[11] Albers CA, Grieve AJ. Test review: Bayley, N. (2006). Bayley Scales of Infant and Toddler Development–Third Edition. San Antonio, TX: Harcourt Assessment. J. of Psychoeduc. Assess. 2007 Jun;25(2):180-90.

[12] Bayley N. Bayley Scales of Infant and Toddler Development. PsychCorp, Pearson; 2006.

[13] Baron IS. Neuropsychological evaluation of the child: Domains, methods, & case studies. Oxford University Press; 2018 May 31.

[14] Van Baar AL, Steenis LJ, Verhoeven M, Hessen DJ. Bayley-III-NL. Technische Handleiding.

[15] International Test Commission. The ITC Guidelines for Translating and Adapting Tests, Second Edition. 2017.

[16] Walker K, Badawi N, Halliday R, Laing S. Brief report: performance of Australian children at one year of age on the Bayley scales of infant and toddler development (version III). Australian Educ. and Dev. Psychol., The. 2010;27(1):54.

[17] Ruiter SA, Spelberg HC, van der Meulen BF, Nakken H. The BSID-II-NL: construction, standardisation, and instrumental utility. Neth. J. of Psychol. 2008 Mar 1;64(1):15-40.

[18] Steenis LJ, Verhoeven M, Hessen DJ, van Baar AL. First steps in developing the Dutch version of the Bayley III: Is the original Bayley III and its item sequence also adequate for Dutch children? Eur. J. of Dev. Psychol. 2014 Jul 4;11(4):494-511.

[19] Steenis LJ, Verhoeven M, Hessen DJ, Van Baar AL. Performance of Dutch children on the Bayley III: a comparison study of US and Dutch norms. PloS one. 2015 Aug 12;10(8):e0132871.

[20] Bayley N. Bayley Scales of Infant and Toddler Development: Administration Manual. Harcourt Assessment; 2006.

[21] IBM Corp. IBM SPSS Statistics for Macintosh. Version 26.0; Armonk, NY: IBM Corp; 2019.

[22] SAS Institute Inc. SAS Computer software. Version 9.4; Cary, NC: SAS Institute Inc; 2013.

[23] Team RC. R: A Language and Environment for Statistical Computing. Version 3.6.1. R Foundation for Statistical Computing. 2019.

[24] Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951 Sep 1;16(3):297-334.

[25] Nunnally JC. Psychometric Theory 3E. Tata McGraw-Hill education; 1994.

[26] Parveen M, Rahman ST, Islam S, Zaman SS, Hamadani J, Khan NZ. Adaptation of items of Bayley scales of infant development-II (BSID-II) suitable for Bangladeshi infants. Dhaka Univ. J. of Biol. Sci. 2014 Aug 20;23(2):187-95.

[27] Hanlon C, Medhin G, Worku B, Tomlinson M, Alem A, Dewey M, Prince M. Adapting the Bayley Scales of infant and toddler development in Ethiopia: evaluation of reliability and validity. Child: Care, Health and Dev. 2016 Sep;42(5):699-708.

[28] Rademeyer V, Jacklin L. A study to evaluate the performance of black South African urban infants on the Bayley Scales of Infant Development III. S. Afr. J. of Child Health. 2013 Jul 3;7(2):54-9.

[29] Kottner, Jan, Byron J. Gajewski, and David L. Streiner. "Guidelines for Reporting Reliability and Agreement Studies (GRRAS)." *International journal of nursing studies* 48.6 (2011): 659.

[30] Azari N, Soleimani F, Vameghi R, Sajedi F, Shahshahani S, Karimi H, Kraskian A, Shahrokhi A, Teymouri R, Gharib M. A psychometric study of the Bayley scales of infant and toddler development in Persian language children. Iranian J. of Child Neurol. 2017;11(1):50.

[31] Zakaria S, Seok CB, Sombuling A, Ahmad MS, Iqbal S. Reliability and validity for Malay version of Bayley scales of infant and toddler development–third edition (Bayley-III): Preliminary study. Int. Proceedings of Economics Dev. and Res. Available online at: http://www.ipedr.com/vol40/024-ICPSB2012-P10017.pdf. 2012.

[32] Nguyen KV. Adaptation of the Bayley Scales of Infant and Toddler Development, (Bayley-III) for Vietnam: A preliminary study. St. John's University (New York); 2017.

[33] Eberhard DM, Simons GF, Fennig CD. Ethnologue: Languages of the World. Dallas, Texas: SIL International; 2020. Available from: http://www.ethnologue.com.

[34] Williams CJ, McLeod S. Speech-language pathologists' assessment and intervention practices with multilingual children. Int. J. of Speech-Lang. Pathol. 2012 Jun 1;14(3):292-305.

[35] Goh SK, Tham EK, Magiati I, Sim L, Sanmugam S, Qiu A, Daniel ML, Broekman BF, Rifkin-Graboi A. Analysis of item-level bias in the Bayley-III language subscales: the validity and utility of standardized language assessment in a multilingual setting. J. of Speech, Lang., and Hear. Res. 2017 Sep 18;60(9):2663-71.

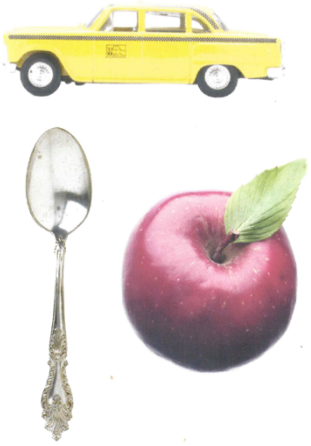
Supplementary Table 1. Item numbers of starting points for each age group by subscale

	Receptive Communication Subscale	Expressive Communication Subscale	Gross Motor Subscale	Fine Motor Subscale	Cognitive Subscale
Age Group 1	Item 6	Item 3	Item 19	Item 19	Item 25
Age Group 2	Item 8	Item 7	Item 22	Item 22	Item 31
Age Group 3	Item 8	Item 7	Item 22	Item 22	Item 31
Age Group 4	Item 10	Item 10	Item 35	Item 26	Item 34
Age Group 5	Item 13	Item 14	Item 39	Item 28	Item 34
Age Group 6	Item 13	Item 14	Item 39	Item 28	Item 34
Age Group 7	Item 13	Item 17	Item 42	Item 28	Item 40
Age Group 8	Item 15	Item 20	Item 45	Item 31	Item 45
Age Group 9	Item 15	Item 20	Item 45	Item 31	Item 45

Supplementary Table 2. Examples of image replacements in the BSID-III-NL picture book and puzzles

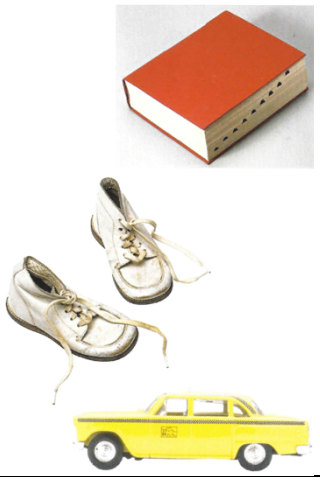
	BSID-III-NL Depiction	Culturally Adapted for the Interior Region

Picture Book Page 1	 	 
Picture Book Page 2	  	  
Picture Book Page 3	   3	  

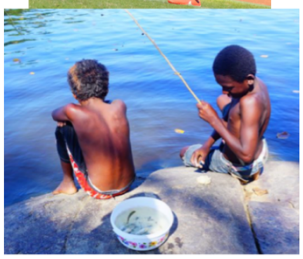
Picture Book Page 4		
Picture Book Page 5		
Picture Book Page 6	 	 

Picture Book Page 7	 	 
Picture Book Page 8	 	 
Picture Book Page 9	  9	 

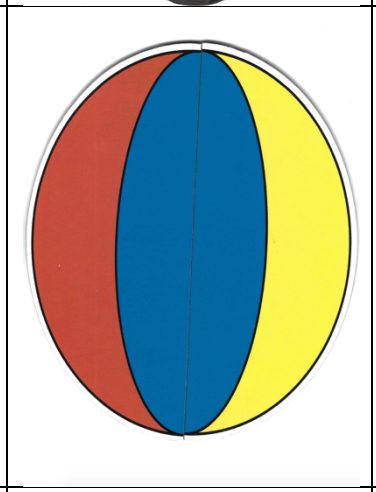
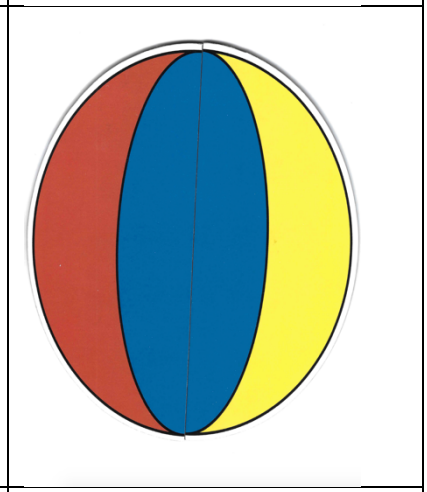
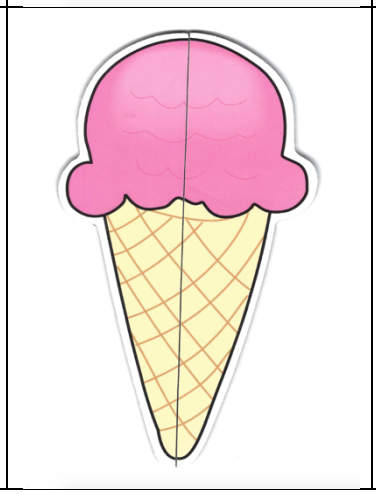
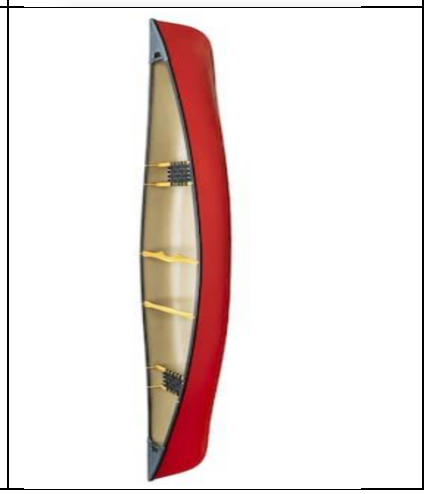
Picture Book Page 10		
Picture Book Page 11		
Picture Book Page 12		

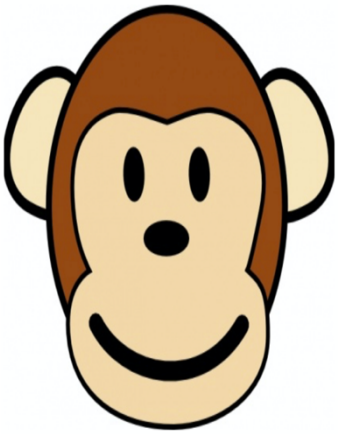
Picture Book Page 13		
Picture Book Page 14		
Picture Book Page 15		

Picture Book Page 16	 	 
Picture Book Page 17	 	 
Picture Book Page 18	  18	 

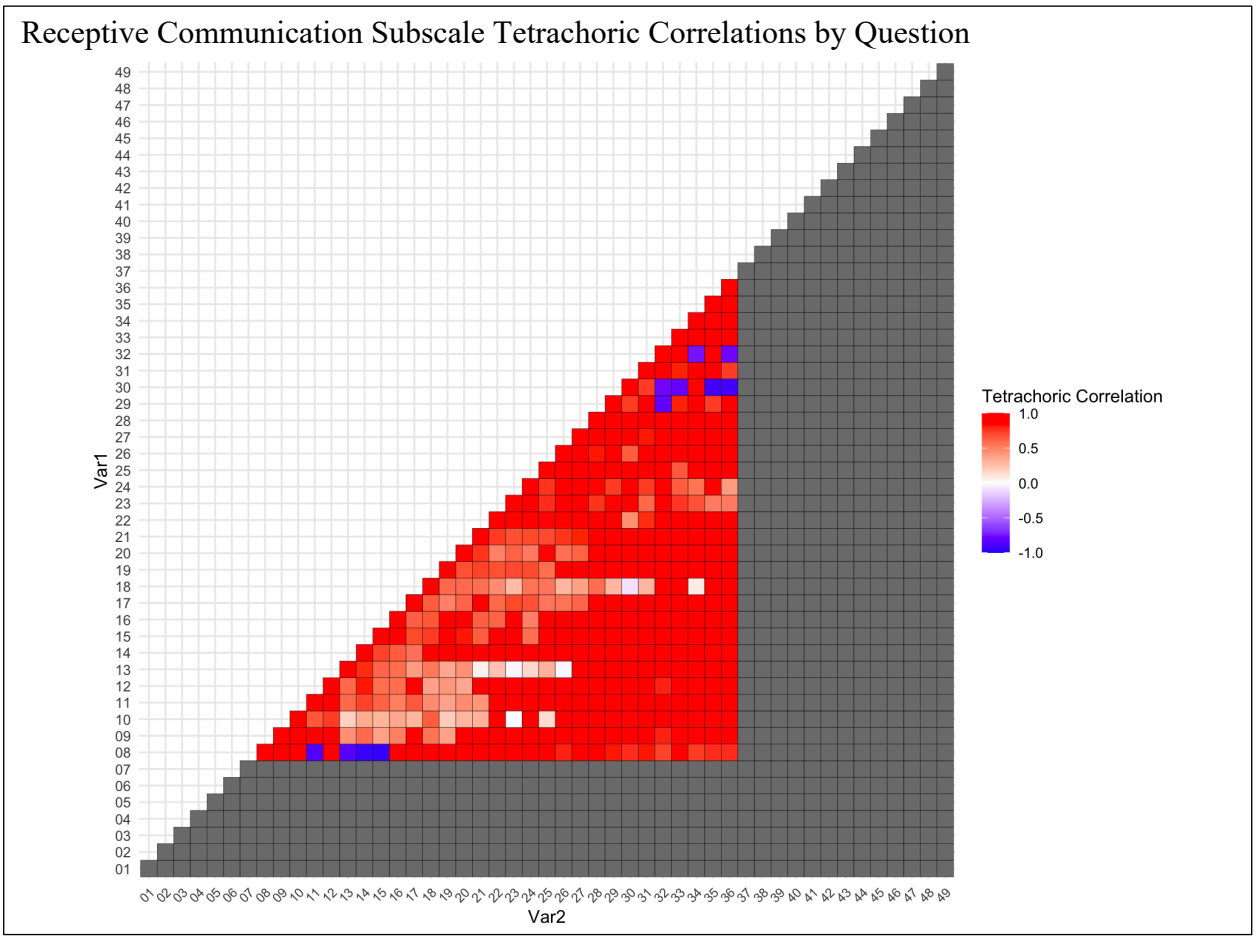
Picture Book Page 19	 	 
Picture Book Page 20	  20	 
Picture Book Page 21	 	 

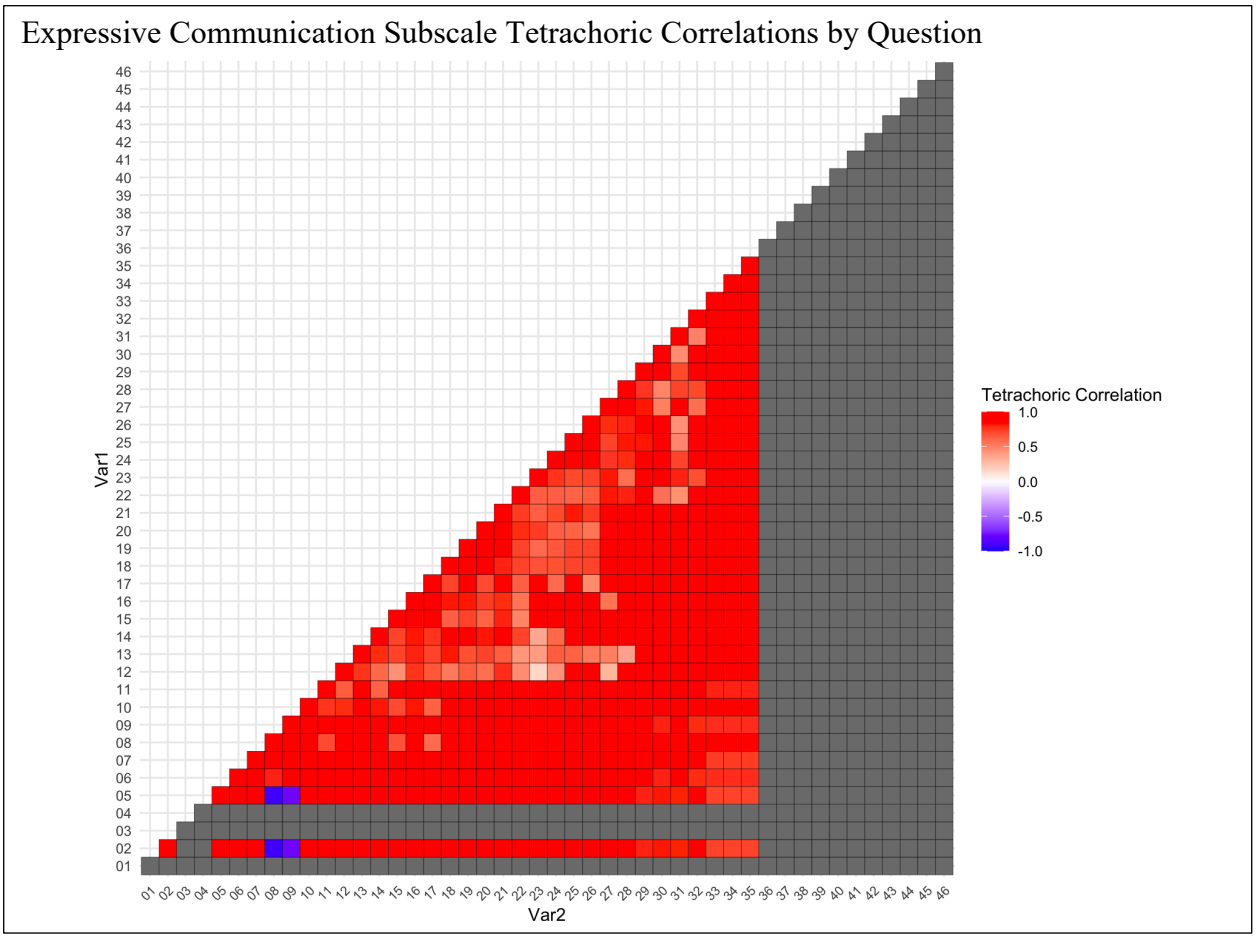
Picture Book Page 22	  	
Picture Book Page 23	  	  
Picture Book Page 24	   24	  

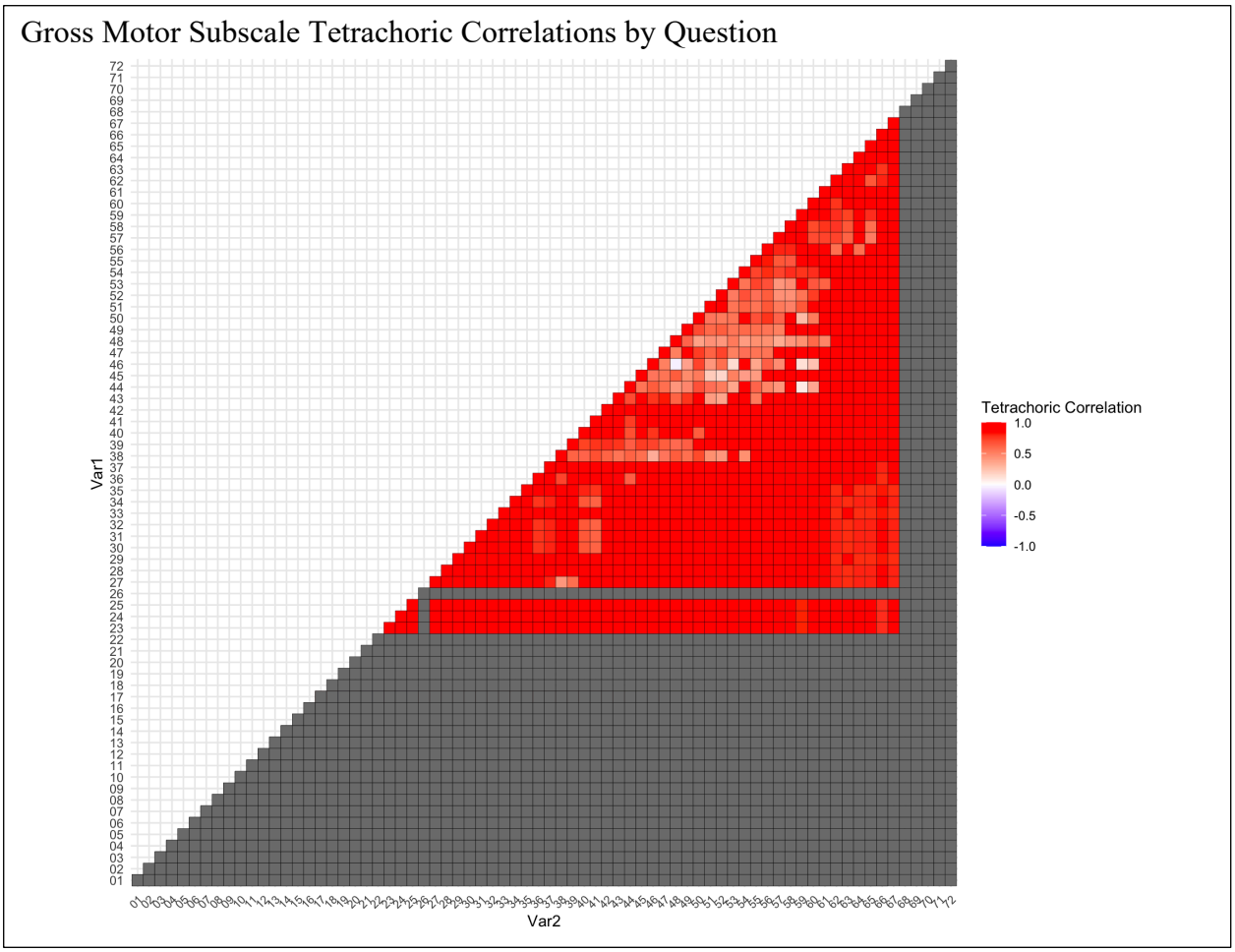
Picture Book Page 25		
Puzzle 1		
Puzzle 2		

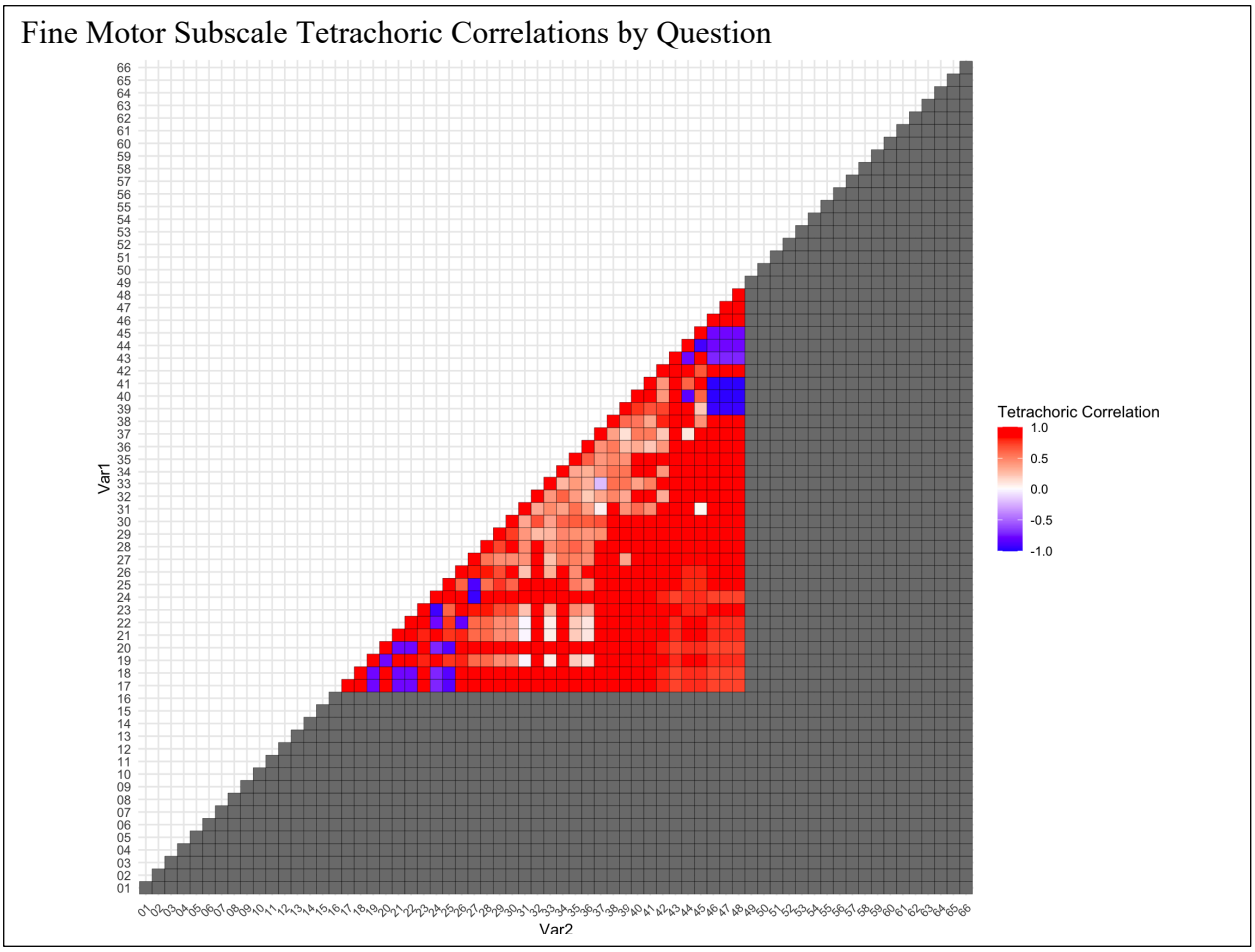
Puzzle 3		

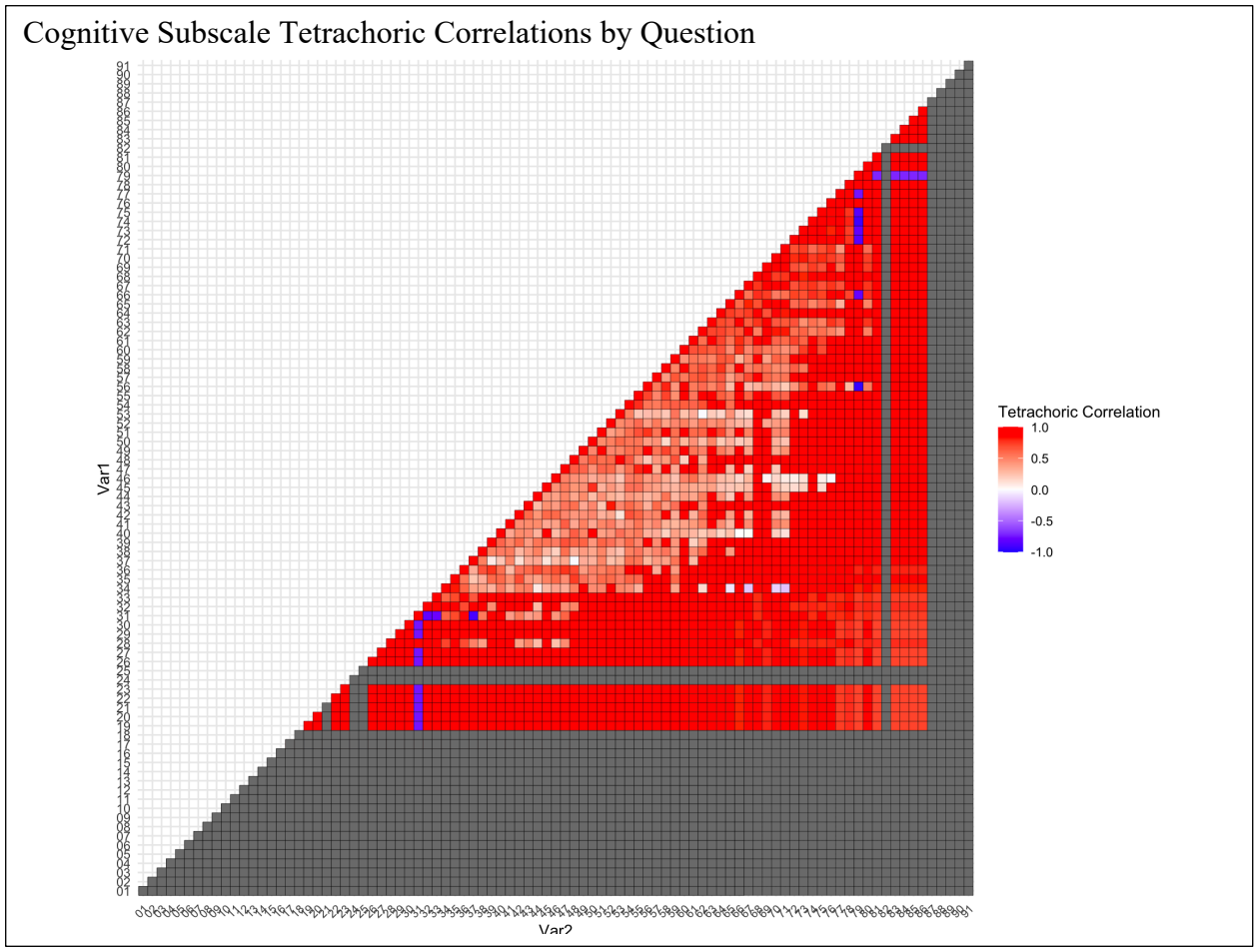
Supplementary Table 3. Tetrachoric correlations of BSID-III-NL subscale items within each subscale in correlation heat map format




Expressive Communication Subscale Tetrachoric Correlations by Question
Var1
Var2
Tetrachoric Correlation
1.0
0.5
0.0
-0.5
-1.0


Gross Motor Subscale Tetrachoric Correlations by Question
Var1
Var2
Tetrachoric Correlation
1.0
0.5
0.0
-0.5
-1.0


Fine Motor Subscale Tetrachoric Correlations by Question
Var1
Var2
Tetrachoric Correlation
1.0
0.5
0.0
-0.5
-1.0



Supplementary Table 4. Range of items of interest from cluster analysis by subscale

Receptive Communication Subscale	
Item Number	**Description of Scoring Criteria in Dutch and English**
14	Kind reageert op juistet manier op ten minste 1 sociaal verzoek
	Child responds in an appropriate manner to at least one spoken request
15	Kind identificeert ten minste 1 voorwerp als antwoord op de vraag
	Child correctly identifies at least one object
16	Kind identificeert het voorwerp dat u benoemt
	Chidl correctly identifies object you name
17	Kind identificeert ten minste 1 plaatje van de testitems
	Child correctly identifies at least one test item picture
18	Kind stopt een ogenblik als reactie op corrigerende woorden tijdens spel
	Child pauses in response to inhibitory words during a play routine
19	Kind identificeert ten minste 3 voorwerpen als antwoord op vraag
	Child correctly identifies at least three objects
20	Kind reageert met pop or beer op ten minste 2 verzoeken op juiste manier
	Child correctly responds to at least two directions with the doll or bear
21	Kind identificeert ten minste 3 plaatjes van de testitems

	Child correctly identifies at least three test item pictures
22	Kind identificeert ten minste 3 genoemde kledingstukken
	Child correctly identifies at least three clothing items
23	Kind identificeert ten minste 1 plaatje
	Child correctly identifies at least one picture
24	Kind wijst ten minste 5 lichaamsdelen aan
	Child corretly points to at least five body parts
25	Kind voert ten minste 1 van de tweevoudige opdrachten in zijn geheel uit
	Child correctly follows at least one two-part diection in its entirety
26	Kind identificeert ten minste 3 plaatjes
	Child correctly identifies at least three pictures
Expressive Communication Subscale	
Item Number	**Description of Scoring Criteria in Dutch and English**
12	Kind produceert ten minste 1 vocalisatie met een expressieve stembuiging
	Child produces at least one vocalization that contains inflections and is expressive
13	Kind imiteert ten minste 4 herhaalde medeklinker-klinker combinaties
	Child imitates at least four different repetitive consonant-vowel combinations
14	Kind produceert ten minste 1 woord
	Child produces at least 1 one-word approximation
15	Kind wijst naar of laat u of de ouder ten minste 1 voorwerp zien
	Child points to or shows you at least one object
16	Kind imiteert ten minste 1 woord, ook al imiteert kind alleen klinnkerklanken
	Child imitates at least one word, even if imitation consists of vowels only
17	Kind neemt ten minste 1 keer initiatief tijdens spel
	Child initiates at least one interaction for play
18	Kind gebruikt ten minste 2 juiste woorden
	Child uses at least two different words appropriately
19	Kind gebruikt ten minste 1 woord om verlangens kenbaar te maken
	Child uses at least one word to make wants known
20	Kind benoemt ten minste 1 voorwerp juist
	Child correctly names at least one object
21	Kind combineert ten minste 1 woord en gebaar
	Child uses at least one word and gesture combination
22	Kind benoemt ten minste 1 plaatje juist
	Child correctly names at least one picture
23	Kind gebruikt ten minste 8 woorden juist
	Child uses at least eight different words appropriately
24	Kind gerbruikt ja **of** nee in ten minste 2 reacties juist
	Child uses *yes* or *no* appropriately in at least two responses
25	Kind imiteert 2-woord zin
	Child imitates a two-word or multiple-word utterance
26	Kind doet ten minste 1 uitspraak die uit minimal 2 woorden bestaat die beiden iets anders betekenen
	Child produces at least one utterance that includes two or more words, each of which denotes a different concept

27	Kind benoemt ten minste 3 voorwerpen juist
	Child correctly names at least three objects
28	Kind benoemt ten minste 5 plaatjes juist
	Child correctly names at least five pictures
Gross Motor Subscale	
Item Number	**Description of Scoring Criteria in Dutch and English**
37	Kind zet gecoördineerde en alternerende stappen, het kind mag hierbij 1 of allebei uw handen ter ondersteuning gebruiken
	Child walks by making coordinated, alternating stepping movements
38	Kind loopt zijwaarts terwijl het zich vasthoudt aan onder- steunend voorwerp
	Child walks sideways while holding onto furniture for support and balance
39	Kind zakt doelbewust en gecontroleerd om van staande houding naar zittende houding te komen
	Child purposely lowers from a standing to a sitting position in a controlled manner
40	Kind staat gedurende ten minste 3 seconden alleen nadat u zign/haar handen heft losgelaten
	Child stands alone for at least 3 seconds after you release his or her hands
41	Kind gaat staan door eerst op buik te draaien of door op handen en voeten te gaan staan zonder ondersteuning
	Child comes to a standing position, rolling first to a prone or quadruped position, without using any support
42	Kind zet ten minste 3 stappen zonder hulp ook al oogt de looppas nog stijf en onvast
	Child takes at least three steps without support, even if gait is stiff-legged and wobbly
43	Kind zet ten minste 5 stappen alleen en laat daarbij coördinatie en balans zien
	Child takes at least five steps independently, displaying coordination and balance
44	Kind gooit bal doelbewust naar voren met onder- of boven-handse beweging
	Child purposely throws ball forward
45	Kind beweegt zonder steun van staande houding naar hurkende houding en weer terug en bewaart daarbij het evenwicht
	Child moves from standing to squatting to standing while maintaining blance without using any support
46	Kind rolt naar sijligging en staat op zonder hulp
	Child rolls to one side and stands without using any support
47	Kind loopt ten minste 3 traptreden omhoog en zet beide voeten op elke traptrede; kind mag gebruik maken van de muur of de trapleuning als ondersteuning
	Child walks up at least three steps, using wall or handrail for support; child places both feet on each step before stepping up to the next
48	Kind loopt zonder hulp ten minste 2 stappen achteruit
	Child takes at least two steps backward unassisted
49	Kind loopt ten minste 3 traptreden omlaag en zet beide voeten op elke traptrede; kind mag gebruik maken van de muur of de trapleuning als ondersteuning

	Child walks down at least three steps, using wall or handrail for support; child places both feet on each step before stepping down to the next
50	Kind rent met goede coördinatie
	Child runs with good coordination
51	Kind balanceert op rechtervoet terwijl u 1 van zijn/haar handen vashoudt
	Child balances on right foot while you hold one of his or her hands
52	Kind balanceert op linkervoet terwijl u 1 van zijn/haar handen vasthoudt
	Child balances on left food while you hold one of his or her hands
53	Kind loopt ten minste 2 stappen zijwaarts zonder hulp
	Child takes at least two steps sideways without support
54	Kind springt vanaf onderste traptrede naar grond
	Child jumps to floor
55	Kind behoudt evenwicht tijdens voorwaarts schoppen tegen bal over ten minste 60 cm
	Child maintains balance while kicking ball in a forward direction at least 2 feet
56	Kind houdt tijdens lopen ten minste 1 voet op looplijn over ten minste 1,5 meter
	Child walks with at least one foot (i.e., left foot or right foot) on path for at least 5 feet
57	Kind loopt ten minste 3 traptreden omhoog zonder gebruik te maken van muur of leuning en zet beide voeten op elke traptrede
	Child walks up three steps without using wall or handrail for support; child places both feet on each step before stepping up to the next
Fine Motor Subscale	
Item Number	**Description of Scoring Criteria in Dutch and English**
30	Kind tekent of krast spontaan en doelbewust op papier
	Child spontaneously and purposely scribbles on the paper
31	Kind stapelt ten minste 2 blokjes in 1 van de pogingen
	Child stacks at least 2 blocks
32	Kind tekent streep in willekeurige richting
	Child produces a stroke in any direction
33	Kind doet 10 graanringetjes 1 voor 1 in flesje binnen 60 seconden
	Child places 10 cereal pieces in bottle in 60 seconds or less, one pellet at a time
34	Kind houdt krijtje of potlood vast en gebruikt vingers in partiële duimoppositie om streep op papier te zetten
	Child grasps crayon or pencil using fingers and partial thumb opposition while making a mark on the paper
35	Kind doet minstens 3 munten in spaarpot
	Child places at least three coins into slot
36	Kind halt alle blokken uit elkaar
	Child takes all the blocks apart
37	Kind houdt krijtje of potlood met driepunts- of vierpuntsgreep vast en zet krassen op papier
	Child grasps crayon or pencil using a static tripod (thumb and two fingers) or quadrupod (thumb and three fingers) grasp while making a mark on the paper

38	Kind stapelt ten minste 6 blokjes in 1 van de pogingen
	Child stacks at least six blocks
39	Kind houdt papier vast met 1 hand terwijl het met andere hand tekent of schrijft
	Child holds paper in place with one hand while he or she scribbles or draws with the other
40	Kind tekent horizontale streep (maximal ongeveer 30^{o} afwijkend van uw horizontale streep)
	Child imitates horizontal stroke on paper within approximately 30 degrees of your horizontal line
Cognitive Subscale	
Item Number	**Description of Scoring Criteria in Dutch and English**
40	Kind vindt armband door eerst onder juiste doekje te kijken zowel aan linker- als aan rechterkant (niet noodzakelijk in dezelfde poging)
	Child finds bracelet by looking first under correct washcloth when hidden on both left and right sides
41	Kind pakt ring op en laat deze aan touwtje hangen zonder dat ring tafeloppervlak raakt
	Child obtains ring and suspends it by string without the ring touching the table
42	Kind halt graanringetje doelbewust uit flesje door zekere inspanning te leveren
	Child purposely removes cereal piece from bottle using some form of direct effort
43	Kind bemachtigt voorwerp via open voorkant binnen 20 seconden
	Child retrieves object through open end of box within 20 seconds
44	Kind probeert in piepbeest te knijpen om geluid te maken
	Child attempts to squeeze toy to make the sound
45	Kind vindt armband door eerst onder juiste doekje te kijken zowel aan linker- als rechterkant (niet noodzakelijk in dezelfde poging)
	Child finds bracelet by looking first under correct washcloth when hidden on both left and right sides
46	Kind draait dop helemaal van flesje af
	Child unscrews lid until it comes off
47	Kind plaatst minstens 1 staafje meerdere keren in hetzelfde of in eer ander gat, of het kind plaatst 2 of meerdere staafjes in het pennenbord binnen 70 seconden; geek ook 1 punt wanneer na 2 succesvolle uitvoeringen, 1 of meerdere staafjes worden verwijderd
	Child places at least one peg two or more times in the same or different hole(s)
48	Kind laat relationeel spel zien door met voorwerpen te spelen zoals bedoeld: naar zichzelf gericht
	Child demonstrates relational play with him- or herself
49	Kind plaatst ten minste 1 schijf correct in puzzelbord binnen 180 seconden
	Child correctly places at least one piece within 180 seconds
50	Kind vindt armband door direct onder juiste doekje te kijken zowel aan linker- als rechterkant (niet noodzakelijk in dezelfde poging)
	Child finds bracelet by looking first under correct washcloth when hidden on both left and right sides

51	Kind plaatst ten minste 1 schijf correct in puzzelbord binnen 150 seconden
	Child correctly places at least one piece within 150 seconds
52	Kind bemachtigt binnen 20 seconden per kant voorwerp via open kant van bakje, zowel met opening lniks als rechts
	Child retrieves object through open end of box when presented on both left and right sides
53	Kind laat relationeel spel zien door tweede person in spel te betrekken met voorwerpen zoals bedoeld
	Child demonstrates relational play, using objects for how they are intended, with others
54	Kind doet alle 9 blokjes in beker in 1 poging
	Child places all nine blocks inside cup at one time
55	King plaast alle 6 staafjes in pennenbord binnen 70 seconden
	Child places all six pegs in pegboard within 70 seconds
56	Kind plaatst alle 3 de schijven correct in puzzelbord binnen 180 seconden
	Child correctly places all three pieces within 180 seconds
57	Kind maakt vegende beweging met potlood om eend te bemachtigen ook al lukt dit niet
	Child uses pencil to attempt to obtain duck
58	Kind plaatst ten minste 4 schijven juist in puzzelbord binnen 150 seconden
	Child correctly places at least four pieces within 150 seconds
59	Kind heft aandacht voor ten minste 15 pagina’s (helft van het hele verhaal)
	Child attends to entire story
60	Kind plaatst alle 3 de schijven juist in gedraaid puzzelbord
	Child correctly places all three pieces while board is in rotated position
61	Kind legt puzzel op juiste manier binnen 90 seconden
	Child correctly assembles object within 90 seconds in either trial
62	Kind plaatst alle 6 staafjes in pennenbord binnen 25 seconden; alle 6 de staafjes moeten tegelijkertijd in het pennenbord staan
	Child places all six pegs in pegboard within 25 seconds
63	Kind legt puzzel op juiste manier binnen 90 seconden in 1 van de pogingen
	Child correctly assembles object within 90 seconds in either trial
64	Kind wijst naar juiste Plaatje op alle 3 de pagina’s
	Child correctly identifies matching picture on at least three pages
65	Kind pakt voorwerp en doet alsof het iets anders is
	Child takes an object and pretends it is something else

CHAPTER 5

Prenatal Environmental Exposures and Cellular Aging Impact the Social-Emotional Development of Surinamese Infants

Background:

Individuals are exposed to a mixture of metals at a given point in time, in possibly toxic concentrations, because metals co-exist in the environment. Some metals, like mercury (Hg), lead (Pb), cadmium (Cd), and manganese (Mn), have been studied for their toxic effects on the human nervous system. Previous research has found even low levels of Hg, Pb, Cd, and Mn to be individually toxic to neurons, but exposure of these components as a mixture is less understood where the toxic effects of individual components may depend on interactions with other neurotoxicants (Hernández et al., 2020; Tsentsevitsky and Petrov, 2021; The World Health Organization, 2023a; The World Health Organization, 2023b).

Exposure of neurotoxicant mixtures on the developing nervous system occurs during pregnancy as some neurotoxic metals cross the placental barrier through passive diffusion (Caserta et al., 2013). During a pregnancy, exposure to neurotoxicants can results in delayed cellular and global neurodevelopment for the offspring (Entringer et al., 2011; Valeri et al., 2017). The individual effects of neurotoxicants on an infant's nervous system have been previously reported, but the mixture of Hg, Pb, Cd, and Mn has not been reported to date.

Cellular responses to neurotoxicant exposure at the molecular level have long-term consequences across organism development. For example, the shortening or lengthening of telomeres has been reported in response to metal exposures (Zhang et al., 2013; Louzon et al., 2019). Telomeres are repetitive, non-coding DNA sequences, responsible for protecting linear chromosomes from genotoxic stressors like Hg, Pb, Cd, and Mn exposure (Cebulska-Wasilewska

et al., 2005; Cheng et al., 2007; Huang et al., 2010; Nicolai et al., 2021; Nagaraju et al., 2022). Altered telomere maintenance in the prenatal period is predictive of accelerated cellular dysfunction and disease susceptibility across the lifespan (Entringer et al., 2011).

Measuring telomere length (TL) as a biological marker of exposure to various neurotoxicants across the lifespan has been investigated in multiple cohorts, where the direction of TL response is dependent upon the type of neurotoxicant exposure (Zhang et al., 2013; Whiteman et al., 2017). Some neurotoxicants shorten TL, increasing susceptibility to aging-related diseases, while others lengthen TL through activation of telomerase and other mechanisms promoting tumorigenesis. Meta-analytic evidence shows that TL is also a biomarker of neurodevelopment and neurological aging (Casavant et al., 2019; Gampawar et al., 2022). However, few studies have utilized advanced statistical methods, such as Bayesian kernel machine regression (BKMR), to investigate telomeres and neurotoxicant effects individually on neurodevelopment (Shah-Kulkarni et al., 2020). Furthermore, there is no study to date which has investigated the combined contributions of telomere biology and neurotoxicants on neurodevelopment, despite the ability to account and control for multiple interactive pathways with recent statistical advances.

Global neurodevelopment of infants is impacted by many prenatal factors, including preterm birth and exposure to Hg, Pb, Cd, and Mn (Martin and Dombrowski, 2008; Tang et al., 2008; Mora et al., 2018; Kim et al., 2020). The development of social and emotional behaviors in infants is especially impacted by prenatal neurotoxicant exposure (Hubbs-Tait et al., 2005; Brucker-Davis et al., 2015; Torche, 2018; Liu et al., 2021). Higher prenatal exposure to Mn alters the functional connectivity of emotional areas of infant brains and higher concentrations of Cd have been associated with broad emotional problems in infants (Sioen et al., 2013; de Water

et al., 2018). Increased prenatal Pb concentrations are associated with neurobehavioral problems in children, including increased emotional reactivity, anxiety, and attention difficulties (Burns et al., 1999; Plusquellec et al., 2010; Liu et al., 2014; Fruh et al., 2019). Furthermore, higher Hg concentrations in pregnancy are associated with neurodevelopmental deficits in infants in prospective population-based studies (Steuerwald et al., 2000; Oken et al., 2005; Jedrychowski et al., 2006; Suzuki et al., 2010; Llop et al., 2012).

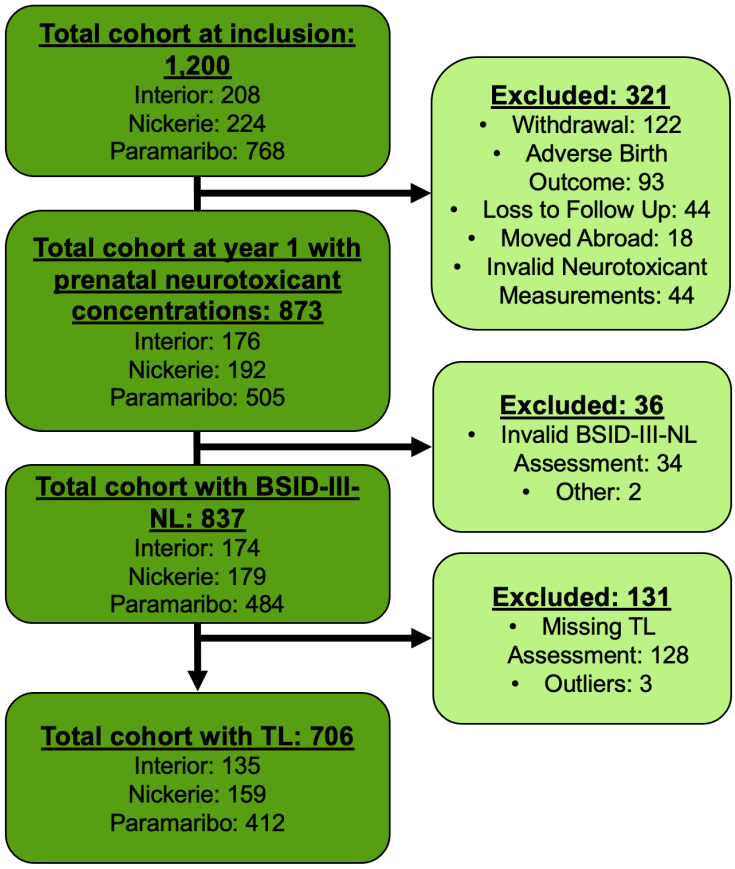


Figure 1. Caribbean Cohort for Research in Environmental and Occupational Health Flow Diagram for the Current Study. The Bayley Scales of Infant and Toddler Development, Third Edition, Dutch Version, represented by "BSID-III-NL" and Telomere Length by "TL".

The Caribbean Consortium for Research in Environmental and Occupational Health (CCREOH)'s prospective environmental population-based birth cohort study was established in the Republic of Suriname, in part, to assess the longitudinal effects of prenatal neurotoxicant

Maternal Demographics (n = 709)
average age of 28.02 years old (σ = 6.24 years)
average BMI of 26.85 (σ = 5.90; missing 4.65%)
Region of Suriname
58.39% Paramaribo (n = 414)
22.57% Nickerie (n = 160)
19.04% Interior (n = 135)
Maternal Parity to Gravidity Ratio
41.89% complete parity/gravidity ratio (n = 297)
32.44% incomplete parity/gravidity ratio (n = 230)
43.33% 1 or 2 living children (n = 307)
22.28% 3 or more living children (n = 158)
25.39% nulliparous (n = 180)
0.28% unknown (n = 2)
Maternal Educational Achievement
4.94% not educated (n = 35)
17.49% primary education (n = 124)
14.95% lower vocational education (n = 106)
0.85% upper vocational education (n = 6)
16.64% lower secondary education (n = 118)
29.48% upper secondary education (n = 209)
15.51% tertiary education (n = 110)
0.14% unknown education (n = 1)
Maternal Ethnicity
18.90% Tribal (n = 134)
20.73% Creole (n = 147)
12.00% Indigenous (n = 85)
20.73% Mixed (n = 147)
6.77% Javanese (n = 48)
20.73% Hindustani (n = 147)
0.14% unknown (n = 1)
Annual Household Income
4.37% <400 SUR annual household income (n = 31)
11.85% 400 to 799 SUR annual household income (n = 84)
19.18% 800 to 1499 SUR annual household income (n = 136)
28.21% 1500 to 2999 SUR annual household income (n = 200)
21.86% 3000 to 4999 SUR annual household income (n = 155)
8.04% 5000 to 9999 SUR annual household income (n = 57)
2.54% 10000 to 14999 SUR annual household income (n = 18)
0.85% >15000 SUR annual household income (n = 6)
3.10% unknown SUR annual household income (n = 22)
Household Size
0.70% 1 in household (n = 5)
10.16% 2 in household (n = 72)
20.31% 3 in household (n = 144)
68.55% 4 or more in household (n = 486)
0.28% unknown (n = 2)
Maternal Relationship Status
88.43% married or living with partner (n = 627)
11.43% not married or living with partner (n = 81)
0.14% unknown status (n = 1)
Maternal Smoking
7.05% smoked within 2 years of pregnancy (n = 50)
88.29% did not smoke within 2 years of pregnancy (n = 626)
4.65% unknown smoking within 2 years of pregnancy (n = 33)
Maternal Drinking
53.17% consumed alcohol within 2 years of pregnancy (n = 377)
42.18% did not consume alcohol within 2 years of pregnancy (n = 299)
4.65% unknown alcohol consumption within 2 years of pregnancy (n = 33)

Table 1. Maternal Demographics. Standard deviation represented by "σ" and body mass index by "BMI".

exposures on Surinamese infant development (Zijlmans et al., 2020). The Republic of Suriname

is a middle-income country (MIC) in northeastern South America named after the Indigenous Surinen tribe (World Bank, 2019). Due to multiple forced movements of slaves and hegiras of immigrants since Dutch colonization in the 17th century, the population of Suriname is highly diverse and includes Tribal, Creole, Indigenous, Javanese, Hindustani, and individuals of mixed descent (Hassankhan et al., 2016).

The majority of the Surinamese economy exists of gold mining and agriculture, in addition to crude oil drilling, fisheries, forestry, and ecotourism (General Bureau of Statistics, 2020). A consequence of these economic activities is unregulated environmental pollution and exposures to occupational hazards (Strategic Environmental Advice / Tropenbos International Suriname, 2017). These include the contamination of water and fish sources due to runoff Hg from gold mining and application of pesticides and herbicides containing Mn and Pb for increased agricultural yields of bananas, rice, and cassava (Gullino et al., 2010; Abdoel Wahid et al., 2017; Rimbaud et al., 2017; Ouboter et al., 2018; Ministry of Agriculture. Animal Husbandry and Fisheries Mancozeb Import Statistics, 2019; Ohiagu et al., 2020; Wickliffe et al., 2021). Subsequently, there are unacceptably high Hg concentrations in popularly consumed fish and Pb levels in the national drinking water distribution system (Ouboter et al., 2012; Ouboter et al., 2018; Suriname Water Company, 2019).

The presence of these neurotoxicants in the food and water consumed by pregnant Surinamese women has negative effects for both mother and offspring (Kort et al., 2022; Koendjbiharie et al., 2023). However, the majority previous research on the effects Hg, Pb, Cd, and Mn exposures on neurodevelopment has been conducted with cross-sectional analyses of single exposure-response relationships and few analyses have also assessed the concurrent role of TL in neurodevelopmental outcomes. This study capitalizes on the comprehensive data

collected from the CCREOH cohort to investigate the effects of the prenatal Hg, Pb, Cd, and Mn exposure mixture and infant TL on infant neurodevelopment.

Methodology:

Caribbean Cohort for Research in Environmental and Occupational Health

The prospective epidemiologic Caribbean Consortium for Research in Environmental and Occupational Health (CCREOH) cohort study recruited pregnant women and their offspring from three geographic locations in Suriname as previously described, including Paramaribo, Nickerie, and the Interior region of Suriname (Zijlmans et al., 2020). Mothers between the ages of 18 and 45 years-old were invited to participate in the study at prenatal clinics and midwife facilities. Recruitment for the study began in December 2016 and completed in July 2019, where the first birth was on February 4th, 2017, and the last birth was on December 4th, 2019. Following informed consent, demographic information on parity and gravidity, maternal age, educational attainment, household income, household size, ethnicity, relationship status, body mass index (BMI), smoking and drinking behaviors, and expected due date were obtained for 1,200 participants (Figure 1; Table 1). A trained research assistant conducted study interviews and questionnaires in Dutch. Whole blood was collected from mothers during pregnancy in trace element vacutainers with potassium EDTA via venipuncture by trained research assistants for assessment of neurotoxicant concentrations which was completed for 873 participants (Figure 1). All participants had singleton births with infant sex, birth weight, length, head circumference and prematurity (less than thirty-seven weeks gestation) recorded at birth. By the third year of the CCREOH study, all infants were brought into one of the research centers between the ages of ten months and forty-one months of age to obtain a buccal swab and assess neurodevelopment which was successful for 837 participants (Figure 1). All study procedures were conducted according to

the principles of the Declaration of Helsinki and the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use, Good Clinical Practice guidelines. The study protocol was reviewed and approved by the Central Committee on Research Involving Human Subjects, the Medical Ethical Committee of the Ministry of Health in

Infant Demographics (n = 709)
48.70% female infants (n = 345)
51.30% male infants (364)
Prematurity
8.60% premature (less than 37 weeks gestation) birth (n = 61)
91.40% non-premature birth (n = 648)
Infant Birth Statistics
average weight of 3,120.02 grams (σ = 484.84 grams)
5.79% low (less than 2,500 grams) birth weight (n = 41)
average length of 48.49 centimeters (σ = 2.29 centimeters)
average head circumference of 32.88 centimeters (σ = 2.29 centimeters)
Infant Age at Bayley Assessment
average age of 18.66 months old (σ = 5.14 years)
minimum of 10.45, maximum of 41.43 months old
Bayley Assessment Percentile Scores
average of 34.46% in Socioemotional (σ = 26.49%)
average of 62.83% in Cognitive (σ = 26.44%)
average of 39.86% in Language (σ = 25.12%)
average of 56.98% in Motor (σ = 26.40%)
Infant Telomere Length
average telomere length of 2.09 (σ = 0.72)
minimum of 0.39, maximum of 14.73
Prenatal Neurotoxicant Metal Concentrations
average cadmium level of 0.22 (σ = 0.13 micrograms per deciliter)
minimum of 0.02, maximum of 0.97 micrograms per deciliter
average manganese level of 15.14 (σ = 5.62 micrograms per deciliter)
minimum of 4.81, maximum of 39.80 micrograms per deciliter
average mercury level of 6.06 (σ = 8.69 micrograms per deciliter)
minimum of 0.18, maximum of 73.40 micrograms per deciliter
average lead level of 3.36 (σ = 3.90 micrograms per deciliter)
minimum of 0.39, maximum of 33.88 micrograms per deciliter

Table 2. Infant Demographics. Standard deviation represented by "σ". The United Sates Environmental Protection Agency action level for blood lead levels is 5µg/dL. The suggested action for blood mercury levels is 3.5µg/dL.

Suriname (VG 023-14), and the Institutional Review Board of Tulane University's School of Public Health and Tropical Medicine (83-093), New Orleans, Louisiana, USA.

Prenatal Neurotoxicant Concentrations

Maternal whole blood samples were processed and stored frozen at -80°C in the Clinical Chemistry Laboratory, Academic Hospital, Paramaribo, Suriname. Samples were shipped frozen on dry ice to the Wisconsin State Laboratory of Hygiene Trace Element Research Laboratory

(Madison, Wisconsin, USA) or the Wadsworth Center Trace Elements Laboratory Core (Albany, New York, USA) using a formal chain-of-custody process. Concentrations of Hg, Pb, Cd, and Mn were determined using magnetic sector inductively coupled plasma mass spectrometry. Standard reference materials, matrix spikes, and method duplicates were used for quality assurance and quality checking. Interlaboratory comparisons were performed to ensure no significant differences in concentrations of Hg, Pb, Cd, and Mn for 30 duplicate samples sent to both laboratories. All internal and external quality control criteria were acceptable. Differences between mean concentrations within individuals where samples were collected across trimesters did not vary, so sample method duplicates were averaged to provide single value concentrations for each neurotoxicant per participant prior to statistical analysis.

Bayley Scales of Infant and Toddler Development, Third Edition, Dutch Version

The Bayley Scales of Infant and Toddler Development, Third Edition, Dutch Version (BSID-III-NL) was administered by trained research staff, including pediatric neuropsychologists, to infants around eighteen months of age, with an age range of BSID-III-NL assessment of ten months to forty-one months of age (Table 2). The cognitive, language, motor, and social-emotional domain assessments of the BSID-III-NL were administered, with validity and reliability of most of these assessments previously reported in this population (McLester-Davis et al., 2021). Primary outcomes utilized in statistical analyses were the composite scores which are standardized to a mean score of 100 and a standard deviation of 15. Each composite score calculation was double-check and confirmed between a minimum of two research personnel.

Social-Emotional Development

The social-emotional domain of the BSID-III-NL was originally developed by child psychiatrist Stanley Greenspan, Ph.D., and adapted to the BSID-III-NL with scoring comparable to the cognitive, language, and motor domains with translation into Dutch (Greenspan, 2004; Bayley, 2006; Van Baar et al., 2014). The social-emotional domain assesses an individual infant's or toddler's "interest in the world, self-regulation, engagement in relationships, interactive and playful emotions, emotional signals or gestures in communication or problem solving, use of ideas to convey feelings, wishes, or intentions, and connection of emotions to ideas logically" as previously described (Breinbauer et al., 2010; Pearson, 2019). The social-emotional BSID-III-NL domain was administered to the primary caregiver via questionnaire format by research personnel where all participants started with the first item and proceeded to the stop point determined by the infant's chronological age with adjustment for premature birth. Each item was a declarative statement and responses were provided to a 6-point Likert scale ranging from "cannot tell if a behavior is displayed" to "the behavior is displayed all of the time." A raw score was created from adding the responses to each applicable item (0 to 6 points per item), then the raw score was converted into a scaled score that created comparable scores across age groups, and then converted into a composite score that is standardized to a Gaussian distribution as previously described, hereby referred to as Social-Emotional Development (Breinbauer et al., 2010).

Telomere Length Measurement

Immediately following administration of the BSID-III-NL assessment, trained research personnel swabbed both interior sides of an infant's mouth using an Isohelix SK1 buccal swab (Cell Projects, Kent, United Kingdom). Swabs were stored with a desiccant at 4°C for up to one year in the Clinical Chemistry Laboratory, Academic Hospital, Paramaribo, Suriname. Samples

were sent to Tulane University at room temperature for DNA extraction and TL analysis. DNA was extracted using the QIAamp DNA Mini Kit protocol (Qiagen, Valencia, CA). Concentration of extracted DNA was quantified with a Qubit dsDNA HS or BR Assay Kit (Invitrogen, Carlsbad, CA) and purity of the DNA was determined using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA). DNA was stored at −80°C following quality check and thawed at room temperature prior to TL analysis.

The average relative buccal cell TL was determined from the telomere repeat copy number to single gene (albumin) copy number (T/S) ratio using an adapted monochrome

ANOVA Summary	Component	CCREOH Region Comparison	Mean Difference	95% CI	p-value
F(2,646) = 299.71 p < 0.001	Mercury	Interior to Paramaribo	1.51	[1.38, 1.64]	<0.001
		Interior to Nickerie	1.60	[1.45, 1.75]	<0.001
F(2,646) = 265.51 p < 0.001	Lead	Interior to Paramaribo	1.37	[1.25, 1.49]	<0.001
		Interior to Nickerie	1.37	[1.22, 1.51]	<0.001
F(2,646) = 68.49 p < 0.001	Cadmium	Paramaribo to Nickerie	0.12	[0.01, 0.23]	0.037
		Paramaribo to Interior	0.67	[0.55, 0.78]	<0.001
		Nickerie to Interior	0.55	[0.41, 0.68]	<0.001
F(2,646) = 68.49 p = 0.015	Manganese	Nickerie to Paramaribo	0.10	[0.03, 0.17]	0.004
F(2,646) = 68.49 p < 0.001	Telomere Length	Paramaribo to Nickerie	0.15	[0.05, 0.25]	0.003
		Interior to Nickerie	0.21	[0.09, 0.33]	<0.001

Table 3. Significant Differences of Mean Component Levels by CCREOH Region. One-way analysis of variance (ANOVA) significant results with least significant differences post-hoc analyses comparing each component's mean by Caribbean Consortium for Research in Environmental and Occupational Health region for prenatal mercury, lead, cadmium manganese, and infant telomere length. Confidence intervals are listed as "CI".

multiplex quantitative real-time polymerase chain reaction (MMqPCR) via a BioRad CFX96 as previously described (Cawthon, 2009; Drury et al., 2014). All samples were performed in triplicate, with a 7-point standard curve (0.0313–2 ng) using pooled control buccal DNA.

Triplicate plates were repeated with all samples in a different well position. Thus, 6 replicates were available for each infant. Buccal TL ratio was determined by the average of the triplicates from both plates. Further information on the TL analysis is available in the Supplemental Methods.

Covariates

For statistical analyses, maternal age, body mass index, household income, household size, infant birth weight, length, head circumference, and chronological age at BSID-III-NL assessment were modeled as continuous variables to adjust for confounding. Furthermore, maternal education (not educated, primary, lower vocational, upper vocational, lower secondary, upper secondary, or tertiary), ethnicity (Tribal, Creole, Indigenous, Mixed, Javanese, or Hindustani), smoking (smoked within 2 years or not), CCREOH region (Paramaribo, Nickerie, or Interior), and infant sex (female or male) were modeled as categorical variables to adjust for confounding. The 4.65% of missing covariate data is assumed to be missing at random (Tables 1 and 2).

Statistical Analyses

Concentrations of Hg, Pb, Cd, and Mn and TL were modeled as natural log-transformed and centered continuous variables to achieve a common scale and account for skewedness. A ratio of maternal parity and gravidity was created by dividing the number of living children by number of total pregnancies for descriptive statistics. Distributional plots and descriptive statistics were examined for concentrations of Hg, Pb, Cd, and Mn and TL by CCREOH region and maternal ethnicity, with one-way analysis of variance (ANOVA) with least significant differences (LSD) post-hoc analyses conducted using IBM Corp's Statistical Package for the Social Sciences (SPSS) version 27.0. The ANOVA of TL means by ethnicity was conducted

after removal of TL outliers (n = 3; Figure 1). Multivariable linear regression analysis predicting Social-Emotional Development was also conducted in SPSS. A probability value (p-value) of 0.05 was used to determine statistical significance. All BKMR analyses were conducted in R Core Team's R version 4.0.3. BKMR was conducted as previously described (Valeri et al., 2017) to assess synergistic and nonlinear effects among the components (concentrations of Hg, Pb, Cd, and Mn, and infant TL). The BSID-III-NL composite scores for cognitive, language, motor, and Socio-Emotional Development were assessed as outcomes using the Gaussian kernel while accounting for multiple-testing penalty (Scott & Berger, 2010).

Results:

Study Population Characteristics

Demographic characteristics of mothers are presented in Table 1. The mean age of mothers was 28 years, with equal Creole, Hindustani, and Mixed ethnicity (20.73%). A majority of mothers had secondary or higher educational attainment (61.63%) as defined in Suriname (UNICEF, 2019). A majority had previous pregnancies (74.61%) with a third of those mothers having lost an offspring in a previous pregnancy (32.44%). The household size was larger than four or more individuals for most pregnant mothers at the time of study enrollment (68.55%) with most married or living with their partner (88.43%) and an average annual income of 1,500 to 2,999 SUR ($47 to $93 USD as of 2023) or more. Only half reported alcohol consumption within 2 years of getting pregnant, while only 78.05% reported smoking cigarettes (Table 1).

Demographic and assessment characteristics of infants are presented in Table 2. There were an approximate equal number of females to males with 8.60% of all infants born premature. At birth, the average weight of infants was 3,120 grams with 5.79% of all infants born underweight. The average length of infants at birth was 48.49 centimeters with an average head

circumference of 32.88 centimeters. Infants were 18 months old on average at BSID-III-NL assessment and had an average 34.46 percentile ranking on Social-Emotional Development. The

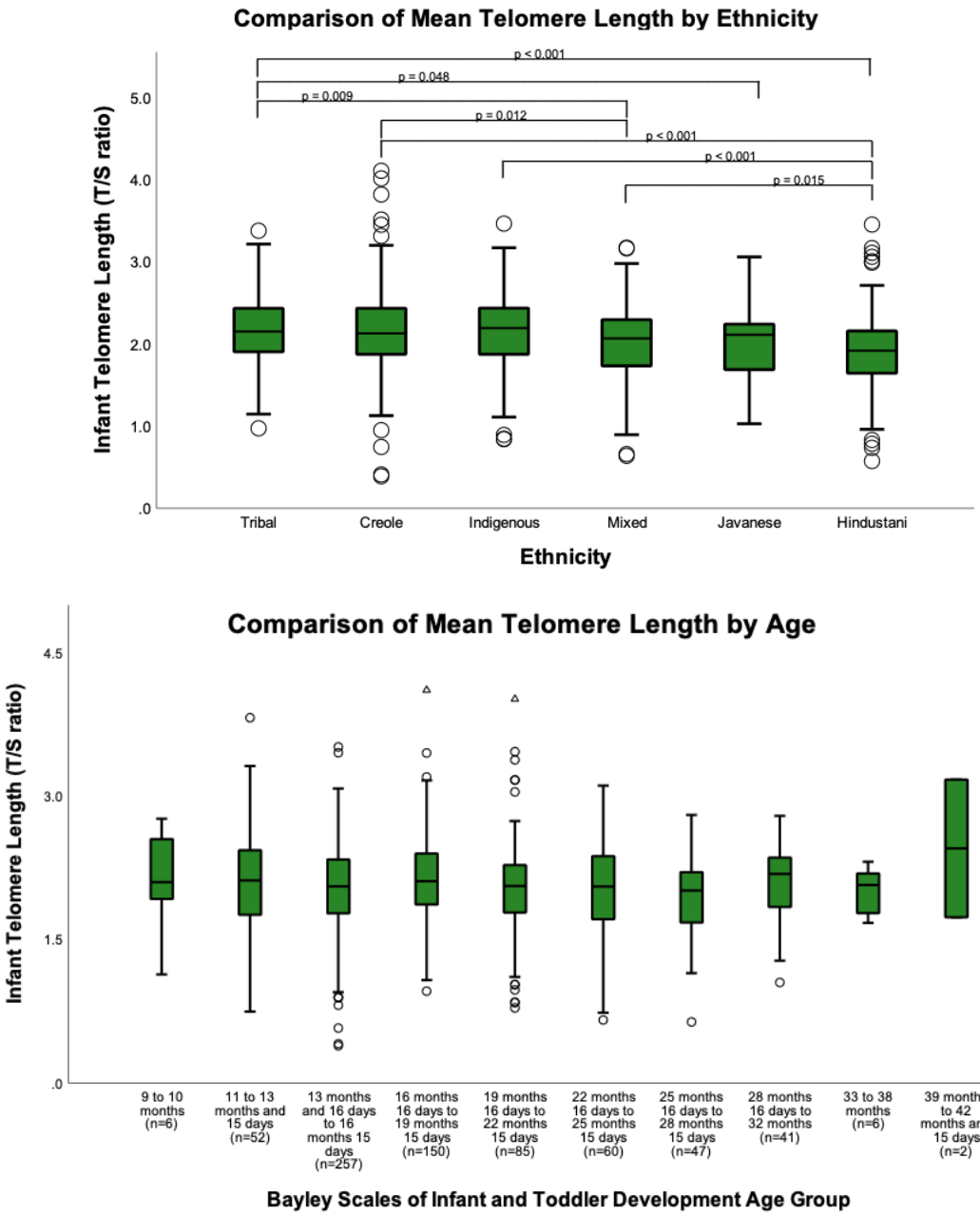


Figure 2. Comparison of Mean Telomere Length by Ethnicity and Age. Significant differences between ethnic groups as shown with p-values above the groups calculated from one-way analysis of covariance with least significant difference correction. Comparison by age excludes outlier telomere length values (n = 706).

average TL was 2.09 (standard deviation of 0.72) with a range from 0.39 to 14.73 including outliers. The non-natural log-transformed values for each neurotoxicant are also listed in Table 2.

Metal Concentrations Vary by Caribbean Consortium for Research in Environmental and Occupational Health Region

Following examination of distributional plots for Hg, Pb, Cd, and Mn concentrations by CCREOH region, a one-way analysis of variance was performed to identify if components of interest were significantly different by CCREOH region in Suriname. The Interior prenatal levels of Hg and Pb were significantly higher than Paramaribo and Nickerie with LSD post-hoc analysis (Table 3). Prenatal levels of Cd were significantly higher in Paramaribo than Nickerie and the Interior and Cd prenatal levels were significantly higher in Nickerie than the Interior. Nickerie had significantly higher Mn prenatal levels than Paramaribo. Infant TL did significantly differ by CCREOH region, where Nickerie had shorter TLs on average than Paramaribo or the Interior.

Telomere Length Varies by Ethnicity and Age

Following examination of TL by maternal ethnicity, a one-way analysis of variance was performed to discern significant differences between ethnicity categories for infant TL. Figure 2 top shows box plots of TL by ethnicity from longest (left) to shortest TL (right) on average. Average Tribal TL was significantly longer than Mixed, Javanese, and Hindustani average TL. Average Creole TL was significantly longer than Mixed average TL. Average Hindustani TL was significantly shorter than Creole, Indigenous, and Mixed average TL. Given the differences in age group sample sizes, there is not sufficient evidence to determine how age relates to TL

across this early period of life from this sample as seen in the bottom of Figure 2.

Model Summary	Outcome Variable	Predictor	β	p-value
$R^2 = 0.033$ F(16,494) = 2.09 p = 0.008	Social-Emotional Development	prenatal mercury*	-0.176	0.015
		prenatal lead	-0.057	0.418
		prenatal cadmium	-0.024	0.665
		prenatal manganese	0.024	0.610
		infant telomere length	-0.044	0.320
		CCREOH region*	0.172	0.008
		infant age*	-0.135	0.003
		infant sex	-0.014	0.750
		infant birth weight	-0.025	0.693
		infant birth length	0.007	0.901
		infant birth head circumference	0.086	0.090
		maternal age	0.048	0.326
		maternal ethnicity	-0.063	0.163
		maternal education	-0.004	0.941
		maternal smoking	0.040	0.368
		maternal body mass index	0.040	0.392

Table 4. Multivariable Linear Regression Results. Caribbean Consortium for Research in Environmental and Occupational Health represented by "CCREOH". Beta values represent the adjusted values.

Multivariable Linear Regression Analysis

Table 4 presents the effect of components of interest and covariates on the Social-Emotional Development of Surinamese infants using multivariable linear regression. Only prenatal levels of Hg, not Pb, Cd, Mn or infant TL, had a significant negative effect on Social-Emotional Development (p = 0.015). The CCREOH region (p = 0.008) and the age of the infant (p = 0.003) were also significant predictors where increased age decreased Social-Emotional Development.

Bayesian Kernel Machine Regression Analysis

Without the assumption of linearity used in the multivariable linear regression analysis, the BKMR analysis yielded an estimate of the exposure-response function for the mixture of Hg, Pb, Cd, and Mn concentrations and infant biological age as indexed by TL on Social-Emotional

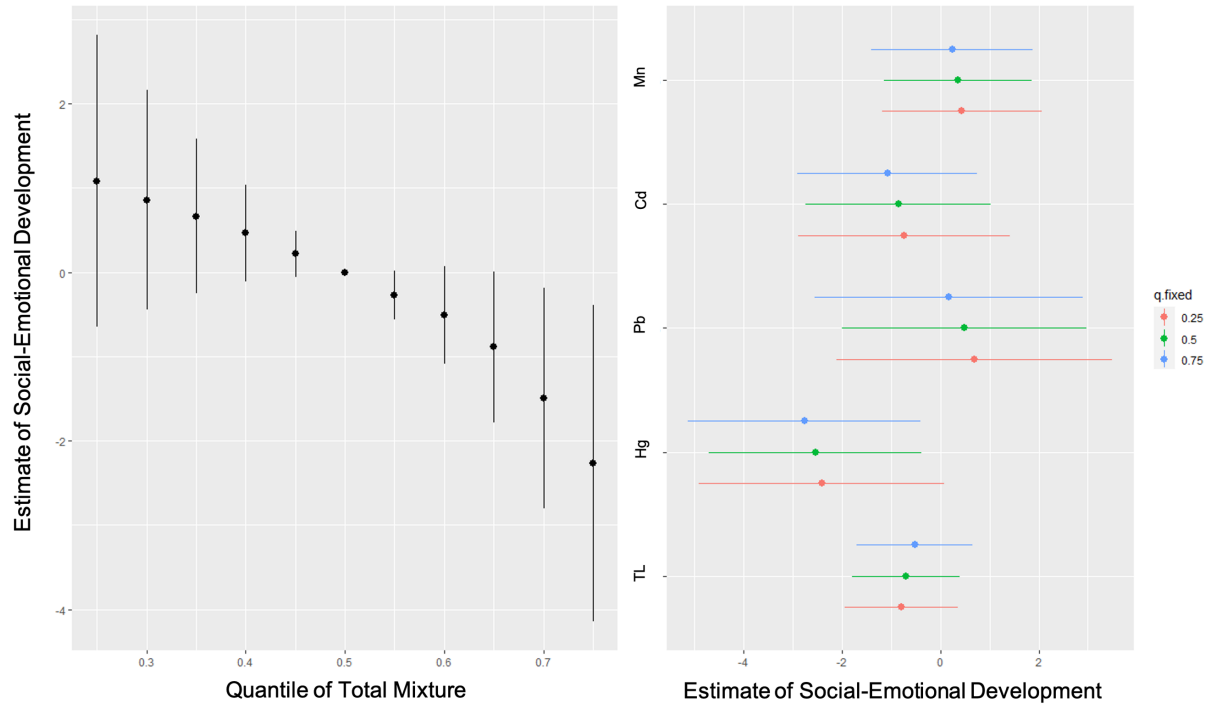


Figure 3. Overall Effect of the Mixture and Individual Components at Different Quantiles on Social-Emotional Development. Quantiles represented by "q", telomere length by "TL", mercury by "Hg", lead by "Pb", cadmium by "Cd", and manganese by "Mn".

Development. The cumulative effect of this mixture estimated by the expected change in Social-Emotional Development with concurrent changes in all mixture components from their median level is seen in the left portion of Figure 3. The cumulative effect of the mixture when components were at 70 quantile or higher significantly decreased Social-Emotional Development. In the right portion of Figure 3, the differences in interquartile range of Hg show significant negative effects of greater Hg prenatal levels (median and 75th percentile) on Social-Emotional Development.

Figure 4 shows the lack of interactive effect between the mixture components, and the non-linear effects of Hg on Social-Emotional Development. Furthermore, Figure 4 shows dose-response relationships of each mixture component and lack of interactions with components on Social-Emotional Development when all other components are at the median, 25th, or 75th

percentile. These results show that Hg prenatal levels are driving the overall effects of the mixture. The findings in Figure 4 suggest that the overall effect may be changed in a non-linear fashion for Hg and TL, Hg and Pb, and Hg and Mn when all other components are at the median. BKMR analyses were also performed for the BSID-III-NL cognitive, language, and motor domains and these non-significant results are available in the Supplemental Findings.

Discussion:

These findings show evidence of neurotoxicity given the mixture of prenatal Hg, Pb, Cd, Mn, and infant TL. Maternal exposure to high Hg concentrations during pregnancy has adverse effects on Social-Emotional Development for Surinamese infants. While not significant, lower amounts of the mixture showed an increase in Social-Emotional Development indicating a need to address exposure to Hg, Pb, Cd, and Mn for pregnant Surinamese individuals to protect offspring neurodevelopment. These findings align with previous research reporting the single prenatal exposures of Hg, Pb, Cd, or Mn on infant neurodevelopment (Burns et al., 1999; Steuerwald et al., 2000; Oken et al., 2005; Jedrychowski et al., 2006; Llop et al., 2012; Sioen et al., 2013; Liu et al., 2014; de Water et al., 2018; Fruh et al., 2019). Few studies have reported mixture effects of any combination of these components, and none have assessed the concurrent role of TL as a part of the causal pathway on neurodevelopment (Plusquellec et al., 2010; Suzuki et al., 2010). Given the significant differences in Hg, Pb, Cd, and Mn prenatal concentrations by CCREOH regions and differences in TL by ethnicity, future research should identify the sensitivity of these models by region and ethnicities in Suriname. While the multivariable regression model provided information on the significance of Hg in this mixture, it is important to note the benefit of BKMR analysis in complex neurotoxicant exposures as seen in this

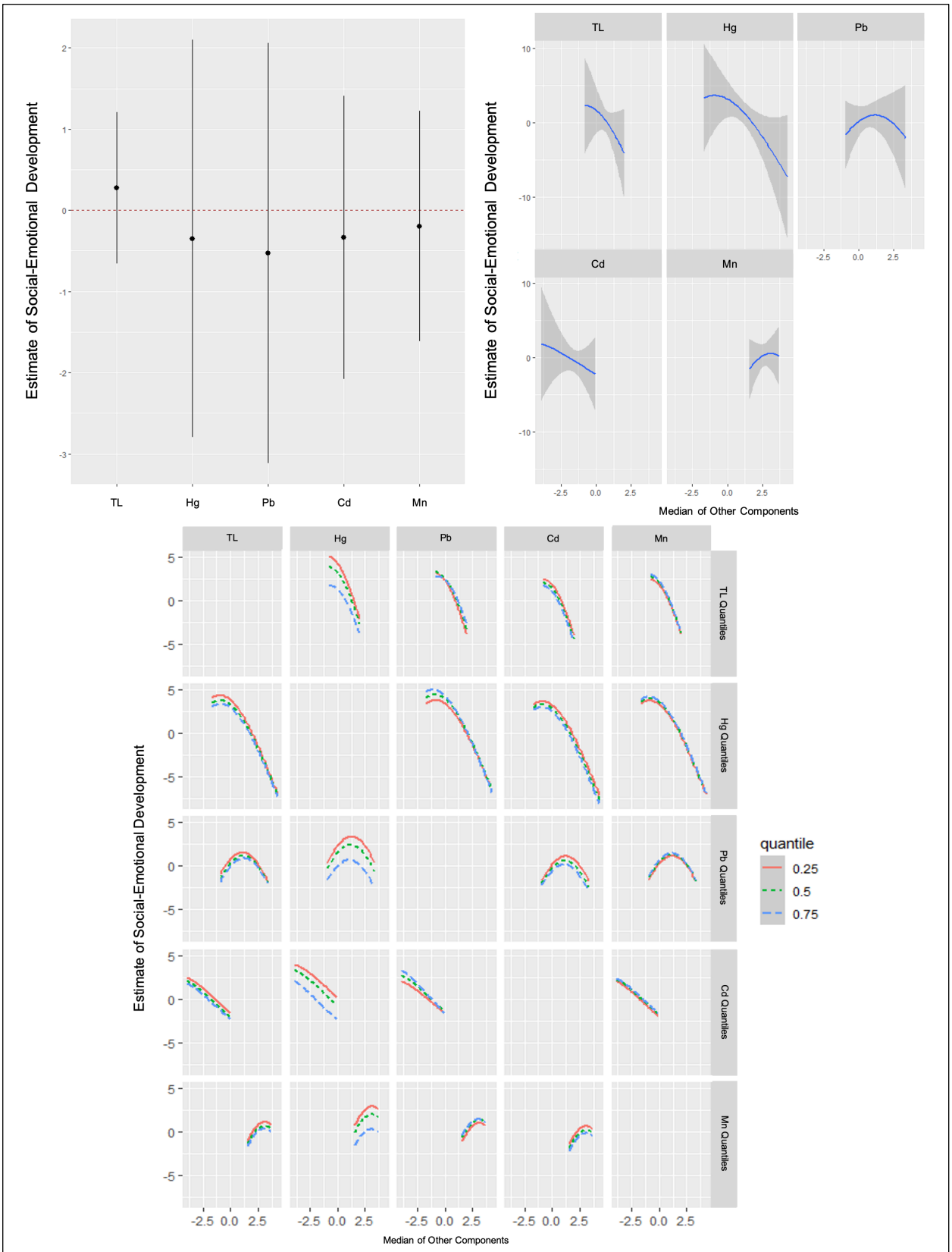


Figure 4. Lack of Interactive Effects Between Individual Components on Social-Emotional Development. Top left: Individual interactive effects by exposure. Top Right: Single-exposure effects on social-emotional development when other components are at the median. Bottom: Effect of an exposure (top row) on social-emotional development at quantiles for another exposure (right column) when others are at the median. Telomere length represented by "TL", mercury by "Hg", lead by "Pb", cadmium by "Cd", and manganese by "Mn".

Surinamese cohort, where Hg does not exhibit a completely linear relationship with Social Emotional Development.

Limitations

There are limitations to this work, including the measurement methodologies and components included in the analyses. First, in order to decrease the burden on the participant and the financial constraints of the study, prenatal blood was collected only once for most participants for assessment of Hg, Pb, Cd, and Mn concentrations. The ability to identify differences in neurodevelopmental outcomes by mixture concentrations of different trimesters of pregnancy, as reported in other studies, is limited with only one time point for neurotoxicant concentration measurement (Shah-Kulkarni et al., 2020). However, a small subset of pregnant participants provided multiple blood samples across trimesters. While differences by trimester have previously reported, no significant differences in neurotoxicant concentrations were found by trimester in the CCREOH participants who provided samples from multiple trimesters (unpublished). Therefore, it is unlikely that the observations found here would be altered if multiple blood samples across trimesters were collected per individual, but it is more likely that a higher dropout or withdrawal rate would be observed in the cohort due to the increased burden of research participation (Lingler et al., 2014). Second, DNA extracted from buccal swabs was utilized for TL measurement, which is not the tissue of direct interest for neurodevelopmental outcomes. However, as brain tissue is not ethically or feasibly obtainable for living participants, TL is meta-analytically correlated across tissues within individuals, and epithelial cells collected by buccal swabs derive from the same embryonic origins of brain tissue, it is likely that the TL measurements from buccal swabs are representative of brain tissue TL (McLester-Davis et al., 2023). Lastly, other genetic factors, micronutrients, and non-metal elemental exposures during

pregnancy have shown interactive effects with neurotoxicants that were not included in this study. For example, Apolipoprotein E genotype modifies the impact of prenatal Hg exposure on behavioral problems (Ng et al., 2013). Analysis of iodine and selenium concentrations in the mixture also minimizes the impact of neurotoxicants on infant development (Brucker-Davis et al., 2015; Liu et al., 2021). Furthermore, folic acid supplements are protective of the effects of Hg on Social-Emotional Development (Kim et al., 2020; Steenweg-de Graaff et al., 2012; Schlotz et al., 2010; Julvez et al., 2009). Therefore, these component and others should be considered in future research assessing the impacts of prenatal neurotoxicant mixtures on infant neurodevelopment.

Strengths

There are several strengths to this study, including analysis methodology, novel findings, and utility for the Republic of Suriname. These findings provide robust information on the effects of the mixture through BKMR analysis which is supported by the multivariable linear regression results. Application of BKMR allows for elucidation of any significant interactions and non-linear effects as seen in research for other populations (Valeri et al., 2017; Shah-Kulkarni et al., 2020). These findings show that Hg exposure during pregnancy is a primary target of intervention for the Republic of Suriname given the significant negative impact on Social-Emotional Development of Surinamese infants. Additionally, these findings are the first to provide information on the mixture effects of prenatal Hg, Pb, Cd, Mn exposure and infant TL on neurodevelopment. These findings further add to the existing literature of these neurotoxicants and cellular development and aging associations with neurodevelopment (Martin and Dombrowski, 2008; Tang et al., 2008; Mora et al., 2018; Kim et al., 2020). Furthermore, these findings provide scientific evidence for Surinamese policymakers of the multigenerational

impacts of environmental and occupational exposures to neurotoxicants specific to the Republic of Suriname.

Implications

The high concentrations of neurotoxicants present in pregnancy and subsequent negative effects on Social-Emotional Development of infants is important for Surinamese health professionals to consider when looking for developmental delays in this population so appropriate interventions can be identified and implemented in a timely manner. Furthermore, Surinamese policymakers should identify and implement limitations on environmental release and occupational exposures to Hg and other neurotoxicant given these findings (Hubbs-Tait et al., 2005; Heemskerk et al., 2016). Additionally, programs should be created in a region-specific manner to address prenatal dietary supplement needs, such as folic acid, that minimize the impact of neurotoxicant exposures for Surinamese infants (Baldewsingh et al., 2021).

Conclusions

In conclusion, these findings contribute to current literature on mixtures of neurotoxicants present in the prenatal period and long-term outcomes in infants. This study shows the neurotoxic effects of the mixture of prenatal Hg, Pb, Cd, Mn, and infant TL. Specifically, maternal exposure to high Hg concentrations during pregnancy has adverse effects on Social-Emotional Development for Surinamese infants. Policy and health interventions are next steps to prevent future neurodevelopmental delays for the Surinamese population.

References:

McLester-Davis, L.W.Y.; Estrada, P.; Hastings, W.J.; Kataria, L.A.; Martin, N.A.; Siebeneicher, J.T.; Tristano, R.I.; Mayne, C.V.; Horlick, R.P.; O'Connell, S.S.; et al. A Review and Meta-Analysis: Cross-Tissue Telomere Length Correlations in Healthy Human. *Under Review, Ageing Research Reviews* **2023**.

The World Health Organization Mercury and Health Available online: https://www.who.int/news-room/fact-sheets/detail/mercury-and-health (accessed on 13 February 2023).

The World Health Organization Lead Poisoning Available online: https://www.who.int/news-room/fact-sheets/detail/lead-poisoning-and-health (accessed on 13 February 2023).

UNICEF Suriname Education Fact Sheets Available online: https://data.unicef.org/resources/suriname-education-fact-sheets-2019/ (accessed on 13 February 2023).

Pearson Bayley-III Developmental Domains Available online: https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/bayley-iii/bayley-iii-domains.pdf.

Greenspan, S. *Greenspan Social-Emotional Growth Chart: A Screening Questionnaire for Infants and Young Children.*; PsychCorp, 2004;

Caserta, D.; Graziano, A.; Lo Monte, G.; Bordi, G.; Moscarini, M. Heavy Metals and Placental Fetal-Maternal Barrier: A Mini-Review on the Major Concerns. *Eur Rev Med Pharmacol Sci* **2013**, *17*, 2198–2206.

Brucker-Davis, F.; Ganier-Chauliac, F.; Gal, J.; Panaïa-Ferrari, P.; Pacini, P.; Fénichel, P.; Hiéronimus, S. Neurotoxicant Exposure during Pregnancy Is a Confounder for Assessment of Iodine Supplementation on Neurodevelopment Outcome. *Neurotoxicology and Teratology* **2015**, *51*, 45–51, doi:10.1016/j.ntt.2015.07.009.

Baldewsingh, G.K.; Hindori-Mohangoo, A.D.; van Eer, E.D.; Covert, H.H.; Shankar, A.; Wickliffe, J.K.; Shi, L.; Lichtveld, M.Y.; Zijlmans, W.C.W.R. Association of Mercury Exposure and Maternal Sociodemographics on Birth Outcomes of Indigenous and Tribal Women in Suriname. *IJERPH* **2021**, *18*, 6370, doi:10.3390/ijerph18126370.

Torche, F. Prenatal Exposure to an Acute Stressor and Children's Cognitive Outcomes. *Demography* **2018**, *55*, 1611–1639, doi:10.1007/s13524-018-0700-9.

Liu, J.; Martin, L.J.; Dinu, I.; Field, C.J.; Dewey, D.; Martin, J.W. Interaction of Prenatal Bisphenols, Maternal Nutrients, and Toxic Metal Exposures on Neurodevelopment of 2-Year-Olds in the APrON Cohort. *Environment International* **2021**, *155*, 106601, doi:10.1016/j.envint.2021.106601.

Schlotz, W.; Jones, A.; Phillips, D.I.W.; Gale, C.R.; Robinson, S.M.; Godfrey, K.M. Lower Maternal Folate Status in Early Pregnancy Is Associated with Childhood Hyperactivity and Peer Problems in Offspring: Folate Status in Pregnancy and Offspring Childhood Behavioural Difficulties. *Journal of Child Psychology and Psychiatry* **2010**, *51*, 594–602, doi:10.1111/j.1469-7610.2009.02182.x.

Julvez, J.; Fortuny, J.; Mendez, M.; Torrent, M.; Ribas-Fitó, N.; Sunyer, J. Maternal Use of Folic Acid Supplements during Pregnancy and Four-Year-Old Neurodevelopment in a Population-Based Birth Cohort. *Paediatric and Perinatal Epidemiology* **2009**, *23*, 199–206, doi:10.1111/j.1365-3016.2009.01032.x.

Martin, R.; Dombrowski, S.C. *Prenatal Exposures: Psychological and Educational Consequences for Children*; Springer: New York, 2008; ISBN 9780387743981.

Tsentsevitsky, A.N.; Petrov, A.M. Synaptic Mechanisms of Cadmium Neurotoxicity. *Neural Regen Res* **2021**, *16*, 1762–1763, doi:10.4103/1673-5374.306067.

Hernández, R.B.; Carrascal, M.; Abian, J.; Michalke, B.; Farina, M.; Gonzalez, Y.R.; Iyirhiaro, G.O.; Moteshareie, H.; Burnside, D.; Golshani, A.; et al. Manganese-Induced Neurotoxicity in Cerebellar Granule Neurons Due to Perturbation of Cell Network Pathways with Potential Implications for Neurodegenerative Disorders. *Metallomics* **2020**, *12*, 1656–1678, doi:10.1039/d0mt00085j.

Hubbs-Tait, L.; Nation, J.R.; Krebs, N.F.; Bellinger, D.C. Neurotoxicants, Micronutrients, and Social Environments: Individual and Combined Effects on Children's Development. *Psychol Sci Public Interest* **2005**, *6*, 57–121, doi:10.1111/j.1529-1006.2005.00024.x.

Breinbauer, C.; Mancil, T.L.; Greenspan, S. CHAPTER 5 - The Bayley-III Social-Emotional Scale. In *Bayley-III Clinical Use and Interpretation*; Weiss, L.G., Oakland, T., Aylward, G.P., Eds.; Practical Resources for the Mental Health Professional; Academic Press: San Diego, 2010; pp. 147–175.

Cheng, A.; Shin-ya, K.; Wan, R.; Tang, S.; Miura, T.; Tang, H.; Khatri, R.; Gleichman, M.; Ouyang, X.; Liu, D.; et al. Telomere Protection Mechanisms Change during Neurogenesis and Neuronal Maturation: Newly Generated Neurons Are Hypersensitive to Telomere and DNA Damage. *J. Neurosci.* **2007**, *27*, 3722–3733, doi:10.1523/JNEUROSCI.0590-07.2007.

Huang, J.; Okuka, M.; McLean, M.; Keefe, D.L.; Liu, L. Telomere Susceptibility to Cigarette Smoke-Induced Oxidative Damage and Chromosomal Instability of Mouse Embryos in Vitro. *Free Radical Biology and Medicine* **2010**, *48*, 1663–1676, doi:10.1016/j.freeradbiomed.2010.03.026.

Cebulska-Wasilewska, A.; Panek, A.; Żabiński, Z.; Moszczyński, P.; Au, W.W. Occupational Exposure to Mercury Vapour on Genotoxicity and DNA Repair. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **2005**, *586*, 102–114, doi:10.1016/j.mrgentox.2005.06.009.

Casavant, S.G.; Cong, X.; Moore, J.; Starkweather, A. Associations between Preterm Infant Stress, Epigenetic Alteration, Telomere Length and Neurodevelopmental Outcomes: A Systematic Review. *Early Human Development* **2019**, *131*, 63–74, doi:10.1016/j.earlhumdev.2019.03.003.

Gampawar, P.; Schmidt, R.; Schmidt, H. Telomere Length and Brain Aging: A Systematic Review and Meta-Analysis. *Ageing Research Reviews* **2022**, *80*, 101679, doi:10.1016/j.arr.2022.101679.

Nicolai, M.M.; Weishaupt, A.-K.; Baesler, J.; Brinkmann, V.; Wellenberg, A.; Winkelbeiner, N.; Gremme, A.; Aschner, M.; Fritz, G.; Schwerdtle, T.; et al. Effects of Manganese on Genomic Integrity in the Multicellular Model Organism Caenorhabditis Elegans. *International Journal of Molecular Sciences* **2021**, *22*, 10905, doi:10.3390/ijms222010905.

Entringer, S.; Epel, E.S.; Kumsta, R.; Lin, J.; Hellhammer, D.H.; Blackburn, E.H.; Wüst, S.; Wadhwa, P.D. Stress Exposure in Intrauterine Life Is Associated with Shorter Telomere Length in Young Adulthood. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, doi:10.1073/pnas.1107759108.

Louzon, M.; Coeurdassier, M.; Gimbert, F.; Pauget, B.; de Vaufleury, A. Telomere Dynamic in Humans and Animals: Review and Perspectives in Environmental Toxicology. *Environment International* **2019**, *131*, 105025, doi:10.1016/j.envint.2019.105025.

Scott, J.G.; Berger, J.O. Bayes and Empirical-Bayes Multiplicity Adjustment in the Variable-Selection Problem. *Ann. Statist.* **2010**, *38*, doi:10.1214/10-AOS792.

Koendjbiharie, A.P.; Hindori-Mohangoo, A.D.; Zijlmans, W.C.W.R.; Wickliffe, J.K.; Shankar, A.; Covert, H.H.; Lichtveld, M.Y.; Grünberg, A.W.; Drury, S.S. The Single and Combined Effects of Prenatal Nonchemical Stressors and Lead Exposure on Neurodevelopmental Outcomes in Toddlers: Results from the CCREOH Environmental Epidemiologic Study in Suriname. *Children* **2023**, *10*, 287, doi:10.3390/children10020287.

Kort, S.A.R.; Wickliffe, J.; Shankar, A.; Shafer, M.; Hindori-Mohangoo, A.D.; Covert, H.H.; Lichtveld, M.; Zijlmans, W. The Association between Mercury and Lead Exposure and Liver and Kidney Function in Pregnant Surinamese Women Enrolled in the Caribbean Consortium for Research in Environmental and Occupational Health (CCREOH) Environmental Epidemiologic Cohort Study. *Toxics* **2022**, *10*, 584, doi:10.3390/toxics10100584.

Nagaraju, R.; Kalahasthi, R.; Balachandar, R.; Bagepally, B.S. Association between Lead Exposure and DNA Damage (Genotoxicity): Systematic Review and Meta-Analysis. *Arch Toxicol* **2022**, *96*, 2899–2911, doi:10.1007/s00204-022-03352-9.

de Water, E.; Proal, E.; Wang, V.; Medina, S.M.; Schnaas, L.; Téllez-Rojo, M.M.; Wright, R.O.; Tang, C.Y.; Horton, M.K. Prenatal Manganese Exposure and Intrinsic Functional Connectivity of Emotional Brain Areas in Children. *NeuroToxicology* **2018**, *64*, 85–93, doi:10.1016/j.neuro.2017.06.006.

Sioen, I.; Den Hond, E.; Nelen, V.; Van de Mieroop, E.; Croes, K.; Van Larebeke, N.; Nawrot, T.S.; Schoeters, G. Prenatal Exposure to Environmental Contaminants and Behavioural Problems at Age 7–8 Years. *Environment International* **2013**, *59*, 225–231, doi:10.1016/j.envint.2013.06.014.

Valeri, L.; Mazumdar, M.M.; Bobb, J.F.; Claus Henn, B.; Rodrigues, E.; Sharif, O.I.A.; Kile, M.L.; Quamruzzaman, Q.; Afroz, S.; Golam, M.; et al. The Joint Effect of Prenatal Exposure to Metal Mixtures on Neurodevelopmental Outcomes at 20–40 Months of Age: Evidence from Rural Bangladesh. *Environ Health Perspect* **2017**, *125*, 067015, doi:10.1289/EHP614.

Zhang, X.; Lin, S.; Funk, W.E.; Hou, L. Republished: Environmental and Occupational Exposure to Chemicals and Telomere Length in Human Studies. *Postgraduate Medical Journal* **2013**, *89*, 722–728, doi:10.1136/postgradmedj-2012-101350rep.

Whiteman, V.E.; Goswami, A.; Salihu, H.M. Telomere Length and Fetal Programming: A Review of Recent Scientific Advances. *Am J Reprod Immunol* **2017**, *77*, e12661, doi:10.1111/aji.12661.

Tang, D.; Li, T.; Liu, J.J.; Zhou, Z.; Yuan, T.; Chen, Y.; Rauh, V.A.; Xie, J.; Perera, F. Effects of Prenatal Exposure to Coal-Burning Pollutants on Children's Development in China. *Environmental Health Perspectives* **2008**, *116*, 674–679, doi:10.1289/ehp.10471.

Mora, A.M.; Córdoba, L.; Cano, J.C.; Hernandez-Bonilla, D.; Pardo, L.; Schnaas, L.; Smith, D.R.; Menezes-Filho, J.A.; Mergler, D.; Lindh, C.H.; et al. Prenatal Mancozeb Exposure, Excess

Manganese, and Neurodevelopment at 1 Year of Age in the Infants' Environmental Health (ISA) Study. *Environ Health Perspect* **2018**, *126*, 057007, doi:10.1289/EHP1955.

Kim, B.; Shah, S.; Park, H.-S.; Hong, Y.-C.; Ha, M.; Kim, Y.; Kim, B.-N.; Kim, Y.; Ha, E.-H. Adverse Effects of Prenatal Mercury Exposure on Neurodevelopment during the First 3 Years of Life Modified by Early Growth Velocity and Prenatal Maternal Folate Level. *Environmental Research* **2020**, *191*, 109909, doi:10.1016/j.envres.2020.109909.

Steenweg–de Graaff, J.; Roza, S.J.; Steegers, E.A.; Hofman, A.; Verhulst, F.C.; Jaddoe, V.W.; Tiemeier, H. Maternal Folate Status in Early Pregnancy and Child Emotional and Behavioral Problems: The Generation R Study. *The American Journal of Clinical Nutrition* **2012**, *95*, 1413–1421, doi:10.3945/ajcn.111.030791.

Burns, J.M.; Baghurst, P.A.; Sawyer, M.G.; McMichael, A.J.; Tong, S.L. Lifetime Low-Level Exposure to Environmental Lead and Children's Emotional and Behavioral Development at Ages 11-13 Years. The Port Pirie Cohort Study. *Am J Epidemiol* **1999**, *149*, 740–749, doi:10.1093/oxfordjournals.aje.a009883.

Plusquellec, P.; Muckle, G.; Dewailly, E.; Ayotte, P.; Bégin, G.; Desrosiers, C.; Després, C.; Saint-Amour, D.; Poitras, K. The Relation of Environmental Contaminants Exposure to Behavioral Indicators in Inuit Preschoolers in Arctic Quebec. *NeuroToxicology* **2010**, *31*, 17–25, doi:10.1016/j.neuro.2009.10.008.

Liu, J.; Liu, X.; Wang, W.; McCauley, L.; Pinto-Martin, J.; Wang, Y.; Li, L.; Yan, C.; Rogan, W.J. Blood Lead Concentrations and Children's Behavioral and Emotional Problems: A Cohort Study. *JAMA Pediatrics* **2014**, *168*, 737–745, doi:10.1001/jamapediatrics.2014.332.

Fruh, V.; Rifas-Shiman, S.L.; Amarasiriwardena, C.; Cardenas, A.; Bellinger, D.C.; Wise, L.A.; White, R.F.; Wright, R.O.; Oken, E.; Claus Henn, B. Prenatal Lead Exposure and Childhood Executive Function and Behavioral Difficulties in Project Viva. *NeuroToxicology* **2019**, *75*, 105–115, doi:10.1016/j.neuro.2019.09.006.

Steuerwald, U.; Weihe, P.; Jørgensen, P.J.; Bjerve, K.; Brock, J.; Heinzow, B.; Budtz-Jørgensen, E.; Grandjean, P. Maternal Seafood Diet, Methylmercury Exposure, and Neonatal Neurologic Function. *The Journal of Pediatrics* **2000**, *136*, 599–605, doi:10.1067/mpd.2000.102774.

Oken, E.; Wright, R.O.; Kleinman, K.P.; Bellinger, D.; Amarasiriwardena, C.J.; Hu, H.; Rich-Edwards, J.W.; Gillman, M.W. Maternal Fish Consumption, Hair Mercury, and Infant Cognition in a U.S. Cohort. *Environmental Health Perspectives* **2005**, *113*, 1376–1380, doi:10.1289/ehp.8041.

Jedrychowski, W.; Jankowski, J.; Flak, E.; Skarupa, A.; Mroz, E.; Sochacka–Tatara, E.; Lisowska–Miszczyk, I.; Szpanowska–Wohn, A.; Rauh, V.; Skolicki, Z.; et al. Effects of Prenatal Exposure to Mercury on Cognitive and Psychomotor Function in One-Year-Old Infants: Epidemiologic Cohort Study in Poland. *Annals of Epidemiology* **2006**, *16*, 439–447, doi:10.1016/j.annepidem.2005.06.059.

Suzuki, K.; Nakai, K.; Sugawara, T.; Nakamura, T.; Ohba, T.; Shimada, M.; Hosokawa, T.; Okamura, K.; Sakai, T.; Kurokawa, N.; et al. Neurobehavioral Effects of Prenatal Exposure to Methylmercury and PCBs, and Seafood Intake: Neonatal Behavioral Assessment Scale Results of Tohoku Study of Child Development. *Environmental Research* **2010**, *110*, 699–704, doi:10.1016/j.envres.2010.07.001.

Llop, S.; Guxens, M.; Murcia, M.; Lertxundi, A.; Ramon, R.; Riano, I.; Rebagliato, M.; Ibarluzea, J.; Tardon, A.; Sunyer, J.; et al. Prenatal Exposure to Mercury and Infant Neurodevelopment in a Multicenter Cohort in Spain: Study of Potential Modifiers. *American Journal of Epidemiology* **2012**, *175*, 451–465, doi:10.1093/aje/kwr328.

Lingler, J.H.; Schmidt, K.L.; Gentry, A.L.; Hu, L.; Terhorst, L.A. A New Measure of Research Participant Burden: Brief Report. *Journal of Empirical Research on Human Research Ethics* **2014**, *9*, 46–49, doi:10.1177/1556264614545037.

Suriname Water Company *Resultaten SWM Waterkwaliteit Na de Zuivering 2019*; Paramaribo, Suriname, 2019;

Ohiagu, F.O.; Lele, K.C.; Chikezie, P.C.; Verla, A.W.; Enyoh, C.E. Bioaccumulation and Health Risk Assessment of Heavy Metals in Musa Paradisiaca, Zea Mays, Cucumeropsis Manii and Manihot Esculenta Cultivated in Onne, Rivers State, Nigeria. *Environ Anal Health Toxicol* **2020**, *35*, e2020011, doi:10.5620/eaht.e2020011.

Rimbaud, D.; Restrepo, M.; Louison, A.; Boukhari, R.; Ardillon, V.; Carles, G.; Lambert, V.; Jolivet, A. Blood Lead Levels and Risk Factors for Lead Exposure among Pregnant Women in Western French Guiana: The Role of Manioc Consumption. *Journal of Toxicology and Environmental Health, Part A* **2017**, *80*, 382–393, doi:10.1080/15287394.2017.1331490.

Ministry of Agriculture Animal Husbandry and Fisheries *Mancozeb Import Statistics*; Paramaribo, Suriname, 2019;

Abdoel Wahid, F.; Wickliffe, J.; Wilson, M.; Van Sauers, A.; Bond, N.; Hawkins, W.; Mans, D.; Lichtveld, M. Presence of Pesticide Residues on Produce Cultivated in Suriname. *Environ Monit Assess* **2017**, *189*, 303, doi:10.1007/s10661-017-6009-0.

Gullino, M.L.; Tinivella, F.; Garibaldi, A.; Kemmitt, G.M.; Bacci, L.; Sheppard, B. Mancozeb: Past, Present, and Future. *Plant Disease* **2010**, *94*, 1076–1087, doi:10.1094/PDIS-94-9-1076.

Ouboter, P.E.; Landburg, G.A.; Quik, J.H.M.; Mol, J.H.A.; van der Lugt, F. Mercury Levels in Pristine and Gold Mining Impacted Aquatic Ecosystems of Suriname, South America. *AMBIO* **2012**, *41*, 873–882, doi:10.1007/s13280-012-0299-9.

Wickliffe, J.K.; Lichtveld, M.Y.; Zijlmans, C.W.; MacDonald-Ottevanger, S.; Shafer, M.; Dahman, C.; Harville, E.W.; Drury, S.; Landburg, G.; Ouboter, P. Exposure to Total and Methylmercury among Pregnant Women in Suriname: Sources and Public Health Implications. *J Expo Sci Environ Epidemiol* **2021**, *31*, 117–125, doi:10.1038/s41370-020-0233-3.

Heemskerk, M.; Negulic, E.; Duijves, C. *Reducing the Use and Release of Mercury by Artisanal and Small- Scale Gold Miners in Suriname; Social Solutions: Paramaribo, Suriname*; Social Solutions, 2016;

Strategic Environmental Advice; Tropenbos International Suriname *Report of the Strategic Environmental and Social Assessment (SESA) Accompanying the Development of the National REDD+ Strategy of the Republic of Suriname*; 2017;

General Bureau of Statistics 9th Environment Statistics Publication 2015-2019 2020.

Ng, S.; Lin; Jeng; Hwang; Hsieh; Chen Cord Blood Mercury, APOE and Child's Behaviour. *Occup Environ Med* **2013**, *70*, A77.1-A77, doi:10.1136/oemed-2013-101717.227.

van Baar, A.L.; Steenis, L.J.P.; Verhoeven, M.; Hessen, D.J. *Bayley-III-NL Technische Handleiding*; Universiteit Utrecht, The Netherlands: Pearson Assessment and Information BV, 2014;

Bayley, N. *Bayley Scales of Infant and Toddler Development*; Psychological Corporation, Pearson, 2006;

World Bank World Development Indicators.

Drury, S.S.; Mabile, E.; Brett, Z.H.; Esteves, K.; Jones, E.; Shirtcliff, E.A.; Theall, K.P. The Association of Telomere Length With Family Violence and Disruption. *Pediatrics* **2014**, *134*, e128–e137, doi:10.1542/peds.2013-3415.

Cawthon, R.M. Telomere Length Measurement by a Novel Monochrome Multiplex Quantitative PCR Method. *Nucleic Acids Research* **2009**, *37*, e21–e21, doi:10.1093/nar/gkn1027.

McLester-Davis, L.W.Y.; Shankar, A.; Kataria, L.A.; Hidalgo, A.G.; van Eer, E.D.; Koendjbiharie, A.P.; Ramjatan, R.; Hatch, V.I.; Middleton, M.A.; Zijlmans, C.W.R.; et al. Validity, Reliability, and Transcultural Adaptations of the Bayley Scales of Infant and Toddler Development (BSID-III-NL) for Children in Suriname. *Early Human Development* **2021**, *160*, 105416, doi:10.1016/j.earlhumdev.2021.105416.

Shah-Kulkarni, S.; Lee, S.; Jeong, K.S.; Hong, Y.-C.; Park, H.; Ha, M.; Kim, Y.; Ha, E.-H. Prenatal Exposure to Mixtures of Heavy Metals and Neurodevelopment in Infants at 6 Months. *Environmental Research* **2020**, *182*, 109122, doi:10.1016/j.envres.2020.109122.

Ouboter, P.E.; Landburg, G.; Satnarain, G.U.; Starke, S.Y.; Nanden, I.; Simon-Friedt, B.; Hawkins, W.B.; Taylor, R.; Lichtveld, M.Y.; Harville, E.; et al. Mercury Levels in Women and Children from Interior Villages in Suriname, South America. *International Journal of Environmental Research and Public Health* **2018**, *15*, 1007, doi:10.3390/ijerph15051007.

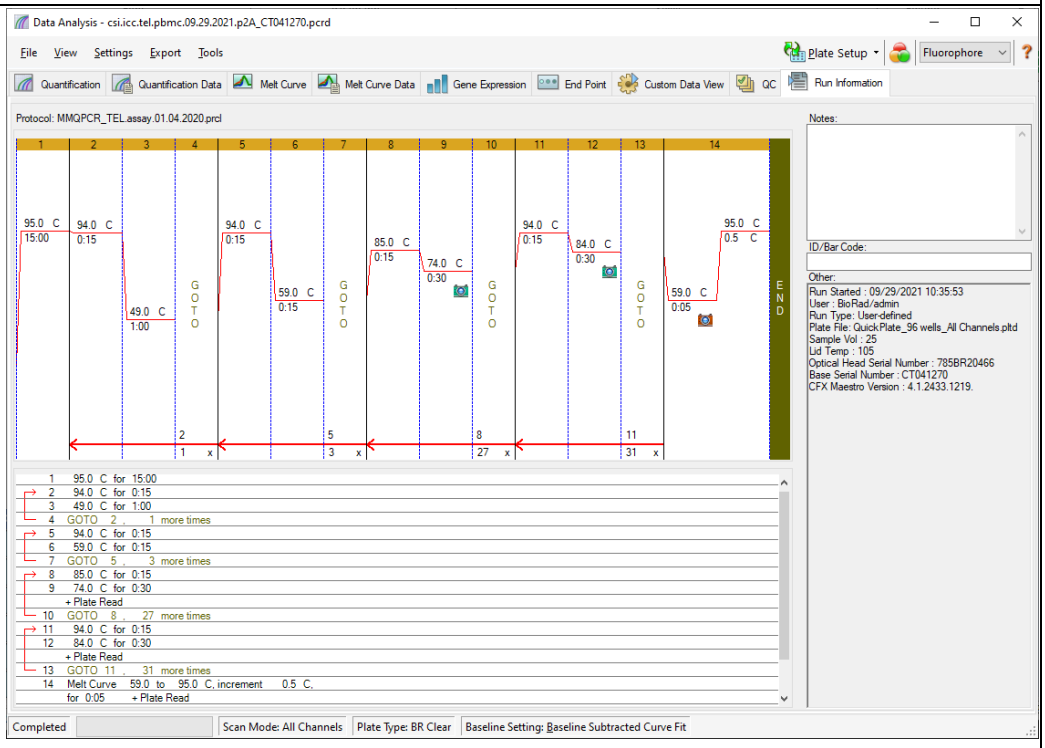
Hassankhan, M.S.; Roopnarine, L.; White, C.; Mahase, R. *Legacy of Slavery and Indentured Labour: Historical and Contemporary Issues in Suriname and the Caribbean*; Routledge, 2016; ISBN 9781351986724.

Zijlmans, W.; Wickliffe, J.; Hindori-Mohangoo, A.; MacDonald-Ottevanger, S.; Ouboter, P.; Landburg, G.; Codrington, J.; Roosblad, J.; Baldewsingh, G.; Ramjatan, R.; et al. Caribbean Consortium for Research in Environmental and Occupational Health (CCREOH) Cohort Study: Influences of Complex Environmental Exposures on Maternal and Child Health in Suriname. *BMJ Open* **2020**, *10*, e034702, doi:10.1136/bmjopen-2019-034702.

Supplemental Methods. Telomere Length Methods

ITEM	DESCRIPTION
Sample Type, Storage, Extraction, and Integrity	
Sample type	DNA extracted from buccal swabs.

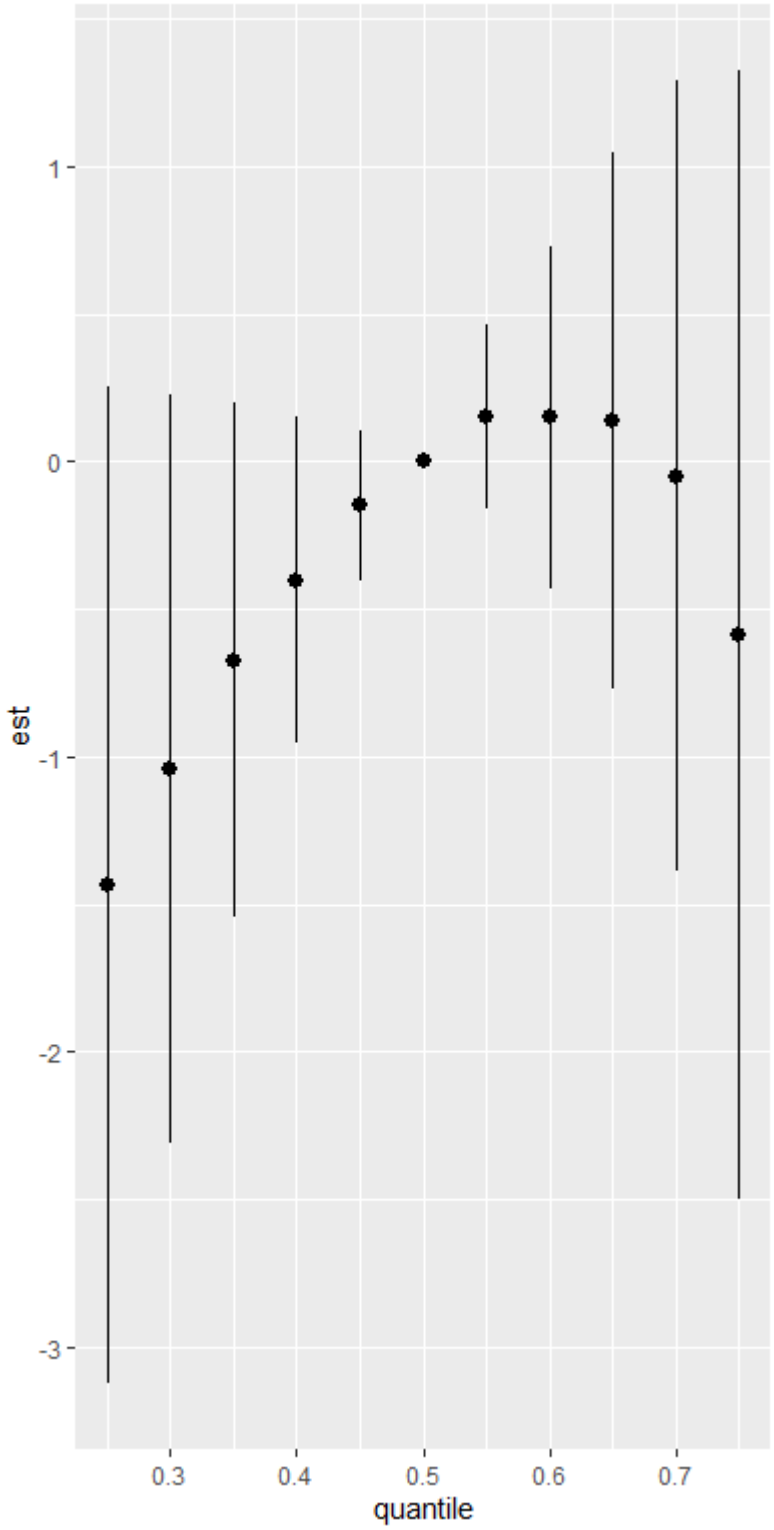
ITEM	DESCRIPTION
Sample storage conditions	The maximum time between sample collection and DNA extraction was 1 year. Swabs were stored at 4°C with a desiccant prior to DNA extraction.
DNA extraction method	DNA was extracted using the QIAamp DNA Mini Kit according to the manufacturer's instructions.
DNA storage conditions, including freeze-thaw cycles	DNA was stored at -80°C prior to telomere length analysis for up to six months. On average there were 3 freeze-thaws for DNA samples between extraction and the MMqPCR assay.
Method of documenting DNA quality and integrity	260/280 and 260/230 ratios were quantified via spectrophotometer for all samples. dsDNA concentration and quality were quantified for all samples with the Qubit assay. No exclusionary criteria were imposed prior to assays.
Percentage of samples specifically tested for DNA quality and integrity	All samples were subjected to quality control via evaluation of 260/280 ratio, 260/230 ratio, and quantity of dsDNA.
MMqPCR Assay	
Method (qPCR, MMqPCR, aTL, etc.)	MMqPCR assays to calculate relative telomere length were structured such that each paired sample (DNA samples from the same individual rom whole blood and from cerebrospinal fluid) was analyzed on the sample plate. Each plate quantified telomere content (T) and single copy gene content (S) using the single copy gene *Albumin.* The ratio of T to S replicates was used to assess relative telomere length. Each run hosted triplicate reactions of 24 samples, 1 standard, and 1 no template control on 96 well plates. A total of 49 MMqPCR assays were performed from September of 2020 to August of 2022 for analysis of all samples.
PCR machine type	CFX96 Touch Real-Time PCR Detection System with 96-well Bio-Rad PCR plates
Source of master mix and reagents, and final reaction volume	The final reaction mix contains 0.75x SYBR Green (Thermo Fisher), 0.9uM forward-reverse telomere primer pair (Integrated DNA Technologies), 0.6uM forward-reverse single copy gene primer pair (Integrated DNA Technologies), 1-10x Gold Buffer (Thermo Fisher), 0.8mM dNTPs, 10mM $MgCl_2$ (Thermo Fisher), 3mM DTT (Research Product International), 1M Betaine (Thermo Fisher), 2.5U/uL AmpliTaq Gold (Thermo Fisher), 1235.2uL PCR grade H_20 (Thermo Fisher), and DNA samples in a 25uL reaction.

<table>
<tr><th>ITEM</th><th>DESCRIPTION</th></tr>
<tr><td>Telomere and single copy gene name, primer sequences and concentration</td><td>
<table>
<tr><th>According to Cawthon 2002:</th><th>Custom Ordered Sequence:</th></tr>
<tr><td colspan="2">Primer specifications include: quantity of XX</td></tr>
<tr><td>Telomere Forward Primer (telg)</td><td>5’ – ACA CTA AGG TTT GGG TTT GGG TTT GGG TTT GGG TTA GTG T – 3’</td></tr>
<tr><td>Telomere Reverse Primer (telc)</td><td>5’ – TGT TAG GTA TCC CTA TCC CTA TCC CTA TCC CTA TCC CTA ACA – 3’</td></tr>
<tr><td>Single Copy Gene Forward Primer (albd2)</td><td>5’ – GCG GGC CCG CGT GGC GGA GCG AGG CCG GAA AAG CAT GGT CGC CT</td></tr>
<tr><td>Single Copy Gene Reverse Primer (albu2)</td><td>5’ – GCC TCG CTC CGG GAG CGC CGC GCG GCC AAA TGC TGC ACA GAA TC</td></tr>
</table>
Albumin
</td></tr>
<tr><td>Full PCR program description including temperature, times, and cycle numbers</td><td> 
</td></tr>
<tr><td>PCR efficiency of single copy gene and telomere primers</td><td>Telomere: $R^2 = 0.999$; Efficiency = 93.65%
Albumin: $R^2 = 0.996$; Efficiency = 95.16%</td></tr>
<tr><td>Source and concentration of control samples and standard curve</td><td>The control sample was comprised of pooled DNA extracted from buccal swabs. The standard curve was run in triplicate in each duplicate plate and consisted of a 2-fold serial dilution of this control sample, ranging from 2e-3 ng/uL to 3.13e-5 ng/uL of DNA across 7 standards, with a no template control.</td></tr>
<tr><td colspan="2">Data Analysis</td></tr>
<tr><td>Mean and standard deviation or</td><td>Average of 2.08, standard deviation of 0.69, and median of 2.07.</td></tr>
</table>

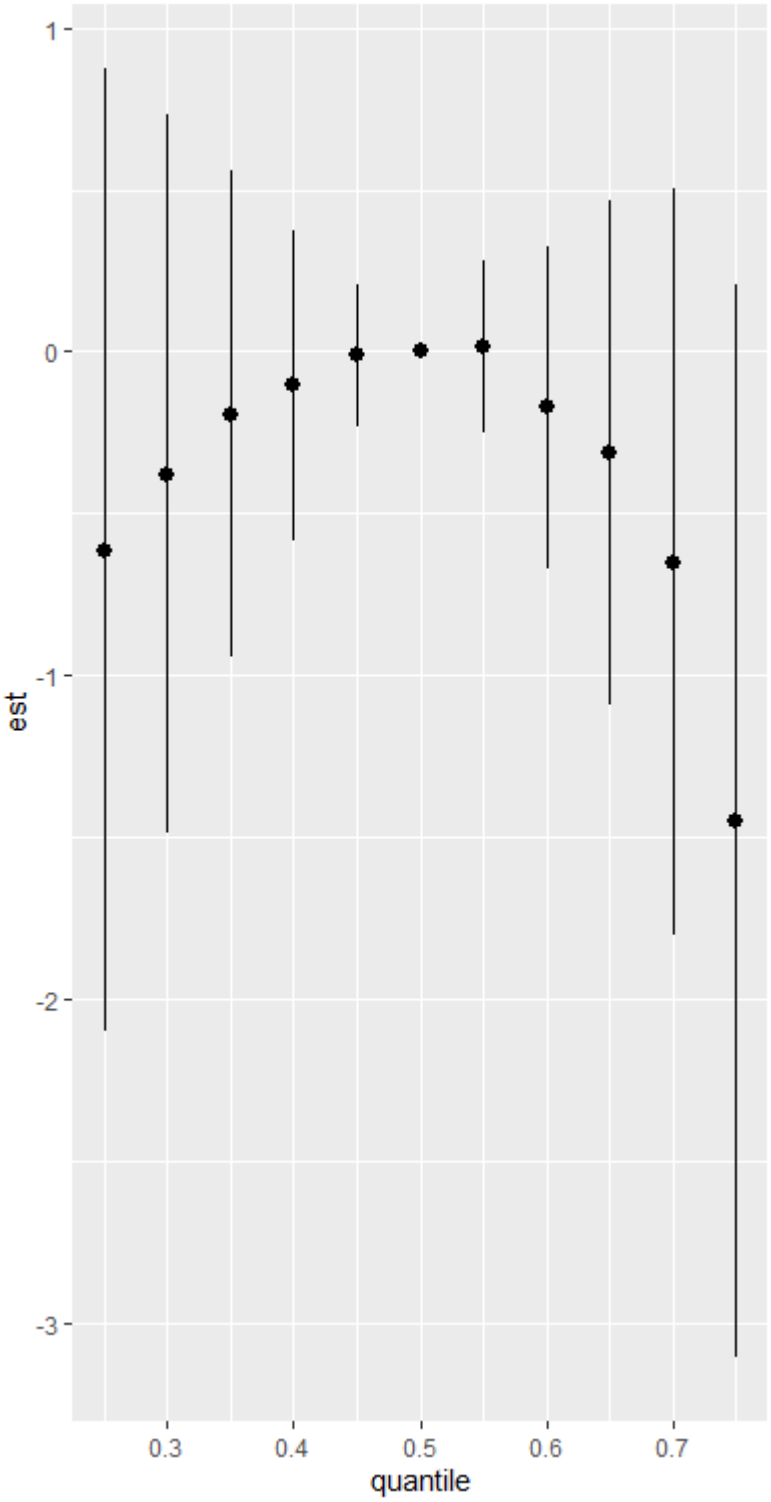
ITEM	DESCRIPTION
median range of telomere lengths	
Number of sample replicates	Each sample was assessed for T and S during a single run across two PCR plates, with three replicates within each plate for a total of 6 measurements. If the sample did not pass quality control criteria described below it was run up to 3 times.
Level of independence of replicates	Replicates were drawn from the same DNA aliquot (i.e., the same tube), with duplicates run in a different well position on a second plate.
Analytic method, considering replicate measurements, to determine final length	Each plate quantified telomere content (T) and single copy gene content (S) using the single copy gene *Albumin.* The ratio of T to S replicates was used to assess relative telomere length.
Method of accounting for variation between replicates	When the coefficient of variation across triplicate estimates of telomere content or genome copy number was greater than 10%, up to one replicate was removed to increase the reproducibility of the data. If 5 replicates could not be obtained for a sample, the sample was rerun. Coefficients of variations were 3.79% for within triplicate variation on average and 2.09% for between duplicate plate variation on average.
Method of accounting for well position effects within plates	Replicates were drawn from the same DNA aliquot (i.e., the same tube), with duplicates run in a different well position on a second plate.
Method of accounting for between plate effects	Coefficients of variations were 3.79% for within triplicate variation on average and 2.09% for between duplicate plate variation on average.
% of samples repeated and % of samples failing QC and excluding from further analyses	Only 20% of samples were repeated once for a minimum of 5 replicates with passing quality control criteria.
Acceptable range of PCR efficiency for single copy gene and	90% – 110% (*within 10% variation of each other*)

ITEM	DESCRIPTION
telomere primers	
ICCs of samples/study groups to address variability	The intra-class correlation (ICC) for all samples passing quality control criteria was calculated as 0.959 (CI: 0.955, 0.963) in accordance with previously described computations (Verhulst, 2020).
T/S ratio transformed to a z-score prior before comparison across methods/studies	Not applicable. No comparison across studies was conducted.
How samples nested within families were accounted for	Samples from the same individual were run on the same plate.

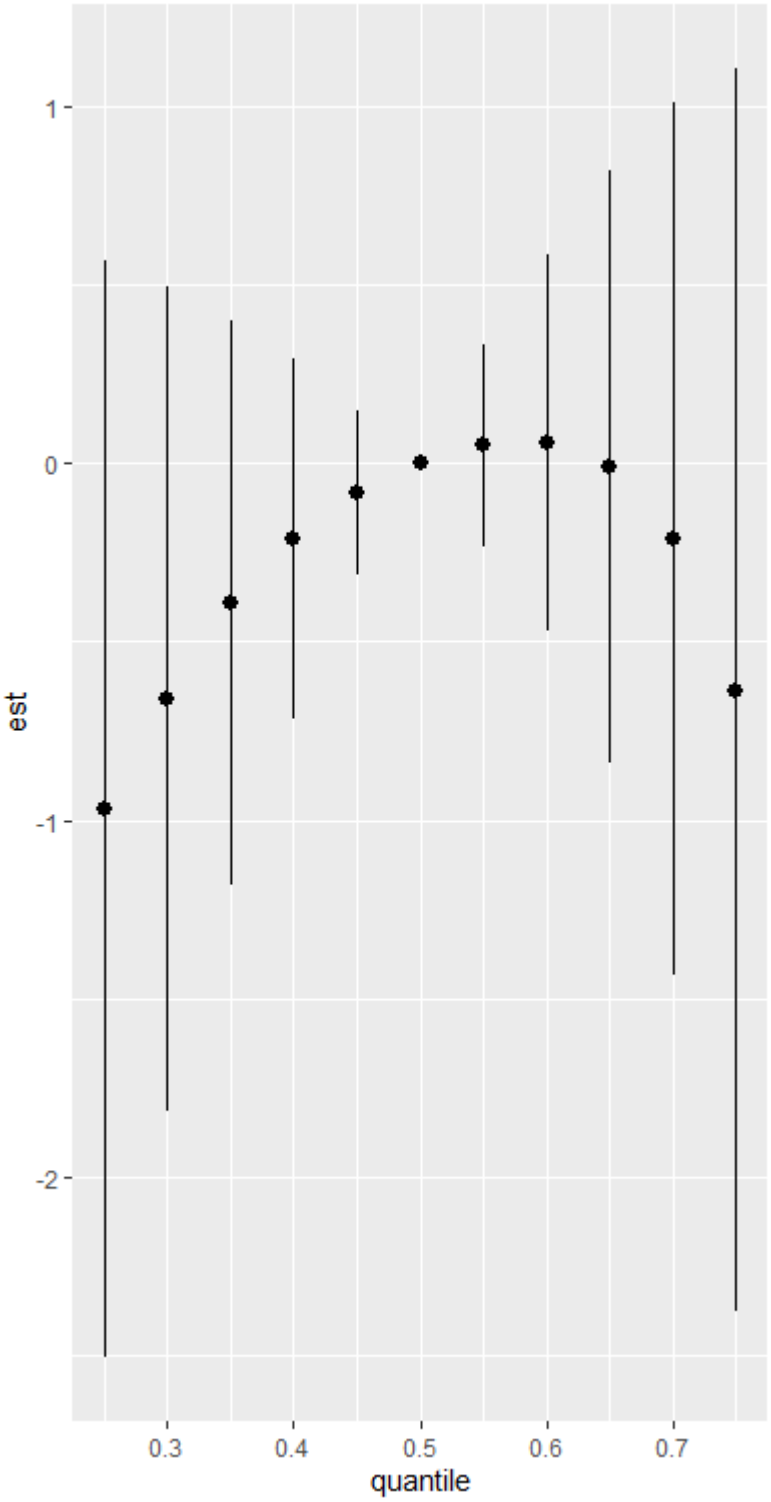
Supplemental Findings. Bayesian Kernel Machine Regression Analyses with BSID-III-NL Cognitive, Language, and Motor Composite Scores



Supplemental Figure 1. BKMR Overall Mixture Results for Cognitive Development.



Supplemental Figure 1. BKMR Overall Mixture Results for Language Development.



Supplemental Figure 1. BKMR Overall Mixture Results for Motor Development.

CHAPTER 6

Telomere length predicts neuropsychological performance in middle-aged and older participants of underrepresented populations from the University of Wisconsin Alzheimer's Disease Research Center

Background:

Telomeres are the protective DNA and protein structures at the ends of human chromosomes which maintain genomic stability. The shortening of telomeres, due to lagging strand synthesis, exposures to genotoxins, and other mechanisms, is an intrinsic part of aging. As such, the measurement of telomere length (TL) is an increasingly popular biomarker of aging. Many biomarkers have been assessed for their ability to predict Alzheimer's disease and related dementias (ADRD) diagnosis and prognosis, including amyloid and levels in cerebrospinal fluid. Unfortunately, no one biomarker to date has been able to comprehensibly capture the abnormal aging in cognitive, memory, visuo-spatial, and verbal domains seen in ADRD (Burns et al., 2002; Mattson, 2004).

The most significant predictor of ADRD is chronological age, where the risk of developing ADRD increases as a person becomes older. While chronological age is not modifiable, biological age can be influenced, as biological age considers environmental exposures, socioeconomic status, and other factors that contribute to the functional decline observed in aging. Behavioral interventions to slow biological aging, as indexed by changes in TL, have proven successful in short term analysis (Buttet et al., 2022). Altering one's biological age, therefore, is a potential therapeutic yet to be investigated longitudinally for prevention of ADRD.

Advanced biological age compared to chronological age has been reported in historically underrepresented populations in research, including American Indian(s) / Alaska Native(s)

(AI/AN) and Black(s) / African American(s) (B/AA; Geronimus, 1992; Elliott et al., 2021). Despite their increased risk for developing ADRD compared to Whites in the United States and their accelerated biological age, AI/AN and B/AA are not proportionately represented in Alzheimer's Disease Research Center (ADRC) cohorts (Mayeda et al., 2016; National Institute on Aging, 2022). The gap in ADRD research from ADRC studies for AI/AN and B/AA is partially fueled by the combination of inadequate representation in both clinical trials and longitudinal prospective studies as well as inadequate culturally responsive educational outreach and engagement (Olson & Albensi, 2020; Gilmore-Bykovskyi et al., 2019). While racial minorities are less likely to participate in research on prevention of cognitive decline, the University of Wisconsin ADRC has dedicated efforts to enrolling individuals from historically unrepresented populations in ADRD research (Zhou et al., 2017; Gleason et al., 2019).

There is a growing interest in biomarkers as predictors of cognitive decline and ADRD that extend beyond parental history and specific genetic markers. For example, a University of Wisconsin ADRC study found that Apolipoprotein E (APOE) ε4 allele positivity was not associated with cognitive decline, but elevations in CSF biomarkers, specifically amyloid and tau, were predicted of cognitive decline (Johnson et al., 2018; Byun et al., 2020). Multiple lines of research support consideration of TL as a measurable predictor of cognitive decline progressing to ADRD. When measured from peripheral blood, TL has been associated with mild cognitive impairment (MCI) and dementia, with meta-analytic associations to Alzheimer's disease (Insel et al., 2012; Grodstein et al., 2008; Kume et al., 2012; Hochstrasser et al., 2012; Liu et al., 2016; Forero et al., 2016; Scarabino et al., 2017; Scarabino et al., 2020). However, discrepancies exist with specific need for additional research with inclusion of AI/AN and B/AA

populations (Thomas et al., 2008; Lukens et al., 2009; Movérare-Skrtic et al., 2012; Mahoney et al., 2019).

Participant Demographics	n	Percent	Mean	Standard Deviation
Race and Ethnicity	n = 188; 0 missing			
AI/AN	24	12.77%		
B/AA	149	79.26%		
B/AA and White and Asian	1	0.53%		
AI/AN and B/AA	3	1.59%		
AI/AN and White	8	4.26%		
AI/AN and B/AA and White	2	1.06%		
AI/AN and White and Hispanic	1	0.53%		
APOE Genotype	n = 155; 33 missing			
APOE ε2 and ε2	1	0.65%		
APOE ε2 and ε3	20	12.90%		
APOE ε3 and ε3	64	41.29%		
APOE ε2 and ε4	4	2.58%		
APOE ε3 and ε4	61	39.35%		
APOE ε4 and ε4	5	3.23%		
Sex	n = 188; 0 missing			
Female	141	75.00%		
Male	47	25.00%		
Age	n = 188; 0 missing		60.88 yrs	9.06 yrs
Height	n = 188; 0 missing		166.40 cm	8.74 cm
Weight	n = 188; 0 missing		91.02 kg	21.89 kg
Systolic Blood Pressure	n = 188; 0 missing		199.12 mmHg	206.04 mmHg
Diastolic Blood Pressure	n = 188; 0 missing		155.61 mmHg	221.13 mmHg
Hypertension Diagnosis	n = 188; 0 missing			
Normal Blood Pressure	23	12.23%		
Elevated Blood Pressure	43	22.87%		
High Blood Pressure (Stage 1)	63	33.51%		
High Blood Pressure (Stage 2)	59	31.39%		
Pulse	n = 109; 79 missing		71.17 bpm	11.72 bpm
Blood Glucose	n = 101; 87 missing		112.90 mg/dL	42.46 mg/dL
Blood Total Cholesterol	n = 188; 0 missing		186.78 mg/dL	38.03 mg/dL
Hypercholesterolemia Diagnosis	n = 188; 0 missing			
Normal Total Cholesterol Level	125	66.49%		
Borderline High Total Cholesterol Level	41	21.81%		
High Total Cholesterol Level	22	11.70%		
Blood Triglycerides	n = 108; 80 missing		110.05 mg/dL	68.49 mg/dL
Hypertriglyceridemia Diagnosis	n = 108; 80 missing			
Normal Triglycerides	89	82.41%		
Borderline High Triglycerides	10	9.26%		
High Triglycerides	9	8.33%		
Blood High-Density Lipoprotein Cholesterol	n = 108; 80 missing		58.44 mg/dL	17.69 mg/dL

Blood Low-Density Lipoprotein Cholesterol	n = 108; 80 missing		106.50 mg/dL	32.51 mg/dL
Waist Circumference	n = 117; 71 missing		103.51 cm	19.52 cm
Hip Circumference	n = 117; 71 missing		112.90 cm	18.19 cm
Waist to Hip Ratio	n = 117; 71 missing		0.92	0.09
Cognitive Status	n = 188; 0 missing			
Normal Cognitive Status	170	90.43%		
Impaired Cognitive Status	12	6.38%		
Mild Cognitive Impairment	6	3.19%		
Family History of Dementia Diagnosis	n = 157; 31 missing			
Lack of Family History (no)	55	35.03%		
Family History Present (yes)	102	64.97%		
Years of Formal Education	n = 188; 0 missing		11.60 yrs	4.07 yrs
Handedness	n = 188; 0 missing			
Left Hand Dominant	20	10.64%		
Ambidextrous	1	0.53%		
Right Hand Dominant	167	88.83%		
Rey Auditory Verbal Learning Test Sum of Trials 1-5	n = 188; 0 missing		44.19	9.25
Trail Making Test Trail A Time	n = 188; 0 missing		33.68 sec	12.61 sec
Trail Making Test Trail B Time	n = 188; 0 missing		104.36 sec	59.26 sec

Table 1. Participant Demographics. Abbreviations correspond to the following: AI/AN = American Indian / Alaska Native; B/AA = Black / African American; APOE = Apolipoprotein E; yrs = years; cm = centimeters; kg = kilograms; mmHg = millimeters of mercury; bpm = beats per minute; mg/dL = milligrams per deciliter; sec = seconds.

Previous research has associated longer TL with better scores on an assessment of executive function, the Trail Making Test (TMT) B, in a mixed sample of healthy White and B/AA individuals (Leibel et al., 2020). However, in an Asian population, shorter TL was associated with better memory functioning on the Rey Auditory Verbal Learning Test (RAVLT) in individuals with MCI (Yu et al., 2020). Research studies assessing the utility of TL to predict cognitive performance on neuropsychological assessments like the TMT and RAVLT have not assessed the relationship for AI/AN populations. Given the need for an enhanced understanding of the relation between neuropsychological performance, ADRD, and TL in historically underrepresented populations, this study assessed the correlation between TL and executive functioning and verbal learning, tasks involving brain regions affected by ADRD, in University of Wisconsin ADRC middle-aged and older AI/AN and B/AA participants.

Methods:

Participants and Study Descriptions

Data were collected from two longitudinal cohort studies of the University of Wisconsin ADRC: 1) the Wisconsin Registry for Alzheimer's Prevention (WRAP) study and 2) the University of Wisconsin ADRC clinical core. The WRAP is a longitudinal observational study of middle-aged adults who were cognitively normal at study entry, and approximately 70% of the cohort has a parental history of ADRD (Clark et al., 2018). The University of Wisconsin ADRC clinical core comprises participants at various ADRD stages ranging from cognitively asymptomatic to having ADRD. In sum, the present investigation included 188 participants who provided whole blood, concurrently completed RAVLT and TMT assessments, and provided relevant covariates including sex, race and ethnicity, age, years of formal education, height, weight, handedness, hypertension, hypercholesterolemia, diabetes, family history of ADRD, APOE genotype, and cognitive status. The full demographics are reported in Table 1. The University of Wisconsin Institutional Review Board approved all study procedures, each participant provided signed informed consent before participation, and all research was completed in accordance with the Helsinki Declaration as previously described (Johnson et al., 2018). The data that support these findings were provided by the University of Wisconsin in a de-identified dataset except for the TL results.

DNA Extraction and TL Measurement

Whole blood samples were collected in 10mL EDTA tubes, mixed by rocking for 5 minutes, and aliquoted into 5mL tubes before being stored at -20oC. Samples were thawed and genomic DNA was extracted at the University of Wisconsin using the Gentra Puregene Blood Kit following the manufacturer’s protocol (Qiagen). Approximately 200ng of DNA per sample

was aliquoted, labeled with a de-identified subject ID, and sent to Tulane University on dry ice overnight for DNA quality assessment and TL analysis. Tulane University's Institutional Review Board assess the TL measurement for these de-identified DNA samples and deemed the research exempt.

Integrity and purity of genomic DNA samples was assessed via BioTek Epoch Microplate Spectrophotometer and with the Invitrogen Qubit dsDNA Broad Range Assay Kit. The average 260/280 ratio for all DNA samples was 1.91, the average 260/230 ratio was 0.99, and the average double-stranded DNA concentration was 149.16 nanogram per microliter (ng/µl). The average relative TL was determined from the telomere repeat copy number to single gene (albumin) copy number (T/S) ratio using an adapted monochrome multiplex quantitate real-time polymerase chain reaction (MMQ-PCR) via a BioRad CFX96 as previously described (Cawthon, 2009; Drury et al., 2014). Samples were performed in triplicate with different well positions on the duplicate plate, using a 7-point standard curve from a peripheral white blood cell DNA standard ranging from 0.0313ng to 2ng, with an average of 1.25 replicates of the standard removed per plate. The average efficiencies of telomere and sing copy gene primers were 94.34% and 95.52% respectively, with an average R^2 of 0.99 for telomere and single copy gene standard replicates. All efficiencies were between 90-110% and between 10% of each other to eliminate plate to plate variability. The average slope and y-intercept for the telomere standard curve were -3.47 and 4.93, respectively. The average slope and y-intercept for the single copy gene standard curve were -3.44 and 12.88, respectively. Coefficients of variations were 4.10% for within triplicate variation on average and 1.90% for between duplicate plate variation on average. Only 13% of samples were repeated once for a minimum of 5 replicates with passing quality control criteria. The intra-class correlation (ICC) for all samples passing quality control

criteria was calculated as 0.887 (CI: 0.862, 0.908) in accordance with previously described computations (Verhulst, 2020).

Neuropsychological Assessments

At WRAP and University of Wisconsin ADRC clinical core visits, both TMT Trials A and B and the RAVLT were administered. TMT Trails A and B consist of 25 circles on paper containing numbers (1-25 in A) or numbers and letters (1-13 and A-L in B). Individuals were instructed to connect the numbers in ascending order in A, and in an ascending pattern alternating numbers and letters in B starting with numbers. The TMT A is an attentional task, while the TMT B is an executive functioning task, with both measured in the amount of time the individual takes to complete the test (Corrigan & Hinkeldey, 1987). While the normative time cutoff for the TMT B is 300 seconds, the cutoff was extended to 600 seconds to allow for greater variability in task performance (Strauss et al., 2006). The RAVLT evaluates verbal learning where individuals were presented with a list of 15 unrelated words to learn and immediately recall aloud over 5 learning trials. After each word list presentation, participants were instructed to recall as many words as possible with each correctly recalled word counting for one point.

Data Analysis

Statistical analyses were conducted with IBM Corp's Statistical Package for the Social Sciences (SPSS) version 27.0 and R Core Team's R version 4.0.3. Bivariate correlations of TL and RAVLT of Trials 1-5, TMT A time, and TMT B time were conducted with Pearson's R. Multivariable regression analyses examined main effects and multivariable interactions of TL, RAVLT of Trials 1-5, TMT A time, and TMT B time. All analyses investigated main effects and covariates including race and ethnicity, age, sex, dominant handedness, height, weight,

hypertension, hypercholesteremia, and cognitive status. The significance level was set to a

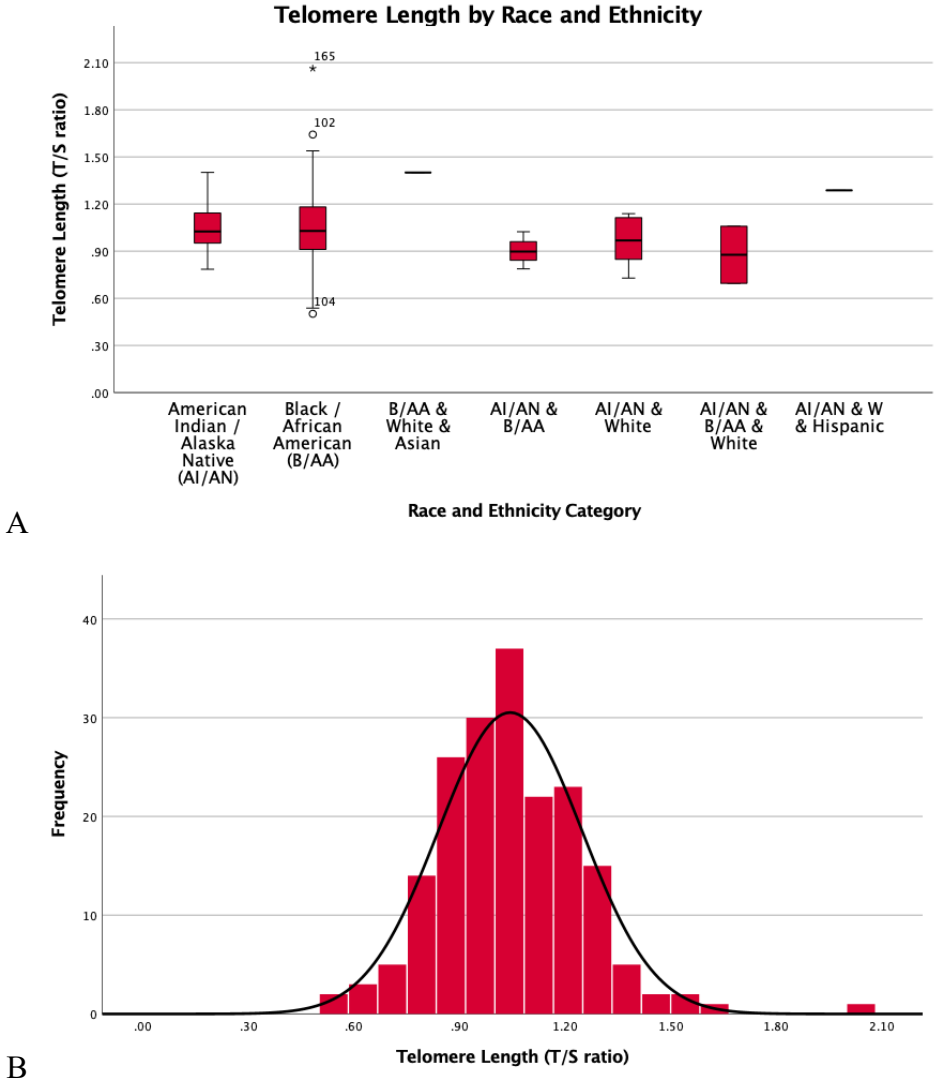


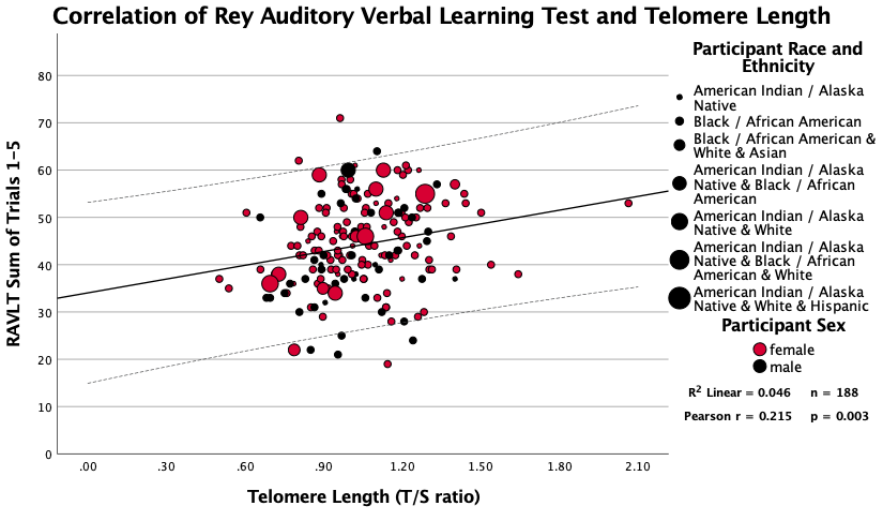
Figure 1. Box plot of telomere length for each race and ethnicity category (A) and the distribution of telomere length for the sample (B).

probability value less than 0.05 for all analyses.

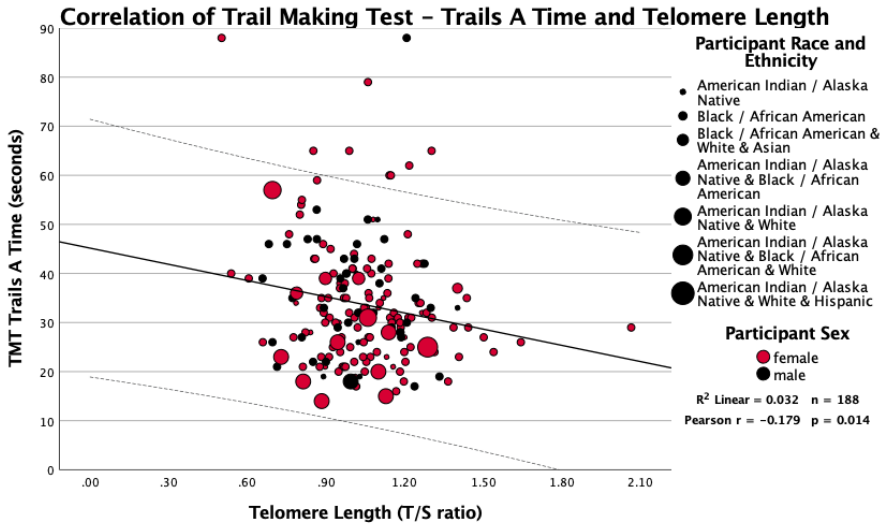
Results:

Description of Participants Demographics and TL Distribution

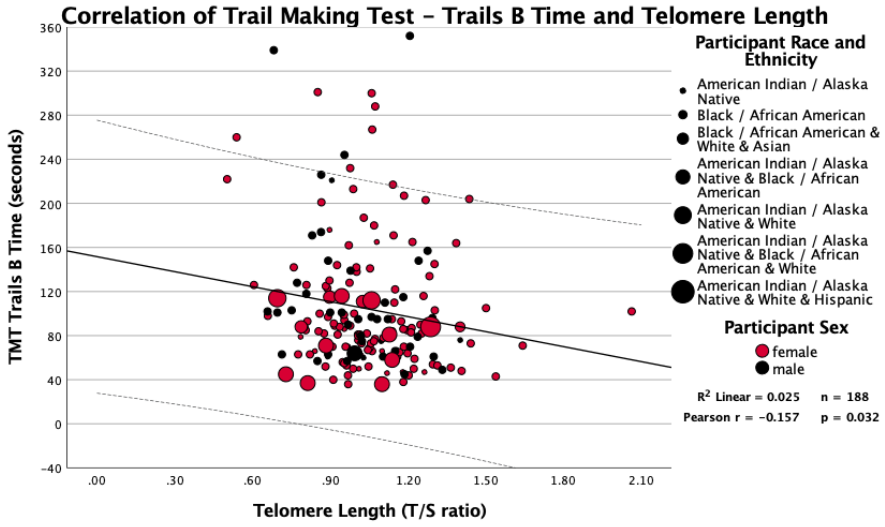
The sample was majority female (75%) with an average age of 60.88 years (σ = 9.06 years). Less than 8% of participants reported more than one racial and ethnic identity, with 12.77% reporting only AI/AN and the remaining 79.26% reporting only B/AA. The average height was



A



B



C

Figure 2. Scatter plots with linear fit lines of telomere length with the Rey Auditory Verbal Learning Test sum of Trials 1-5 (A), the Trails Making Test Trails A time (B) and the Trails Making Test Trails B time (C). Larger data points represent low frequency racial and ethnic categories.

166.40 centimeters (cm; σ = 8.74 cm), the average weight was 91.02 kilograms (kg; σ = 21.89 kg), and the average waist to hip ratio was 0.92 (σ = 0.09). The majority of participants had elevated (22.87%) or high blood pressure (64.89%), with an average pulse of 71.17 beats per minute (bpm; σ = 11.72 bpm). The average blood glucose level was 112.90 milligrams per deciliter (mg/dL; σ = 42.46 mg/dL), the majority of participants had normal total cholesterol blood levels (66.49%), and normal triglycerides blood levels (82.41%). The majority genotype for Apolipoprotein E was two ε3 alleles, followed by one ε3 allele and one ε4 allele, where 45.16% of the total sample has at least one ε4 allele. However, the majority of participants had a positive family history of dementia (64.97%). Few were diagnosed with MCI (3.19%) at the time of neuropsychological assessment, with 90.43% presenting cognitively normal. The majority of participants were right-handed (88.83%) with an average of 11.60 years (σ = 4.07 years) of formal education. The average number of words recalled across RAVLT Trials 1-5 was 44.19 (σ = 9.25). The average time of participants for completing the TMT Trail A was 33.68 seconds (σ = 12.61 seconds) and for the TMT Trail B the average time to completion was 104.36 seconds (σ = 59.26 seconds).

There were no significant differences between race and ethnic groups for overall mean TL per category, though some categories only contained one individual (Figure 1A). The distribution of TL for the sample resembled a normal Gaussian distribution (Figure 1B) with a skewness of 0.71 and a kurtosis of 2.89. The average TL was 1.04 (σ = 0.21) with a minimum of 0.50 to a maximum of 2.06.

Neuropsychological Assessment and TL Correlations

The RAVLT sum of Trials 1-5 was significantly positively correlated with TL, where increases in the number of words recalled across RAVLT Trials 1-5 correlated to longer

telomeres (Figure 1A). Furthermore, the TMT was significantly correlated with TL for both Trails A and B times (Figure 1B and 1C). Less time spent on completing both Trails A and B was correlated with longer telomeres.

Model Summary	Outcome Variable	Predictor	β	p-value
$R^2 = 0.120$ $F(10,177) = 3.540$ $p < 0.001$	Rey Auditory Verbal Learning Test Sum of Trials 1-5	Telomere length	0.210	0.005
		Sex	-0.240	0.016
		Cognitive status	-0.207	0.004
$R^2 = 0.113$ $F(10,177) = 3.3777$ $p < 0.001$	Trail Making Test Trail A Time (seconds)	Telomere length	-0.158	0.033
		Weight	-0.155	0.037
		Cognitive status	0.188	0.009
$R^2 = 0.042$ $F(7,180) = 2.157$ $p = 0.040$	Trail Making Test Trail B Time (seconds)	Telomere length	-0.175	0.017
Race, sex, weight, height, and cognitive status included in all models				

Table 2. Multivariable Linear Regression Models with Significant Predictors' Betas and Probability Values. The coefficients of determination listed are adjusted values for the total variables included in the model.

Regression Analysis

Multivariable linear regression was used to test if telomere length, race and ethnicity, age, sex, dominant handedness, height, weight, hypertension, hypercholesteremia, and cognitive status significantly predicted the RAVLT sum score of Trials 1-5 (Table 2). The overall regression was statistically significant ($R^2 = 0.120$, $F(10, 177) = 3.540$, $p < 0.001$) and telomere length significantly predicted RAVLT sum score of Trials 1-5 ($\beta = 0.210$, $p = 0.005$). Race and ethnicity, age, height, weight, hypertension, and hypercholesteremia did not significantly predict RAVLT sum score of Trials 1-5 ($p > 0.05$). However, sex ($\beta = -0.240$, $p = 0.016$), dominant handedness ($\beta = 0.155$, $p = 0.027$), and cognitive status ($\beta = -0.207$, $p = 0.004$) also significantly predicted RAVLT sum score of Trials 1-5.

Additionally, multiple linear regression was used to test if telomere length, race and ethnicity, age, sex, dominant handedness, height, weight, hypertension, hypercholesteremia, and cognitive status significantly predicted the TMT Trails A time (Table 2). The overall regression

was statistically significant ($R^2 = 0.113$, $F(10, 177) = 3.3777$, $p < 0.001$) and telomere length significantly predicted TMT Trails A time ($\beta = -0.158$, $p = 0.033$). Race and ethnicity, age, sex, dominant handedness, height, hypertension, and hypercholesteremia did not significantly predict TMT Trails A time ($p > 0.05$). However, weight ($\beta = -0.155$, $p = 0.037$) and cognitive status ($\beta = 0.188$, $p = 0.009$) also significantly predicted TMT Trails A time.

Furthermore, multiple linear regression was used to test if telomere length, race and ethnicity, sex, dominant handedness, height, weight, and cognitive status significantly predicted the TMT Trails B time (Table 2). The overall regression was statistically significant ($R^2 = 0.042$, $F(7, 180) = 2.157$, $p = 0.040$) and telomere length significantly predicted TMT Trails B time ($\beta = -0.175$, $p = 0.017$). Race and ethnicity, sex, dominant handedness, height, weight, and cognitive status did not significantly predict TMT Trails B time ($p > 0.05$).

Cognitive Status, Genotype, and TL

Lastly, there were no significant differences in overall mean TL by cognitive status group (Figure 3A). Similarly, there were no significant difference for the 153 participants with APOE genotype data for overall mean TL by allelic category (Figure 3B). The data for the 33 participants missing APOE genotype is not random, as samples from these participants are currently being genotyped in accordance with Wisconsin ADRC procedures (Johnson et al., 2018).

Discussion:

These findings suggest that TL is a strong predictor of cognitive performance for historically underrepresented populations. Here, TL significantly predicted TMT Trails A and B time in this sample of historically underrepresented populations in ADRD research, consistent with previous findings with different racial and ethnic characteristics (Leibel et al., 2020). The

average times to complete the TMT Trails A and B were longer here than previously reported for a younger B/AA and White sample, which confirms the ability of the TMT to detect age-related changes in executive functioning. While performance on the RAVLT is affected by conditions like ADRD and RAVLT scores have been significantly associated with TL before, here TL also significantly predicted performance on the RAVLT regardless of cognitive status (Mitrushina et al., 1989; Yu et al., 2020). As the first report to elucidate the relation of TL and executive function and memory for AI/AN, our findings here are specific for middle-aged and older adults. Future research on this relationship across the life course is warranted, especially if longitudinal analysis of interventions to slow biological aging and prevent ADRD are implemented at different life stages.

At birth, significant differences in average TL are seen between B/AA and White newborns given disparities in prenatal exposures, where B/AA female infants are born with the longest telomeres (McLester-Davis et al., 2019). However, TL did not differ across race and ethnicity categories for this middle-aged and older adult population. These findings align with results supporting the weathering hypothesis, where the rate of TL shortening is accelerated in historically underrepresented populations in research because of environmental and socioeconomic exposures specific to the structural disadvantages faced by these populations in the United States (Geronimus, 1992; Elliot et al., 2021). Therefore, the lack of differences in TL by race and ethnicity category reported here is expected given the age of this sample (mean of 60.88 years) where social determinants of health throughout the lifespan have eliminated differences in TL across racial and ethnic identities. These findings align with previous reports on the utility of TL to indirectly capture cumulative life experiences contributing to aging and age-related diseases like ADRD (Leibel et al., 2020; Chen et al., 2022). These findings also

suggest that the success of interventions to slow biological aging may be dependent on the life stage of participants, such as before ADRD pathology begins or following exposures that accelerate biological aging early in life (Buttet et al., 2022). Future research should therefore identify differences in these behavioral and socioeconomic interventions by age group.

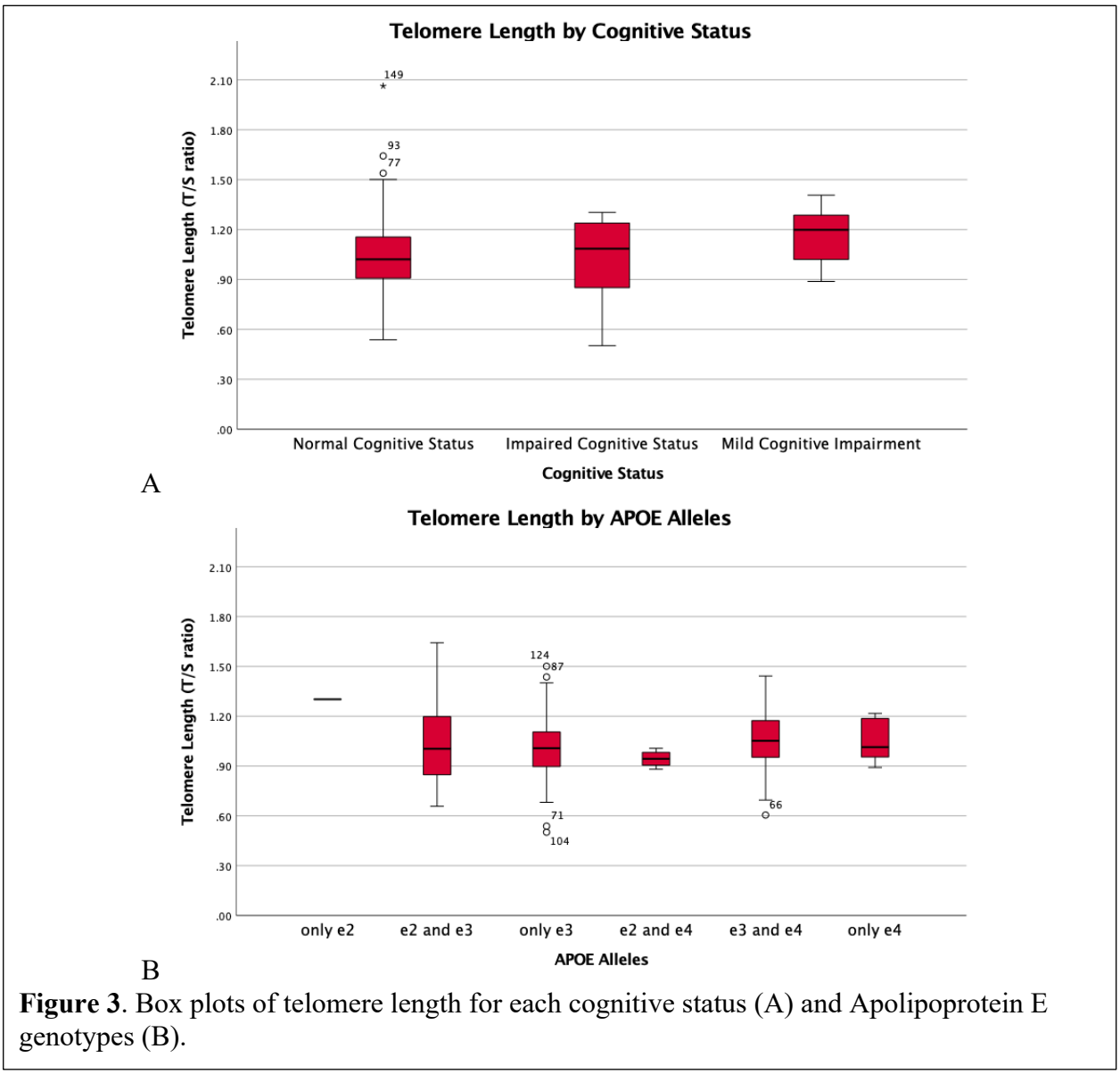


Figure 3. Box plots of telomere length for each cognitive status (A) and Apolipoprotein E genotypes (B).

Similarly, future corroboration of these findings for males is warranted given the unequitable male population analyzed here (25%). It is unsurprising that this sample is majority female as women are more likely to participate in ADRD research, in part because women live

longer on average than men and are therefore more likely to develop ADRD (Martinkova et al., 2021; Olson & Albensi, 2020). However, the equitable utility of interventions to slow biological aging and prevent ADRD should be assessed for all sex and gender identities (Buttet et al., 2022).

While previous cross-sectional reports have identified differences in TL and cognitive status, where those with cognitive impairment have shorter TLs, no significant difference in mean TL based on cognitive status was found in the present study possibly due to differences in the racial composition of this sample (Honig et al., 2006; Scarabino et al., 2017). Alternatively, this is perhaps due to most participants in this sample having normal cognitive status, and future research should clarify the relationship of TL and cognitive status with longitudinal analysis of those progressing to ADRD. Similarly, there were no significant differences in TL given APOE genotype despite previous positive reports (Honig et al., 2006; Takata et al., 2012; Mahoney et al., 2019). While one of these studies involved B/AA participants, previous TL comparisons by APOE genotype have not included AI/AN participants. Importantly, the association of APOE genotype with ADRD risk in AI/AN populations is not as significant as reported in non-Hispanic White populations (Henderson et al., 2002). Therefore, if TL is truly capturing cognitive decline and risk for developing ADRD, previously reported associations with APOE genotype may not hold true in diverse populations samples like the one in this report.

These novel findings are the first to provide information for AI/AN on the relationship of biological aging and neuropsychological performance specific to their population. The inclusion of AI/AN participants was possible through culturally appropriate community-based research practices of the University of Wisconsin ADRC's Native American Outreach team (Gleason et al., 2019). Future ADRC research should hire AI/AN to the research team, identify culturally

appropriate community-based research practices for the AI/AN communities in the ADRC area, and increase the ethical representation of AI/AN participants, communities, and Tribal Nations in ADRD research especially at the investigator level.

References:

Burns A, Byrne EJ, Maurer K. Alzheimer's disease. The Lancet 2002;360:163–5. https://doi.org/10.1016/S0140-6736(02)09420-5.

Buttet M, Bagheri R, Ugbolue UC, Laporte C, Trousselard M, Benson A, et al. Effect of a lifestyle intervention on telomere length: A systematic review and meta-analysis. Mechanisms of Ageing and Development 2022;206:111694. https://doi.org/10.1016/j.mad.2022.111694.

Byun MS, Yi D, Lee JH, Lee J, Lee H, Kim YK, et al. Telomere length, brain tau deposition and cognitive decline: Two-year follow-up study: Biomarkers (non-neuroimaging) / Novel biomarkers. Alzheimer's & Dementia 2020;16. https://doi.org/10.1002/alz.045919.

Cawthon RM. Telomere length measurement by a novel monochrome multiplex quantitative PCR method. Nucleic Acids Research 2009;37:e21–e21. https://doi.org/10.1093/nar/gkn1027.

Chen R, Weuve J, Misra S, Cuevas A, Kubzansky LD, Williams DR. Racial Disparities in Cognitive Function Among Middle-Aged and Older Adults: The Roles of Cumulative Stress Exposures Across the Life Course. The Journals of Gerontology: Series A 2022;77:357–64. https://doi.org/10.1093/gerona/glab099.

Clark LR, Berman SE, Norton D, Koscik RL, Jonaitis E, Blennow K, et al. Age-accelerated cognitive decline in asymptomatic adults with CSF β-amyloid. Neurology 2018;90:e1306–15. https://doi.org/10.1212/WNL.0000000000005291.

Corrigan JD, Hinkeldey NS. Relationships between Parts A and B of the Trail Making Test. J Clin Psychol 1987;43:402–9. https://doi.org/10.1002/1097-4679(198707)43:4<402::AID-JCLP2270430411>3.0.CO;2-E.

Drury SS, Mabile E, Brett ZH, Esteves K, Jones E, Shirtcliff EA, et al. The Association of Telomere Length With Family Violence and Disruption. Pediatrics 2014;134:e128–37. https://doi.org/10.1542/peds.2013-3415.

Elliott ML, Caspi A, Houts RM, Ambler A, Broadbent JM, Hancox RJ, et al. Disparities in the pace of biological aging among midlife adults of the same chronological age have implications for future frailty risk and policy. Nat Aging 2021;1:295–308. https://doi.org/10.1038/s43587-021-00044-4.

Forero DA, González-Giraldo Y, López-Quintero C, Castro-Vega LJ, Barreto GE, Perry G. Meta-analysis of Telomere Length in Alzheimer's Disease. GERONA 2016;71:1069–73. https://doi.org/10.1093/gerona/glw053.

Geronimus AT. The weathering hypothesis and the health of African-American women and infants: evidence and speculations. Ethn Dis 1992;2:207–21.

Gilmore-Bykovskyi AL, Jin Y, Gleason C, Flowers-Benton S, Block LM, Dilworth-Anderson P, et al. Recruitment and retention of underrepresented populations in Alzheimer's disease research: A systematic review. Alzheimer's & Dementia: Translational Research & Clinical Interventions 2019;5:751–70. https://doi.org/10.1016/j.trci.2019.09.018.

Gleason CE, Martin W, Strong L, Summers M, Lambrou NH, Zuelsdorff M, et al. F4-06-02: IMPROVING DEMENTIA OUTCOMES IN INDIAN COUNTRY: THE ONEIDA NATION ALZHEIMER'S DISEASE PROJECT. Alzheimer's & Dementia 2019;15:P1227–P1227. https://doi.org/10.1016/j.jalz.2019.06.4734.

Grodstein F, Oijen M van, Irizarry MC, Rosas HD, Hyman BT, Growdon JH, et al. Shorter Telomeres May Mark Early Risk of Dementia: Preliminary Analysis of 62 Participants from the Nurses' Health Study. PLOS ONE 2008;3:e1590. https://doi.org/10.1371/journal.pone.0001590.

Henderson JN, Crook R, Crook J, Hardy J, Onstead L, Carson-Henderson L, et al. Apolipoprotein E4 and tau allele frequencies among Choctaw Indians. Neuroscience Letters 2002;324:77–9. https://doi.org/10.1016/S0304-3940(02)00150-7.

Hochstrasser T, Marksteiner J, Humpel C. Telomere length is age-dependent and reduced in monocytes of Alzheimer patients. Experimental Gerontology 2012;47:160–3. https://doi.org/10.1016/j.exger.2011.11.012.

Honig LS, Schupf N, Lee JH, Tang MX, Mayeux R. Shorter telomeres are associated with mortality in those with APOE ϵ4 and dementia. Ann Neurol 2006;60:181–7. https://doi.org/10.1002/ana.20894.

Insel KC, Merkle CJ, Hsiao C-P, Vidrine AN, Montgomery DW. Biomarkers for Cognitive Aging Part I: Telomere Length, Blood Pressure and Cognition Among Individuals with Hypertension. Biological Research For Nursing 2012;14:124–32. https://doi.org/10.1177/1099800411406433.

Johnson SC, Koscik RL, Jonaitis EM, Clark LR, Mueller KD, Berman SE, et al. The Wisconsin Registry for Alzheimer's Prevention: A review of findings and current directions. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 2018;10:130–42. https://doi.org/10.1016/j.dadm.2017.11.007.

Kume K, Kikukawa M, Hanyu H, Takata Y, Umahara T, Sakurai H, et al. Telomere length shortening in patients with dementia with Lewy bodies: Telomere shortening in DLB. European Journal of Neurology 2012;19:905–10. https://doi.org/10.1111/j.1468-1331.2011.03655.x.

Leibel DK, Shaked D, Beatty Moody DL, Liu HB, Weng N, Evans MK, et al. Telomere length and cognitive function: Differential patterns across sociodemographic groups. Neuropsychology 2020;34:186–98. https://doi.org/10.1037/neu0000601.

Letang SK, Lin SS-H, Parmelee PA, McDonough IM. Ethnoracial disparities in cognition are associated with multiple socioeconomic status-stress pathways. Cognitive Research: Principles and Implications 2021;6:64. https://doi.org/10.1186/s41235-021-00329-7.

Liu M, Huo YR, Wang J, Wang C, Liu S, Liu S, et al. Telomere Shortening in Alzheimer's Disease Patients. Ann Clin Lab Sci 2016;46:260–5.

Lukens JN, Van Deerlin V, Clark CM, Xie SX, Johnson FB. Comparisons of telomere lengths in peripheral blood and cerebellum in Alzheimer's disease. Alzheimer's & Dementia 2009;5:463–9. https://doi.org/10.1016/j.jalz.2009.05.666.

Mahoney ER, Dumitrescu L, Seto M, Nudelman KNH, Buckley RF, Gifford KA, et al. Telomere length associations with cognition depend on Alzheimer's disease biomarkers. Alzheimer's & Dementia: Translational Research & Clinical Interventions 2019;5:883–90. https://doi.org/10.1016/j.trci.2019.11.003.

Martinkova J, Quevenco F-C, Karcher H, Ferrari A, Sandset EC, Szoeke C, et al. Proportion of Women and Reporting of Outcomes by Sex in Clinical Trials for Alzheimer Disease: A Systematic Review and Meta-analysis. JAMA Network Open 2021;4:e2124124. https://doi.org/10.1001/jamanetworkopen.2021.24124.

Mattson MP. Pathways towards and away from Alzheimer's disease. Nature 2004;430:631–9. https://doi.org/10.1038/nature02621.

Mayeda ER, Glymour MM, Quesenberry CP, Whitmer RA. Inequalities in dementia incidence between six racial and ethnic groups over 14 years. Alzheimer's & Dementia 2016;12:216–24. https://doi.org/10.1016/j.jalz.2015.12.007.

McLester-Davis LWY, Jones CW, Esteves KC, Drury SS. Racial and sex disparities in newborn telomere length predicted by maternal adverse childhood experiences 2019.

Mitrushina M, Satz P, Van Gorp W. Some putative cognitive precursors in subjects hypothesized to be at-risk for dementia. Archives of Clinical Neuropsychology 1989;4:323–33. https://doi.org/10.1016/0887-6177(89)90023-1.

Movérare-Skrtic S, Johansson P, Mattsson N, Hansson O, Wallin A, Johansson J-O, et al. Leukocyte Telomere Length (LTL) is reduced in stable mild cognitive impairment but low LTL is not associated with conversion to Alzheimer's Disease: A pilot study. Experimental Gerontology 2012;47:179–82. https://doi.org/10.1016/j.exger.2011.12.005.

National Institute on Aging. National Alzheimer's Coordinating Center (NACC) 2022. https://naccdata.org/requesting-data/data-summary/uds.

Olson NL, Albensi BC. Race- and Sex-Based Disparities in Alzheimer's Disease Clinical Trial Enrollment in the United States and Canada: An Indigenous Perspective. ADR 2020;4:325–44. https://doi.org/10.3233/ADR-200214.

Scarabino D, Broggio E, Gambina G, Corbo RM. Leukocyte telomere length in mild cognitive impairment and Alzheimer's disease patients. Experimental Gerontology 2017;98:143–7. https://doi.org/10.1016/j.exger.2017.08.025.

Scarabino D, Peconi M, Broggio E, Gambina G, Maggi E, Armeli F, et al. Relationship between proinflammatory cytokines (Il-1beta, Il-18) and leukocyte telomere length in mild cognitive impairment and Alzheimer's disease. Experimental Gerontology 2020;136:110945. https://doi.org/10.1016/j.exger.2020.110945.

Strauss E, Sherman EMS, Spreen O. A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary. Oxford University Press; 2006.

Takata Y, Kikukawa M, Hanyu H, Koyama S, Shimizu S, Umahara T, et al. Association Between ApoE Phenotypes and Telomere Erosion in Alzheimer's Disease. The Journals of Gerontology: Series A 2012;67A:330–5. https://doi.org/10.1093/gerona/glr185.

Thomas P, O' Callaghan NJ, Fenech M. Telomere length in white blood cells, buccal cells and brain tissue and its variation with ageing and Alzheimer's disease. Mechanisms of Ageing and Development 2008;129:183–90. https://doi.org/10.1016/j.mad.2007.12.004.

Verhulst S. Improving comparability between qPCR-based telomere studies. Mol Ecol Resour 2020;20:11–3. https://doi.org/10.1111/1755-0998.13114.

Yu J, Kanchi MM, Rawtaer I, Feng L, Kumar AP, Kua E-H, et al. The functional and structural connectomes of telomere length and their association with cognition in mild cognitive impairment. Cortex 2020;132:29–40. https://doi.org/10.1016/j.cortex.2020.08.006.

Zhou Y, Elashoff D, Kremen S, Teng E, Karlawish J, Grill JD. African Americans are less likely to enroll in preclinical Alzheimer's disease clinical trials. Alzheimer's & Dementia: Translational Research & Clinical Interventions 2017;3:57–64. https://doi.org/10.1016/j.trci.2016.09.004.

CHAPTER 7

The correlation of telomere length from whole blood and cerebrospinal fluid circulating white blood cells from participants of the University of Wisconsin Alzheimer's Disease Research Center

Short Communication:

Identification of individuals at increased risk for Alzheimer's disease and related dementias (ADRD) through biological measurements has been challenging as there are no biomarkers available that provide 100% predictive insight for disease progression. Telomere length (TL) is an increasingly popular biomarker in ADRD research with recent meta-analytic associations of shorter TL and Alzheimer's disease diagnosis (Forero et al., 2016). As a widely recognized aging biomarker, TL shortens with each cell division, decreases with chronological age, and captures genetic, environmental, psychosocial, and lifestyle factors impacting cellular aging (Zhu and van der Harst, 2014). Unfortunately, associations of TL and ADRD identification and prognosis are inconsistent, where some reports show longer TL associated with ADRD for certain tissue samples selected for TL measurement (Franco et al., 2006; Thomas et al., 2008; Lukens et al., 2009; Mahady et al., 2021).

Cerebrospinal fluid (CSF) has been utilized as a biomarker sample source for ADRD research with crucial results aiding in diagnosis and predicting ADRD progression (Kester et al., 2009; Wallin et al., 2010; Dhiman et al., 2019). CSF is a dynamic mixture containing multiple cell types which provide immune functions, aid in elimination of cellular waste for the central nervous system, and cushion neurons from potential impacts with the skull (Frans et al., 1981; Ransohoff and Engelhardt, 2012). To date, TL, while a biomarker of aging and associated with ADRD, has not been assessed in the cells of CSF. T cells, B cells, monocytes, and dendritic cells

in CSF are naturally positioned to interact more intricately with brain aging and ADRD pathogenesis than their peripheral counterparts, such as those found in whole blood (Pashenkov et al., 2001; de Graaf et al., 2011; Ransohoff and Engelhardt, 2012). Understanding the relationship between TL measured from CSF and TL measured from a peripherally obtainable sample provides information for future study designs researching TL as a ADRD biomarker. In addition to measuring TL from CSF, comparison of CSF biomarkers to peripherally obtainable samples is prioritized in ADRD research given that participants are more likely to consent to whole blood draw than lumbar puncture (LP), the cost of LPs is unreasonable for large population-based studies, and CSF analysis in resource-limited settings is not always feasible (Karlawish et al., 2009; Blazel et al., 2020). With these considerations, this study defines, within an individual, the correlation of TL from whole blood and TL from CSF from participants at the University of Wisconsin Alzheimer's Disease Research Center (ADRC).

The University of Wisconsin ADRC studies concurrently collected CSF via LP and whole blood with participant consent and approval from the University of Wisconsin Institutional Review Board as previously described (Madrid et al., 2018; Clark et al., 2018). The ADRC extracted DNA from CSF and whole blood samples and shipped the de-identified DNA samples on dry ice overnight to Tulane University for telomere analysis. Integrity and purity of DNA samples were assessed via 260/280 and 260/230 ratios and concentration of double-stranded DNA. The average relative TL was determined from the telomere repeat copy number to single gene (albumin) copy number (T/S) ratio using an adapted monochrome multiplex quantitate real-time polymerase chain reaction via a BioRad CFX96 as previously described (Cawthon, 2009; Drury et al., 2014; McLester-Davis et al., 2023a). All samples were performed in triplicate with different well positions on the duplicate plate, using a 7-point standard curve

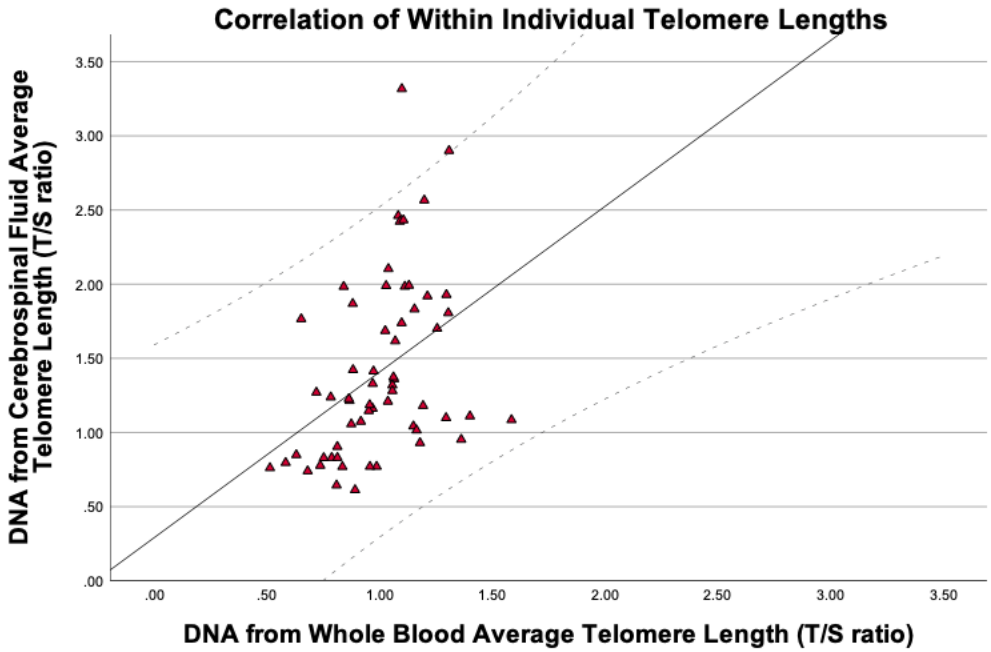


Figure 1. Positive correlation of cerebrospinal fluid telomere length and whole blood telomere length within individuals. Each triangle represents one participant (n = 60) with a solid linear best fit trendline and dashed 95% confidence intervals.

from a peripheral blood mononuclear cell DNA standard ranging from 0.0313ng to 2ng and average efficiencies of telomere and single copy gene primers 93.32% and 98.24%, respectively, with further details of the assay available in Supplementary Table 1. Coefficients of variations were 3.97% for within triplicate variation and 2.88% for between duplicate plate variation. The computed intraclass correlation coefficient for all samples was 0.954 [CI: 0.940, 0.965] as previously described (Verhulst, 2020). Pearson correlations, paired sample T-test, and descriptive statistics for participant demographics were conducted with IBM Corp's SPSS version 27.0 and R Core Team's R version 4.0.3. Significance was set at $\alpha = 0.05$. The ADRC provided a de-identified dataset containing participant sex, race and ethnicity, chronological age, cognitive diagnosis, family history of ADRD, years of formal education, and Apolipoprotein E genotype matching the DNA samples assayed.

There were 60 participants from which CSF and whole blood TLs were measured. This

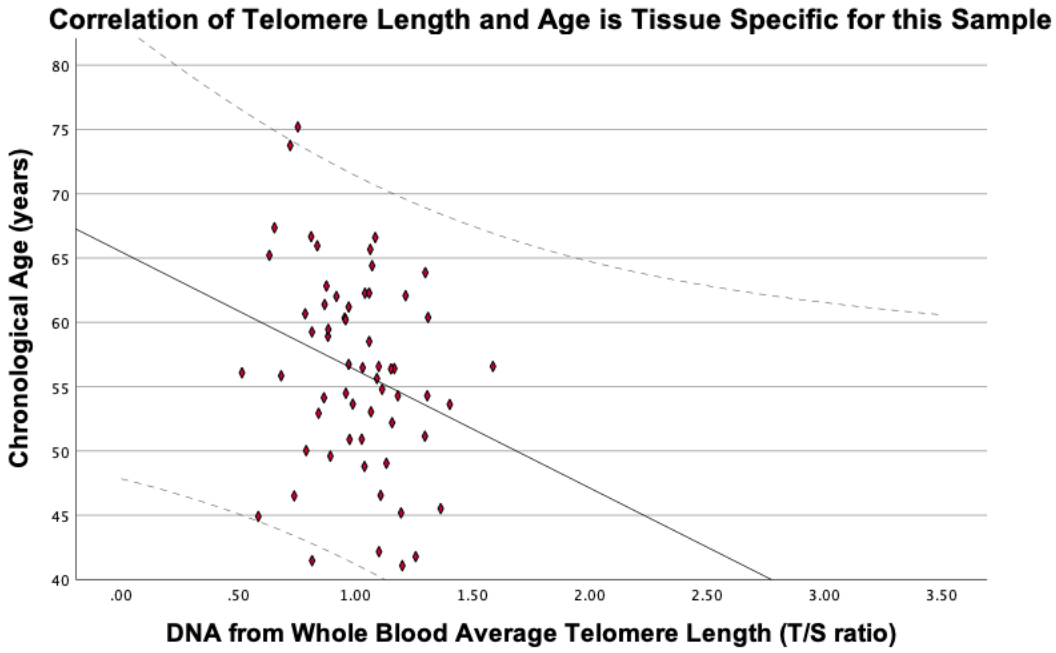


Figure 2. Negative correlation of whole blood telomere length and chronological age. Each diamond represents one participant (n = 60) with a solid linear best fit trendline and dashed 95% confidence intervals.

sample was majority female (61.70%), majority White (96.70%; 3.3% Black), and most were cognitively normal (98.30%; 1.70% Mild Cognitive Impairment). At least 40% had one or more ε4 allele and the majority had a positive family history of ADRD (58.30%). The average age was 56.28 years (σ=0.99 years; range: 41-75) and the average number of formal education years was 11.13 (σ=0.53 years). The whole blood TL values ranged from a minimum of 0.51 to a maximum of 1.59, while the CSF TL values ranged from a minimum of 0.61 to a maximum of 3.32. The paired sample T-test indicated a significant difference between the whole blood TL (mean=1.00, σ=0.03) and CSF TL (mean=1.41, σ=0.08) means ($t(59)=|5.771|$, $p<0.001$). However, whole blood TL and CSF TL within an individual was significantly correlated (Pearson's $r=0.402$, $p=0.001$) as seen in Figure 1 with 95% confidence intervals of the correlation. While this moderate correlation between the two tissues suggests either TL

measurement can be inferred from the other TL measurement, chronological age of the participant was only significantly correlated with whole blood TL (Pearson's r=-0.257, p=0.048) and not CSF TL (Pearson's r=-0.119, p=0.364) in this sample of middle-aged and older participants as seen in Figure 2 with 95% confidence intervals of the correlation.

Similar to other TL tissue correlation reports, these results show that TL measured from CSF can be inferred from TL measured from whole blood for the same individual (Demanelis et al., 2020; McLester-Davis et al., 2023b). However, these findings suggest that there are potential confounding factors producing differences in mean TL for each sample type. For example, the populations of T cells, B cells, and monocytes recruited to the CSF may differ in their replicative history and metabolic activity than their counterparts present in whole blood, which would lead to differences in TL (de Graaf et al., 2011; Croese et al., 2021). Furthermore, CSF has increased numbers of monocytes, neutrophils, and granulocytes when the brain is experiencing an inflammatory process like aging or ADRD, which may be occurring in individuals at risk of ADRD prior to diagnosis, such as this middle-aged and older sample with ε4 alleles and positive family history (Stalder et al., 2005; Lunnon et al., 2012; Ransohoff and Engelhardt, 2012; Roostaei et al., 2021). These changes in cell populations of the CSF would also contribute to differences in mean TL by sample source. Given these limitations of the current study, future research should identify the role of different cell populations in CSF on the correlation between TL measurements in a larger sample size and conduct analysis of this correlation for a sample with majority ADRD diagnosis and the relationship to other CSF biomarkers. Additionally, these results support the use of whole blood TL to capture systematic changes associated with aging and ADRD as peripheral blood cells are exposed to more varied environmental, psychosocial, and lifestyle factors that are captured with TL than those cells specifically recruited to the CSF

(Ayub et al., 2021; Andreu-Sánchez et al., 2022). Overall, TL measured from peripherally obtainable whole blood is reflective of the sample germane to ADRD research, CSF, and these findings support use of peripheral measurements of TL as a biomarker of protective and risk factors in future ADRD studies.

References:

Andreu-Sánchez, S., Aubert, G., Ripoll-Cladellas, A., Henkelman, S., Zhernakova, D.V., Sinha, T., Kurilshikov, A., Cenit, M.C., Jan Bonder, M., Franke, L., Wijmenga, C., Fu, J., van der Wijst, M.G.P., Melé, M., Lansdorp, P., Zhernakova, A., 2022. Genetic, parental and lifestyle factors influence telomere length. Commun Biol 5, 1–14. https://doi.org/10.1038/s42003-022-03521-7

Ayub, M., Jin, H.K., Bae, J., 2021. The blood cerebrospinal fluid barrier orchestrates immunosurveillance, immunoprotection, and immunopathology in the central nervous system. BMB Rep 54, 196–202. https://doi.org/10.5483/BMBRep.2021.54.4.205

Blazel, M.M., Lazar, K.K., Van Hulle, C.A., Ma, Y., Cole, A., Spalitta, A., Davenport-Sis, N., Bendlin, B.B., Wahoske, M., Illingworth, C., Gleason, C.E., Edwards, D.F., Blazel, H., Asthana, S., Johnson, S.C., Carlsson, C.M., 2020. Factors Associated with Lumbar Puncture Participation in Alzheimer's Disease Research. Journal of Alzheimer's Disease 77, 1559–1567. https://doi.org/10.3233/JAD-200394

Cawthon, R.M., 2009. Telomere length measurement by a novel monochrome multiplex quantitative PCR method. Nucleic Acids Research 37, e21–e21. https://doi.org/10.1093/nar/gkn1027

Clark, L.R., Berman, S.E., Norton, D., Koscik, R.L., Jonaitis, E., Blennow, K., Bendlin, B.B., Asthana, S., Johnson, S.C., Zetterberg, H., Carlsson, C.M., 2018. Age-accelerated cognitive decline in asymptomatic adults with CSF β-amyloid. Neurology 90, e1306–e1315. https://doi.org/10.1212/WNL.0000000000005291

Croese, T., Castellani, G., Schwartz, M., 2021. Immune cell compartmentalization for brain surveillance and protection. Nat Immunol 22, 1083–1092. https://doi.org/10.1038/s41590-021-00994-2

de Graaf, M.T., Smitt, P.A.E.S., Luitwieler, R.L., van Velzen, C., van den Broek, P.D.M., Kraan, J., Gratama, J.W., 2011. Central memory CD4+ T cells dominate the normal cerebrospinal fluid. Cytometry 80B, 43–50. https://doi.org/10.1002/cyto.b.20542

Demanelis, K., Jasmine, F., Chen, L.S., Chernoff, M., Tong, L., Delgado, D., Zhang, C., Shinkle, J., Sabarinathan, M., Lin, H., Ramirez, E., Oliva, M., Kim-Hellmuth, S., Stranger, B.E., Lai, T.-P., Aviv, A., Ardlie, K.G., Aguet, F., Ahsan, H., GTEx Consortium, Doherty, J.A., Kibriya, M.G., Pierce, B.L., Aguet, F., Anand, S., Ardlie, K.G., Gabriel, S., Getz, G.A., Graubert, A., Hadley, K., Handsaker, R.E., Huang, K.H., Kashin, S., Li, Xiao, MacArthur, D.G., Meier, S.R., Nedzel, J.L., Nguyen, D.T., Segrè, A.V., Todres, E., Balliu, B., Barbeira, A.N., Battle, A., Bonazzola, R., Brown, A., Brown, C.D., Castel, S.E., Conrad, D.F., Cotter, D.J., Cox, N., Das,

S., de Goede, O.M., Dermitzakis, E.T., Einson, J., Engelhardt, B.E., Eskin, E., Eulalio, T.Y., Ferraro, N.M., Flynn, E.D., Fresard, L., Gamazon, E.R., Garrido-Martín, D., Gay, N.R., Gloudemans, M.J., Guigó, R., Hame, A.R., He, Y., Hoffman, P.J., Hormozdiari, F., Hou, L., Im, H.K., Jo, B., Kasela, S., Kellis, M., Kim-Hellmuth, S., Kwong, A., Lappalainen, T., Li, Xin, Liang, Y., Mangul, S., Mohammadi, P., Montgomery, S.B., Muñoz-Aguirre, M., Nachun, D.C., Nobel, A.B., Oliva, M., Park, YoSon, Park, Yongjin, Parsana, P., Rao, A.S., Reverter, F., Rouhana, J.M., Sabatti, C., Saha, A., Stephens, M., Stranger, B.E., Strober, B.J., Teran, N.A., Viñuela, A., Wang, G., Wen, X., Wright, F., Wucher, V., Zou, Y., Ferreira, P.G., Li, G., Melé, M., Yeger-Lotem, E., Barcus, M.E., Bradbury, D., Krubit, T., McLean, J.A., Qi, L., Robinson, K., Roche, N.V., Smith, A.M., Sobin, L., Tabor, D.E., Undale, A., Bridge, J., Brigham, L.E., Foster, B.A., Gillard, B.M., Hasz, R., Hunter, M., Johns, C., Johnson, M., Karasik, E., Kopen, G., Leinweber, W.F., McDonald, A., Moser, M.T., Myer, K., Ramsey, K.D., Roe, B., Shad, S., Thomas, J.A., Walters, G., Washington, M., Wheeler, J., Jewell, S.D., Rohrer, D.C., Valley, D.R., Davis, D.A., Mash, D.C., Branton, P.A., Barker, L.K., Gardiner, H.M., Mosavel, M., Siminoff, L.A., Flicek, P., Haeussler, M., Juettemann, T., Kent, W.J., Lee, C.M., Powell, C.C., Rosenbloom, K.R., Ruffier, M., Sheppard, D., Taylor, K., Trevanion, S.J., Zerbino, D.R., Abell, N.S., Akey, J., Chen, L., Demanelis, K., Doherty, J.A., Feinberg, A.P., Hansen, K.D., Hickey, P.F., Jasmine, F., Jiang, L., Kaul, R., Kibriya, M.G., Li, J.B., Li, Q., Lin, S., Linder, S.E., Pierce, B.L., Rizzardi, L.F., Skol, A.D., Smith, K.S., Snyder, M., Stamatoyannopoulos, J., Tang, H., Wang, M., Carithers, L.J., Guan, P., Koester, S.E., Little, A.R., Moore, H.M., Nierras, C.R., Rao, A.K., Vaught, J.B., Volpi, S., 2020. Determinants of telomere length across human tissues. Science 369, eaaz6876. https://doi.org/10.1126/science.aaz6876

Dhiman, K., Blennow, K., Zetterberg, H., Martins, R.N., Gupta, V.B., 2019. Cerebrospinal fluid biomarkers for understanding multiple aspects of Alzheimer's disease pathogenesis. Cell. Mol. Life Sci. 76, 1833–1863. https://doi.org/10.1007/s00018-019-03040-5

Drury, S.S., Mabile, E., Brett, Z.H., Esteves, K., Jones, E., Shirtcliff, E.A., Theall, K.P., 2014. The Association of Telomere Length With Family Violence and Disruption. Pediatrics 134, e128–e137. https://doi.org/10.1542/peds.2013-3415

Forero, D.A., González-Giraldo, Y., López-Quintero, C., Castro-Vega, L.J., Barreto, G.E., Perry, G., 2016. Meta-analysis of Telomere Length in Alzheimer's Disease. GERONA 71, 1069–1073. https://doi.org/10.1093/gerona/glw053

Franco, S., Blasco, M.A., Siedlak, S.L., Harris, P.L.R., Moreira, P.I., Perry, G., Smith, M.A., 2006. Telomeres and telomerase in Alzheimer's disease: Epiphenomena or a new focus for therapeutic strategy? Alzheimer's & Dementia 2, 164–168. https://doi.org/10.1016/j.jalz.2006.03.001

Frans, L., Schipholt, L., Stadhouders, A.M., 1981. Techniques a Simple Method for the Collection of Cells from Cerebrospinal Fluid for Electron Microscopy. Ultrastructural Pathology 2, 175–181. https://doi.org/10.3109/01913128109064246

Karlawish, J., Rubright, J., Casarett, D., Cary, M., Ten Have, T., Sankar, P., 2009. Older Adults' Attitudes Toward Enrollment of Non-competent Subjects Participating in Alzheimer's Research. AJP 166, 182–188. https://doi.org/10.1176/appi.ajp.2008.08050645

Kester, M.I., Vlies, A.E. van der, Blankenstein, M.A., Pijnenburg, Y. a. L., Elk, E.J. van, Scheltens, P., Flier, W.M. van der, 2009. CSF biomarkers predict rate of cognitive decline in Alzheimer disease. Neurology 73, 1353–1358. https://doi.org/10.1212/WNL.0b013e3181bd8271

Lukens, J.N., Van Deerlin, V., Clark, C.M., Xie, S.X., Johnson, F.B., 2009. Comparisons of telomere lengths in peripheral blood and cerebellum in Alzheimer's disease. Alzheimer's & Dementia 5, 463–469. https://doi.org/10.1016/j.jalz.2009.05.666

Lunnon, K., Ibrahim, Z., Proitsi, P., Lourdusamy, A., Newhouse, S., Sattlecker, M., Furney, S., Saleem, M., Soininen, H., Kłoszewska, I., Mecocci, P., Tsolaki, M., Vellas, B., Coppola, G., Geschwind, D., Simmons, A., Lovestone, S., Dobson, R., Hodges, A., 2012. Mitochondrial Dysfunction and Immune Activation are Detectable in Early Alzheimer's Disease Blood. JAD 30, 685–710. https://doi.org/10.3233/JAD-2012-111592

Madrid, A., Hogan, K.J., Papale, L.A., Clark, L.R., Asthana, S., Johnson, S.C., Alisch, R.S., 2018. DNA Hypomethylation in Blood Links B3GALT4 and ZADH2 to Alzheimer's Disease. JAD 66, 927–934. https://doi.org/10.3233/JAD-180592

Mahady, L.J., He, B., Malek-Ahmadi, M., Mufson, E.J., 2021. Telomeric alterations in the default mode network during the progression of Alzheimer's disease: Selective vulnerability of the precuneus. Neuropathology Appl Neurobio 47, 428–440. https://doi.org/10.1111/nan.12672

McLester-Davis, L.W.Y., Magruder, M.G., Roy, T.R., Martin, N.A., Hastings, W.J., Verhulst, S., Drury S.S. Precision MMqPCR Telomere Length Measurement Methodology. Journal of Visualized Experiments: Genetics, 2023a, in preparation.

McLester-Davis, L.W.Y., Estrada, P., Kataria, L.A., Martin, N.A., Siebeneicher, J.T., Tristano, R.I., Mayne, C.V., Horlick, R.P., O'Connell, S.S., Drury, S.S. Cross-Tissue Telomere Length Correlations in Healthy Humans: A Systematic Review and Meta-Analysis. Ageing Research Reviews, 2023b, under review.

Pashenkov, M., Huang, Y.-M., Kostulas, V., Haglund, M., Söderström, M., Link, H., 2001. Two subsets of dendritic cells are present in human cerebrospinal fluid. Brain 124, 480–492. https://doi.org/10.1093/brain/124.3.480

Ransohoff, R.M., Engelhardt, B., 2012. The anatomical and cellular basis of immune surveillance in the central nervous system. Nat Rev Immunol 12, 623–635. https://doi.org/10.1038/nri3265

Roostaei, T., Diaconu, C., Touil, H., Harbison, C., Zhang, Y., Epstein, S., Tuddenham, J., Thakur, K., Bryois, J., Wiendl, H., Meyer Zu Hörste, G., Malhotra, D., Riley, C., Menon, V., De Jager, P.L., 2021. Defining the architecture of cerebrospinal fluid cellular communities in neuroinflammatory diseases (preprint). Immunology. https://doi.org/10.1101/2021.11.01.466797

Stalder, A.K., Ermini, F., Bondolfi, L., Krenger, W., Burbach, G.J., Deller, T., Coomaraswamy, J., Staufenbiel, M., Landmann, R., Jucker, M., 2005. Invasion of Hematopoietic Cells into the Brain of Amyloid Precursor Protein Transgenic Mice. J. Neurosci. 25, 11125–11132. https://doi.org/10.1523/JNEUROSCI.2545-05.2005

Thomas, P., O' Callaghan, N.J., Fenech, M., 2008. Telomere length in white blood cells, buccal cells and brain tissue and its variation with ageing and Alzheimer's disease. Mechanisms of Ageing and Development 129, 183–190. https://doi.org/10.1016/j.mad.2007.12.004

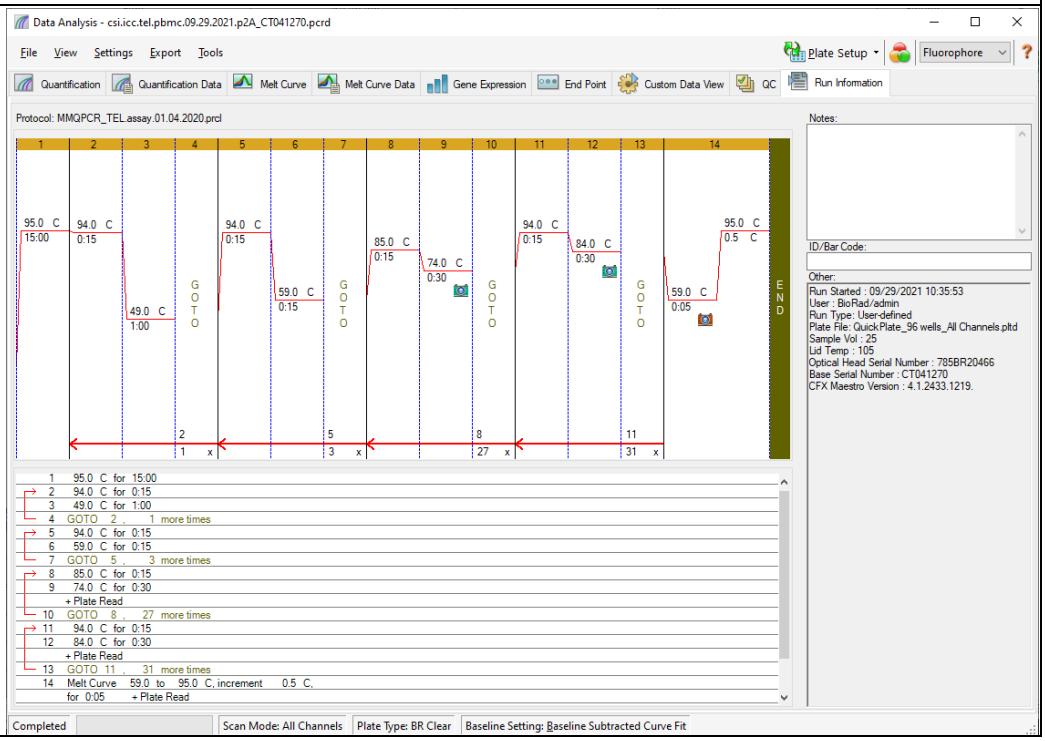
Verhulst, S., 2020. Improving comparability between qPCR-based telomere studies. Mol Ecol Resour 20, 11–13. https://doi.org/10.1111/1755-0998.13114

Wallin, Å.K., Blennow, K., Zetterberg, H., Londos, E., Minthon, L., Hansson, O., 2010. CSF biomarkers predict a more malignant outcome in Alzheimer disease. Neurology 74, 1531–1537. https://doi.org/10.1212/WNL.0b013e3181dd4dd8

Zhu, H., van der Harst, P., 2014. Telomere Biology in Senescence and Aging, in: Inflammation, Advancing Age and Nutrition. Elsevier, pp. 71–84. https://doi.org/10.1016/B978-0-12-397803-5.00007-1

Supplementary Table 1. Telomere Length Measurement Methodology Summary

ITEM	DESCRIPTION
Sample Type, Storage, Extraction, and Integrity	
Sample type	DNA extracted from whole blood samples and cerebrospinal fluid samples were obtained from the University of Wisconsin Alzheimer's Disease Research Center.
Sample storage conditions	The time between sample collection and DNA extraction was 1 year. Blood samples and cerebrospinal fluid pellets were stored at -20°C prior to DNA extraction.
DNA extraction method	DNA was extracted from whole blood using the Gentra Puregene Blood Kit according to the manufacturer's protocol.
DNA storage conditions, including freeze-thaw cycles	DNA was stored at -80°C and shipped on dry ice overnight to Tulane University for telomere length analysis in 3 shipments. On average there were 2 freeze-thaws for DNA samples between extraction and the MMqPCR assay. The first thaw was conducted to confirm receipt, reorganize samples, and quantify DNA with spectrophotometer and Qubit dsDNA assay upon arrival at Tulane University. The second thaw was done to dilute the sample if necessary and aliquot into the MMqPCR assay. No samples were not rerun on the MMqPCR assay. DNA samples were stored at -20°C for an average of 2 months between DNA quantification and the MMqPCR assay.
Method of documenting DNA quality and integrity	260/280 and 260/230 ratios were quantified via spectrophotometer for all samples. dsDNA concentration and quality were quantified for all samples with the Qubit assay. No exclusionary criteria were imposed prior to assays.
Percentage of samples specifically tested for DNA quality and integrity	All samples were subjected to quality control via evaluation of 260/280 ratio, 260/230 ratio, and quantity of dsDNA.
MMqPCR Assay	

<table>
<tr><th>ITEM</th><th>DESCRIPTION</th></tr>
<tr><td>Method (qPCR, MMqPCR, aTL, etc.)</td><td>MMqPCR assays to calculate relative telomere length were structured such that each paired sample (DNA samples from the same individual rom whole blood and from cerebrospinal fluid) was analyzed on the sample plate. Each plate quantified telomere content (T) and single copy gene content (S) using the single copy gene Albumin. The ratio of T to S replicates was used to assess relative telomere length. Each run hosted triplicate reactions of 24 samples, 1 standard, and 1 no template control on 96 well plates. A total of 14 MMqPCR assays were performed from 6/20/2022 to 1/16/2023 for analysis of all samples.</td></tr>
<tr><td>PCR machine type</td><td>CFX96 Touch Real-Time PCR Detection System with 96-well Bio-Rad PCR plates</td></tr>
<tr><td>Source of master mix and reagents, and final reaction volume</td><td>The final reaction mix contains 0.75x SYBR Green (Thermo Fisher), 0.9uM forward-reverse telomere primer pair (Integrated DNA Technologies), 0.6uM forward-reverse single copy gene primer pair (Integrated DNA Technologies), 1-10x Gold Buffer (Thermo Fisher), 0.8mM dNTPs, 10mM $MgCl_2$ (Thermo Fisher), 3mM DTT (Research Product International), 1M Betaine (Thermo Fisher), 2.5U/uL AmpliTaq Gold (Thermo Fisher), 1235.2uL PCR grade H_20 (Thermo Fisher), and DNA samples in a 25uL reaction.</td></tr>
<tr><td>Telomere and single copy gene name, primer sequences and concentration</td><td>
<table>
<tr><th>According to Cawthon 2002:</th><th>Custom Ordered Sequence:</th></tr>
<tr><td colspan="2">Primer specifications include: quantity of XX</td></tr>
<tr><td>Telomere Forward Primer (telg)</td><td>5’ – ACA CTA AGG TTT GGG TTT GGG TTT GGG TTT GGG TTA GTG T – 3’</td></tr>
<tr><td>Telomere Reverse Primer (telc)</td><td>5’ – TGT TAG GTA TCC CTA TCC CTA TCC CTA TCC CTA TCC CTA ACA – 3’</td></tr>
<tr><td>Single Copy Gene Forward Primer (albd2)</td><td>5’ – GCG GGC CCG CGT GGC GGA GCG AGG CCG GAA AAG CAT GGT CGC CT</td></tr>
<tr><td>Single Copy Gene Reverse Primer (albu2)</td><td>5’ – GCC TCG CTC CGG GAG CGC CGC GCG GCC AAA TGC TGC ACA GAA TC</td></tr>
</table>
Albumin</td></tr>
<tr><td>Full PCR program description including temperature, times, and cycle numbers</td><td> 
</td></tr>
</table>

ITEM	DESCRIPTION
PCR efficiency of single copy gene and telomere primers	Telomere: $R^2 = 0.99$; Efficiency = 94.34% Albumin: $R^2 = 0.99$; Efficiency = 95.52%
Source and concentration of control samples and standard curve	The control sample was comprised of pooled DNA extracted peripheral blood mononuclear cell. The standard curve was run in triplicate in each duplicate plate and consisted of a 2-fold serial dilution of this control sample, ranging from 2e-3 ng/uL to 3.13e-5 ng/uL of DNA across 7 standards, with a no template control.
Data Analysis	
Mean and standard deviation or median range of telomere lengths	Average of 1.04, standard deviation of 0.21, minimum of 0.50 and maximum of 2.06.
Number of sample replicates	Each sample was assessed for T and S during a single run across two PCR plates, with three replicates within each plate for a total of 6 measurements. If the sample did not pass quality control criteria described below it was run up to 3 times.
Level of independence of replicates	Replicates were drawn from the same DNA aliquot (i.e., the same tube), with duplicates run in a different well position on a second plate.
Analytic method, considering replicate measurements, to determine final length	Each plate quantified telomere content (T) and single copy gene content (S) using the single copy gene *Albumin.* The ratio of T to S replicates was used to assess relative telomere length.
Method of accounting for variation between replicates	When the coefficient of variation across triplicate estimates of telomere content or genome copy number was greater than 10%, up to one replicate was removed to increase the reproducibility of the data. If 5 replicates could not be obtained for a sample, the sample was rerun. Coefficients of variations were 4.10% for within triplicate variation on average and 1.90% for between duplicate plate variation on average. Only 13% of samples were repeated once for a minimum of 5 replicates with passing quality control criteria.
Method of accounting for well position effects within plates	Replicates were drawn from the same DNA aliquot (i.e., the same tube), with duplicates run in a different well position on a second plate.

ITEM	DESCRIPTION
Method of accounting for between plate effects	Coefficients of variations were 4.10% for within triplicate variation on average and 1.90% for between duplicate plate variation on average.
% of samples repeated and % of samples failing QC and excluding from further analyses	Only 13% of samples were repeated once for a minimum of 5 replicates with passing quality control criteria.
Acceptable range of PCR efficiency for single copy gene and telomere primers	90% – 110% (*within 10% variation of each other*)
ICCs of samples/study groups to address variability	The intra-class correlation (ICC) for all samples passing quality control criteria was calculated as 0.887 (CI: 0.862, 0.908) in accordance with previously described computations (Verhulst, 2020).
T/S ratio transformed to a z-score prior before comparison across methods/studies	Not applicable. No comparison across studies was conducted.
How samples nested within families were accounted for	Samples from the same individual were run on the same plate.

DISCUSSION OF THE DISSERTATION

This dissertation investigated the utility of telomere length (TL), with precision measurement, as a biomarker of neurodevelopment and neuropsychological performance for minoritized populations. The first three chapters address limitations of TL measurement methodology, specifically quantitative polymerase chain reaction (qPCR) based measurement. From these chapters, there is guidance on reporting minimums of qPCR methodology, instructions for conducting the monochrome multiplex qPCR (MMqPCR) method utilized in this dissertation, and the strong ability to infer TL across tissues within an individual with the supplemental correlations for 102 different tissue types. The middle two chapters assess TL as a biomarker of neurodevelopment in the diverse population found in the Caribbean Consortium for Research in Environmental and Occupational Health (CCREOH) Cohort. These chapters provide a valid, reliable, and culturally appropriate measurement of infant and toddler neurodevelopment for Suriname and evidence of neurotoxic effects of prenatal environmental and occupational exposures with TL as a putative biomarker of neurodevelopment. The last two chapters assess TL as a biomarker of neuropsychological performance in minoritized participants of the Wisconsin Alzheimer's Disease Research Center (WADRC) studies. These chapters support the use of peripherally measured TL to capture systemic changes, such as chronological and biological age, and changes of interest for diseases hallmarked by cognitive decline, such as Alzheimer's disease and related dementias (ADRD). Taken together, these novel results contribute to the existing literature and address knowledge gaps in the utility of TL for neuroscience research across the lifespan for minoritized populations.

Precision Measurement of Telomere Length: From Methodology to Meta-Analytic Findings

The first chapter defines the importance of precision, reproducibility, sample size considerations, and overall scientific rigor when measuring TL from human DNA. Adherence to these reporting guidelines and considerations of scientific rigor are seen in the subsequent chapters of the dissertation with intraclass correlation coefficients (ICCs) for TL measurements in the high 0.9s. As a measure of repeatability in the assay, the rigor and reporting metrics used in this dissertation allow the same findings to be obtained by another technician with access to the same samples for TL measurement. To increase the ability for other technicians with the MMqPCR method and TL research laboratories interested in measuring TL with the MMqPCR method, the third chapter identifies the quality control steps and tips necessary to achieve high ICCs for duplicate DNA samples extracted from the same biologic source. As source tissue influences the utility of TL as a biomarker for many diseases and exposures of interest, the second chapter elucidates tissue specific sources of variation within individuals and provides recommendations for researchers selecting a tissue for TL measurement. These first three chapters enhance the scientific rigor of the TL research field, through recommendations, examples of best practices, and data supportive of critical thought prior to TL measurement to reduce spurious findings in the field.

While qPCR reporting recommendations existed previously, the new reporting recommendations put forth in the first chapter are specific to TL research and therefore intrinsically succeed in improving TL research findings when applied.[11] Similarly, the second chapter is not the first to report a correlation of TL between tissues within the same healthy individual.[12] However, the second chapter originally reports the first meta-correlation of TL across 102 tissues within healthy individuals, providing robust evidence of the positive correlation and clarifying reasons for conflicting prior reports. The third chapter also provides

written and visual protocol steps for TL measurement as previously reported by other laboratories; however, the third chapter is the first methodological report of best practices for conducting the MMqPCR TL measurement.[13] These first three chapters contribute greatly to the broad TL research field, where TL is applied as a biomarker for more disciplines than neuroscience.

Telomere Length as a Biomarker of Neurodevelopment in the Diverse Population of the CCREOH Cohort

However, as a neuroscience dissertation, this work naturally led to investigating the utility of TL in neurodevelopment. The fourth chapter of this dissertation created a culturally appropriate neurodevelopmental assessment for Surinamese infants and was applied in the CCREOH cohort study. While the Bayley Scales of Infant and Toddler Development have been utilized in global settings, many items relied on exposures to settings, such as roadways with taxicabs and domesticated cats, that Surinamese infants born and raised in the Amazon rainforest do not have exposure to in their upbringing, making cultural adaptations necessary prior to neurodevelopmental assessment of Surinamese infants.[14] Circling back to investigating TL as a biomarker of neurodevelopment, the fifth chapter defines the role of an infant's biological age, as indexed by TL, in conjunction with a neurotoxicant mixture on the global neurodevelopment of Surinamese infants. The fifth chapter is the first work to investigate TL and exposures on neurodevelopment using Bayesian kernel machine regression analysis. While the overall mixture had a significant negative effect on neurodevelopment, TL may not have a significant role in the casual pathway for neurodevelopment at this age as TL was not driving the mixture's effects.

As this is a diverse cohort, the associations of infant TL with ancestral ethnicity support the weathering hypothesis.[15] The weathering hypothesis has been reported and developed within

the historical context of the United States, however, these findings expand the weathering hypothesis to a different colonialized country. These findings of the fifth chapter show the historically and chronically disadvantaged ethnic groups of Suriname having longer TLs at one year of age than their counterparts regardless of region of the country. While the diverse ethnic groups of Suriname have multiple complex histories, interactions, and disadvantages given by colonialism, the ethnic groups with the oldest contact with colonialism, the descendants of enslaved Africans, have the longest TLs at birth in this cohort.

These findings that expand the weathering hypothesis to a non-United States context, can also be applied to underlying causes of the neurotoxicant effects seen in the findings of chapter five. Primarily, prenatal care for Tribal and Indigenous mothers residing in the Interior region and Amazon rainforest sections of Suriname are economically disadvantaged and rely on small-scale artisanal gold mining using mercury to amalgamate the gold for sale as a primary economic income, despite the occupational hazards and pollutive impacts on their local environment. Furthermore, Tribal and Indigenous mothers often are at a disadvantage for their prenatal care, as health practitioners lack the means to reach them throughout their pregnancy and provide dietary supplements like iodine and folic acid that would decrease the negative effects of their neurotoxicant exposures on their infant's neurodevelopment.

Therefore, in addition to chapter four providing a culturally appropriate measurement of neurodevelopment for Surinamese health practitioners, chapter five provides Surinamese policymakers with scientific evidence of the longitudinal and multigenerational negative effects of mercury, lead, cadmium, and manganese exposure. This evidence should be leveraged to identify programs to prevent future decreases in social and emotional health of infants through increased national prenatal care and policies preventing environmental and occupational

exposures to these neurotoxicants. However, there are significant limitations to the fifth chapter which shall be addressed here, in the larger context of the work.

While TL is a putative biomarker of neurodevelopment, with a few publications for both associations with neurotoxicant exposures in the prenatal period and postnatal TL measurement, and associations with infant, toddler, and adolescent chronological age, TL was a not a significant predictor in the fifth chapter of this work.[16-18] However, TL did play a role in creating the significant decrease in a domain of neurodevelopment as part of a larger mixture with non-linear observations. Interestingly, longer TLs were predictive of lower neurodevelopmental scores in this population, which can be interpreted by a few different theories. Firstly, longer TLs may represent a lack of biological adaptation to neurotoxicant exposure. Here, the appropriate adaptation to the neurotoxicant mixture exposure in the prenatal period that longitudinally protects neurodevelopment could be cellular attrition with shortening of TL to clear the neurotoxicants from the infant's cells. In this first hypothesis, TL is capturing biological adaptations, or lack thereof, in response to environmental exposures. Alternatively, the higher levels of the neurotoxicant mixture exposure could be directly lengthening telomeres in a tumorigenesis pathway as previously reported for other heavy metals. Concurrently, in this second hypothesis, the neurotoxicant mixture is directly negatively affecting global neurodevelopment and TL is a capturing a unique pathway of the mixture where the infants are at increased susceptibility to cancer and other diseases in their lifetimes. A third hypothesis that is a likely explanation is that the finding of longer TLs associated with lower neurodevelopmental scores is that the Bayesian kernel machine regression (BKMR) results are spurious. While these statistical analyses are inherently designed to protect against the risk of multiple comparisons, the BKMR analysis was conducted for four different neurodevelopmental

assessment outcomes which increases the likelihood of spurious findings.[19] However, the BKMR results show an overall effect of the mixture which supports the need to address high mercury levels as the effects of mercury are significant in the closer to real world scenario simulated with BKMR where multiple components were simultaneously analyzed.

In order to confirm one of these hypotheses, further longitudinal research for the CCREOH cohort is needed. In fact, some infants have a second later buccal swab collection, on or after the third year of life, that could be leveraged to assess longitudinal changes in infant TL and clarify the findings in the fifth chapter of this dissertation. Such a future analysis would greatly contribute to the parse literature on the rate of TL shortening in the early years of life.

Furthermore, it is possible that TL was not a significant predictor of neurodevelopment at this age because adaption to the environmental neurotoxicant stressors included increased presence of telomerase and maintenance of TL at a potential cost of later life TL. This results in TL being a better predictor of later life neurological function when considering the contributions of all factors in a lifespan perspective. While TL was not as a robust biomarker of neurodevelopment in the CCREOH cohort as reported in other studies, the findings from the sixth and seventh chapters suggest that TL is a valuable biomarker for neuroscience at the opposite end of the lifespan.

Telomere Length as a Biomarker of Neuropsychological Performance in Minoritized Participants of the WADRC Studies

Given that neurodevelopment is only one, albeit large, section of neuroscience research, a more comprehensive assessment of the utility of TL in neuroscience research would include other research areas, so naturally this dissertation led to the investigation of the most researched neuroscience topic – Alzheimer's disease and related dementias (ADRD). While a personally

important research topic, ADRD is also under researched for minoritized populations in the United States and the utility of TL in ADRD is not well understood.[20-22] However, the seventh chapter reaffirms findings of the second chapter in a novel application to ADRD research and the sixth chapter confirms TL as an adequate biomarker for brain aging in African Americans and American Indians / Alaska Natives in middle to old age.

The sixth chapter is the first study to measure TL from DNA extracted from cerebrospinal fluid. This novel contribution further aligns with the previous meta-analysis in chapter two on cross-tissue TL correlations that lacked analysis of cerebrospinal fluid TL. This sixth chapter shows cerebrospinal fluid TL was significantly positively correlated with whole blood TL from the same individual. However, future research should elucidate any differences in this correlation across the lifespan in a longitudinal manner, given the mid-life average age of the sample studied, and future work should identify any discrepancies of the correlation within individuals of different cognitive statuses. The chapter seven results further support the use of TL in future ADRD research and provide central nervous system relevance to the results of chapter six.

In a trend of novel contributions, the sixth chapter of this dissertation provides the first report of TL for middle to old-aged American Indians / Alaska Natives in relation to neuropsychological aging. These results, in the opposite finding from the fifth chapter, also contribute to the literature supporting the weathering hypothesis. Here, there were no significant differences in average TL by ethnic group. This is aligned with the weathering hypothesis where exposures to social and economic disadvantages during the lifespan have eliminated any TL differences present at birth which is an adaptation to the weathering expected during the lifespan of minoritized populations.[15] This, in combination with the significant TL differences across

ethnic groups in chapter five, supports the utility of TL as a biomarker of weathering which is often investigated and alluded to in neuroscience research.

It is important to note that the populations studied here, diverse Surinamese infants and American Indian / Alaska Native and African American / Black United States older adults, cannot be equated to one another for their obvious differences in race and ethnicity definitions and histories. However, while different, these groups do have shared facets of history including experienced structural racism and colonization. In Suriname, the process and subsequent effects of colonization include minimal presence of Surinen descendants from which the country takes its name, the abolishment of slavery years prior to the United States' equivalent action, and socioeconomic reliance on transactions that often put the health of Suriname's population secondary to immediate financial needs. Minoritized populations cannot be grouped together given their different actions of survivance as non-dominant groups, but they often share environmental and psychosocial risk factors due to their steps of resilience in the face of structural racism and exclusion. Given these shared risk factors which predict worse health outcomes, the evidence in this dissertation collaboratively demonstrates the utility of TL in describing weathering for these distinct minoritized populations.

Furthermore, the sixth chapter shows the predictive utility of TL as a biomarker of neuropsychological decline with advancing age. These findings uniquely position TL as an adequate biomarker for the large minoritized population advancing in age to identify and slow the progression of their biological aging to preserve their cognitive functioning in later life. Therefore, TL is a valuable biomarker in neuroscience research especially at the later life stages. With precision measurement and reproducibility of findings across laboratories, TL could be implemented in the future in clinical testing to identify those with advanced biological aging

compared to their chronological age and intervene with behavioral changes to slow biological aging.

Future Research Directions Given These Findings

Taken together, these dissertation chapters successfully investigated the utility of TL as a biomarker for neurodevelopment and neuropsychological performance with mixed findings, as similarly seen across TL research. Importantly, while both ends of the lifespan were investigated, longitudinal analysis following prenatal exposures of infants, their neurodevelopment, subsequent social and environmental exposures, to neurocognitive decline in old age have not been conducted with scientifically rigorous TL measurement. Hence, identifying the role of early life exposures hypothesized to contribute to risk of ADRD, and the utility of TL in capturing this, still alludes the scientific literature. For example, there is sparse literature on the ability of educational attainment in the United States to predict ADRD later in life for minoritized populations, for whom cognitive function associated with school years is often actually obtained outside of the classroom.[23] Similarly, APOE genotype status and the relationship with ADRD pathogenesis has mixed findings for minoritized populations, in part due to differences in the frequency of alleles present.[24] Given that not all APOE genotypes were available at the time of completion of this dissertation, future work on the role of APOE genotype in the relationship between TL and neuropsychological performance is warranted and future ADRD research must include minoritized populations. Both educational attainment and APOE genotype status are two examples of future research directions which should be investigated for their effects biological effects on TL via cellular mechanisms. As a biomarker, TL offers a great opportunity for future researchers to deduce cellular pathways, utilizing cell cultures, animal models, etc., that produce the neurodevelopment and neuropsychological performance findings reported here.

This dissertation provides evidence of TL as a valuable biomarker for consideration in future neuroscience human research when measured with precision and scientific rigor, with validity in minoritized populations. As seen in the meta-analysis in this dissertation, race and ethnicity are regularly overlooked in TL research, despite the science rooted in humanity and aimed to benefit human health outcomes for all. It is vital to validate TL findings for minoritized populations going forward, and report sample demographics as all doctorates have been trained to do, given the continued existence of disparities in health outcomes. Risks for negative health outcomes are known to differ by race and ethnicity in the United States, given the effects of colonization and racism that continue to operate as designed today, and researchers who research for humans must decolonize their science to center these complex systems in the future if the work aims to better health outcomes for humans. Future research should investigate the utility of TL for other neuroscience research areas, given the promising preliminary findings in other neurodegenerative diseases and developmental disorders. Furthermore, both the CCREOH and the WADRC are longitudinal cohort studies, and future research should attempt to replicate these findings at a later time point as well as provide insight into the TL changes observed across time for these cohorts. In conclusion, TL is a useful biomarker for neuroscience research given the findings of this dissertation, with implications of these findings stretching across the lifespan and applicable to diverse populations.

LIST OF APPENDICES AND SUPPLEMENTAL MATERIALS

Chapter 1.

Supplemental Table 1. Telomere Research Network Reporting Guidelines…….**41**

Supplemental Table 2. Morinha Reporting Guidelines…….**42**

Supplemental Table 3. Reporting guidelines rubric……..**44**

Supplemental Table 4. Individual paper results of reporting guidelines……...**45**

Supplemental Methods. Calculation of the repeatability (ICC) of telomere length measures…..**46**

Supplemental References……...**50**

Chapter 2.

Appendix A. PubMed Search Strategy……..**95**

Appendix B. Example Data Abstraction Spreadsheet……...**96**

Appendix C. Telomere Research Network Telomere Length Data Quality Assurance Checklist……**97**

Appendix D. List of Tissue Types Included in the Meta-Analysis……**97**

Supplementary Materials in Heat Map Format……**100**

Chapter 3.

Supplemental Video. Precision MMqPCR Telomere Length Measurement Methodology……**135**

Available via this link: https://tulane.box.com/s/42nhpmfob3chyn9fsgecvuc4xiikh4o7

Chapter 4.

Supplementary Table 1. Item numbers of starting points for each age group by subscale……..**162**

Supplementary Table 2. Examples of image replacements in the BSID-III-NL picture book and puzzles……..**162**

Supplementary Table 3. Tetrachoric correlations of BSID-III-NL subscale items within each subscale in correlation heat map format..……**172**

Supplementary Table 4. Range of items of interest from cluster analysis by subscale……...**176**

Chapter 5.

Supplemental Methods. Telomere Length Methods……**208**

Supplemental Findings. Bayesian Kernel Machine Regression Analyses with BSID-III-NL Cognitive, Language, and Motor Composite Scores……...**212**

Chapter 6.

Not Applicable.

Chapter 7.

Supplementary Table 1. Telomere Length Measurement Methodology Summary…….**243**

LIST OF REFERENCES

Chapter 1 References……**34**

Chapter 2 References……**92**

Chapter 3 References……...**134**

Chapter 4 References……...**160**

Chapter 5 References……...**202**

Chapter 6 References……...**231**

Chapter 7 References……...**240**

Other References:

[1] Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L., Wu, J.R., 1988. A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. Proc. Natl. Acad. Sci. U.S.A. 85, 6622–6626. https://doi.org/10.1073/pnas.85.18.6622

[2] Hemann, M.T., Strong, M.A., Hao, L.-Y., Greider, C.W., 2001. The Shortest Telomere, Not Average Telomere Length, Is Critical for Cell Viability and Chromosome Stability. Cell 107, 67–77. https://doi.org/10.1016/S0092-8674(01)00504-9

[3] Shay, J.W., Wright, W.E., 2019. Telomeres and telomerase: three decades of progress. Nat Rev Genet 20, 299–309. https://doi.org/10.1038/s41576-019-0099-1

[4] Fice, H.E., Robaire, B., 2019. Telomere Dynamics Throughout Spermatogenesis. Genes (Basel) 10, 525. https://doi.org/10.3390/genes10070525

[5] De Pauw, E.S.D., Verwoerd, N.P., Duinkerken, N., Willemze, R., Raap, A.K., Fibbe, W.E., Tanke, H.J., 1998. Assessment of telomere length in hematopoietic interphase cells using in situ hybridization and digital fluorescence microscopy. Cytometry 32, 163–169. https://doi.org/10.1002/(SICI)1097-0320(19980701)32:3<163::AID-CYTO1>3.0.CO;2-L

[6] Liu, M., Huo, Y.R., Wang, Junwei, Wang, C., Liu, Shuling, Liu, Shuai, Wang, Jinhuan, Ji, Y., 2016. Telomere Shortening in Alzheimer's Disease Patients. Ann Clin Lab Sci 46, 260–265.

[7] Forero, D.A., González-Giraldo, Y., López-Quintero, C., Castro-Vega, L.J., Barreto, G.E., Perry, G., 2016. Meta-analysis of Telomere Length in Alzheimer's Disease. GERONA 71, 1069–1073. https://doi.org/10.1093/gerona/glw053

[8] 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

[9] Haycock, P.C., Heydon, E.E., Kaptoge, S., Butterworth, A.S., Thompson, A., Willeit, P., 2014. Leucocyte telomere length and risk of cardiovascular disease: systematic review and meta-analysis. BMJ 349, g4227. https://doi.org/10.1136/bmj.g4227

[10] Mundstock, E., Sarria, E.E., Zatti, H., Mattos Louzada, F., Kich Grun, L., Herbert Jones, M., Guma, F.T.C.R., Mazzola in Memoriam, J., Epifanio, M., Stein, R.T., Barbé-Tuana, F.M., Mattiello, R., 2015. Effect of obesity on telomere length: Systematic review and meta-analysis: Effect of Obesity on Telomeres. Obesity 23, 2165–2174. https://doi.org/10.1002/oby.21183

[11] Morinha, F., Magalhães, P., Blanco, G., 2020. Standard guidelines for the publication of telomere qPCR results in evolutionary ecology. Mol Ecol Resour 20. https://doi.org/10.1111/1755-0998.13152

[12] Demanelis, K., Jasmine, F., Chen, L.S., Chernoff, M., Tong, L., Delgado, D., Zhang, C., Shinkle, J., Sabarinathan, M., Lin, H., Ramirez, E., Oliva, M., Kim-Hellmuth, S., Stranger, B.E., Lai, T.-P., Aviv, A., Ardlie, K.G., Aguet, F., Ahsan, H., GTEx Consortium, Doherty, J.A., Kibriya, M.G., Pierce, B.L., Aguet, F., Anand, S., Ardlie, K.G., Gabriel, S., Getz, G.A., Graubert, A., Hadley, K., Handsaker, R.E., Huang, K.H., Kashin, S., Li, Xiao, MacArthur, D.G., Meier, S.R., Nedzel, J.L., Nguyen, D.T., Segrè, A.V., Todres, E., Balliu, B., Barbeira, A.N., Battle, A., Bonazzola, R., Brown, A., Brown, C.D., Castel, S.E., Conrad, D.F., Cotter, D.J., Cox, N., Das, S., de Goede, O.M., Dermitzakis, E.T., Einson, J., Engelhardt, B.E., Eskin, E., Eulalio, T.Y., Ferraro, N.M., Flynn, E.D., Fresard, L., Gamazon, E.R., Garrido-Martín, D., Gay, N.R., Gloudemans, M.J., Guigó, R., Hame, A.R., He, Y., Hoffman, P.J., Hormozdiari, F., Hou, L., Im, H.K., Jo, B., Kasela, S., Kellis, M., Kim-Hellmuth, S., Kwong, A., Lappalainen, T., Li, Xin, Liang, Y., Mangul, S., Mohammadi, P., Montgomery, S.B., Muñoz-Aguirre, M., Nachun, D.C., Nobel, A.B., Oliva, M., Park, YoSon, Park, Yongjin, Parsana, P., Rao, A.S., Reverter, F., Rouhana, J.M., Sabatti, C., Saha, A., Stephens, M., Stranger, B.E., Strober, B.J., Teran, N.A., Viñuela, A., Wang, G., Wen, X., Wright, F., Wucher, V., Zou, Y., Ferreira, P.G., Li, G., Melé, M., Yeger-Lotem, E., Barcus, M.E., Bradbury, D., Krubit, T., McLean, J.A., Qi, L., Robinson, K., Roche, N.V., Smith, A.M., Sobin, L., Tabor, D.E., Undale, A., Bridge, J., Brigham, L.E., Foster, B.A., Gillard, B.M., Hasz, R., Hunter, M., Johns, C., Johnson, M., Karasik, E., Kopen, G., Leinweber, W.F., McDonald, A., Moser, M.T., Myer, K., Ramsey, K.D., Roe, B., Shad, S., Thomas, J.A., Walters, G., Washington, M., Wheeler, J., Jewell, S.D., Rohrer, D.C., Valley, D.R., Davis, D.A., Mash, D.C., Branton, P.A., Barker, L.K., Gardiner, H.M., Mosavel, M., Siminoff, L.A., Flicek, P., Haeussler, M., Juettemann, T., Kent, W.J., Lee, C.M., Powell, C.C., Rosenbloom, K.R., Ruffier, M., Sheppard, D., Taylor, K., Trevanion, S.J., Zerbino, D.R., Abell, N.S., Akey, J., Chen, L., Demanelis, K., Doherty, J.A., Feinberg, A.P., Hansen, K.D., Hickey, P.F., Jasmine, F., Jiang, L., Kaul, R., Kibriya, M.G., Li, J.B., Li, Q., Lin, S., Linder, S.E., Pierce, B.L., Rizzardi, L.F., Skol, A.D., Smith, K.S., Snyder, M., Stamatoyannopoulos, J., Tang, H., Wang, M., Carithers, L.J., Guan, P., Koester, S.E., Little, A.R., Moore, H.M., Nierras, C.R., Rao, A.K., Vaught, J.B., Volpi, S., 2020. Determinants of telomere length across human tissues. Science 369, eaaz6876. https://doi.org/10.1126/science.aaz6876

[13] Axelrad, M.D., Budagov, T., Atzmon, G., 2013. Telomere Length and Telomerase Activity; A Yin and Yang of Cell Senescence. JoVE 50246. https://doi.org/10.3791/50246-v

[14] Hambleton, R.K., Merenda, P.F., Spielberger, C.D., 2004. Adapting Educational and Psychological Tests for Cross-Cultural Assessment. Psychology Press.

[15] Geronimus, A.T., 1992. The weathering hypothesis and the health of African-American women and infants: evidence and speculations. Ethn Dis 2, 207–221.

[16] Cowell, W., Colicino, E., Tanner, E., Amarasiriwardena, C., Andra, S.S., Bollati, V., Kannan, S., Ganguri, H., Gennings, C., Wright, R.O., Wright, R.J., 2020. Prenatal toxic metal mixture exposure and newborn telomere length: Modification by maternal antioxidant intake. Environmental Research 190, 110009. https://doi.org/10.1016/j.envres.2020.110009

[17] Benetos, A., Verhulst, S., Labat, C., Lai, T.-P., Girerd, N., Toupance, S., Zannad, F., Rossignol, P., Aviv, A., 2019. Telomere length tracking in children and their parents: implications for adult onset diseases. FASEB J 33, 14248–14253. https://doi.org/10.1096/fj.201901275R

[18] Feiler, M.O., Patel, D., Li, H., Meacham, P.J., Watson, G.E., Shamlaye, C., Yeates, A., Broberg, K., van Wijngaarden, E., 2018. The association between early-life relative telomere length and childhood neurodevelopment. NeuroToxicology 65, 22–27. https://doi.org/10.1016/j.neuro.2018.01.005

[19] Scott, J.G., Berger, J.O., 2010. Bayes and empirical-Bayes multiplicity adjustment in the variable-selection problem. Ann. Statist. 38. https://doi.org/10.1214/10-AOS792

[20] Lukens, J.N., Van Deerlin, V., Clark, C.M., Xie, S.X., Johnson, F.B., 2009. Comparisons of telomere lengths in peripheral blood and cerebellum in Alzheimer's disease. Alzheimer's & Dementia 5, 463–469. https://doi.org/10.1016/j.jalz.2009.05.666

[21] Thomas, P., O' Callaghan, N.J., Fenech, M., 2008. Telomere length in white blood cells, buccal cells and brain tissue and its variation with ageing and Alzheimer's disease. Mechanisms of Ageing and Development 129, 183–190. https://doi.org/10.1016/j.mad.2007.12.004

[22] Gilmore-Bykovskyi, A.L., Jin, Y., Gleason, C., Flowers-Benton, S., Block, L.M., Dilworth-Anderson, P., Barnes, L.L., Shah, M.N., Zuelsdorff, M., 2019. Recruitment and retention of underrepresented populations in Alzheimer's disease research: A systematic review. Alzheimer's & Dementia: Translational Research & Clinical Interventions 5, 751–770. https://doi.org/10.1016/j.trci.2019.09.018

[23] Suchy-Dicey, A.M., Oziel, K., Sawyer, C., Olufadi, Y., Ali, T., Fretts, A.M., Umans, J.G., Shibata, D.K., Longstreth, W.T., Rhoads, K. and Buchwald, D.S., 2022. Educational and clinical associations with longitudinal cognitive function and brain imaging in American Indians: the Strong Heart Study. Neurology, 99(24), pp.e2637-e2647.

[24] Henderson JN, Crook R, Crook J, Hardy J, Onstead L, Carson-Henderson L, et al. Apolipoprotein E4 and tau allele frequencies among Choctaw Indians. Neuroscience Letters 2002;324:77–9. https://doi.org/10.1016/S0304-3940(02)00150-7.