TELOMERE RESEARCH NETWORK

NIH NIA

NIH NIEHS

TRN Recommendations for the measurement of telomere length in population studies

RECOMMENDATION 1: Assay precision has substantial effects on statistical power. There is evidence of significant variability in the Intraclass Correlation Coefficients (ICC, e.g. assay precision) of assays measuring telomere length. When conducting telomere length studies in human populations, investigators need to carefully consider, in addition to other methodologic parameters such as effect size, the influence of this variability in ICCs when calculating statistical power and determining sample size. Assay precision is relevant to both cross sectional and longitudinal studies of telomere length.

Background: The Telomere Research Network (TRN) is a collaborative effort between telomere researchers and the NIH to establish best practices and methodologic guidelines for population based studies of telomere length in relation to psychosocial and environmental exposures and a predictor of later health outcomes. As this multi-year effort conducts methodologically rigorous cross laboratory and cross method comparison studies, we expect to provide data-driven recommendations, as well as up-dates, as needed, to existing recommendations. These cross methodologic studies, and these recommendations are designed to enhance the reproducibility and rigor of the field.

Key considerations: Measurement precision is one of four key factors determining statistical power. Imprecise measurement (measurement error) increases variation in the data, and thereby reduces statistical power. As telomere length studies are examining small differences within individuals, and there is large between individual variation, all assays need to account for the ICC when conducting power analyses. When the level of measurement error is known, as expressed in the ICC, this effect of assay precision can be accounted for when calculating the sample size required to reach a specified statistical power. To avoid underestimating measurement error, the ICC should be based on replicates with the greatest possible independence (ideally repeated samples from the same individual with independent DNA extractions). ICCs should be determined for each cohort because they are assay and lab specific. To assist investigators in incorporation of ICC calculations when determining required sample size for proposed investigations of telomere length in human populations the TRN has provided a generic example of the relation between ICC and statistical power in our published manuscript and on the TRN web page. Further guidance is located on the TRN web page resources related to study design and we provide the statistical code in R to calculate the ICC from replicate measurements of telomere length. Additional information about the impact of assay precision on telomere length studies can also be found in this manuscript.

The relation between ICC and statistical power is both applicable and unique to all assays that utilize repeated measurements. A specific ICC needs to be calculated, and reported, for each study (and each sample type), and cannot be reliably inferred from previous cohorts, the ICCs of other assays, nor the same assay in different laboratories. While the use of repeated measurements to enhance reproducibility is a well-established method across biological assays, telomere length measurements are unusually sensitive to differences in measurement

precision, in particular when examining variation in telomere dynamics, because telomere dynamics are small relative to between-individual variation. As with all statistical power considerations, in addition to the ICC, investigators should be cognizant of expected effect size and the normative expected change in TL over time in longitudinal studies.

Hypothetical example: (Lindrose et al 2021, PloSOne) These examples are provided to guide thinking on the sample size required to investigate differences in telomere length or telomere shortening with sufficient power. Figure A: The sample size required to test effect sizes of 150, 200 and 300 bp in a cross-sectional study with a t-test with a power of 0.9, as a function of measurement error as expressed in the ICC. Figure B. 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 and a four-year follow-up period. Figure C. This is like B but utilizes an eight-year follow-up period. In both figure B and C, baseline telomere shortening was simulated assuming a Poisson distribution with mean/variance of 25, and the population standard deviation of telomere length was maintained at 650bp at both time points. Calculations assumed a realistic (true) standard deviation of 650 bp, sample sizes were equally distributed over the two groups to be compared, and power analysis was done using G-Power.