Power analysis methods accounting for high intra-individual variability in measured cognition

Nicolai D. Ayasse & Sonya Eremenco
Critical Path Institute

Introduction

- Power analyses, conducted to determine the required sample size of a trial, traditionally account for between-person variability and within-person changes over time, and generally assume relative intra-individual stability, homogeneity of variances between groups, and a normally distributed latent variable (i.e., the concept of interest that cannot be directly measured).
  - When these assumptions are untrue, traditional power analyses may underestimate the required sample size, potentially resulting in under-powered trials.
- Measures of cognitive function, such as those used in Alzheimer's disease, challenge these assumptions.
  - They have been consistently observed to produce high intra-individual- compared to between-person variability, a pattern that is exacerbated in people with dementia, a population of great interest for clinical trials.
- Additionally, it is common for cognitive tasks to use task-related time measurements such as response time (e.g., as a proxy for processing speed), which tend to produce skewed rather than normal distributions.

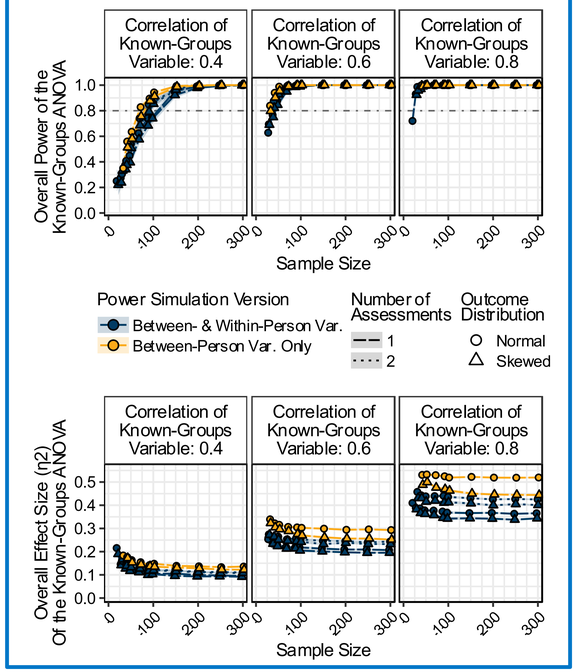
Methods

- Conducted a Monte Carlo power simulation study with 500 replications per condition.
  - With and without including parameters for within-person variability.
  - Within-person variability population parameters based on published parameters from ecological momentary assessments of cognition collected in older adults with and without mild cognitive impairment (Cerino, Katz, et al., 2021, doi: 10.3389/fdgth.2021.758031).
- Assumed cognitive assessments would occur at clinic visits, with the first two occurring one week apart and subsequent visits occurring six months apart.
- The power analysis was conducted for the Virtual Reality Functional Capacity Assessment Tool – Short List (VRFCAT-SLx), but the same methods are applicable to other measures of cognitive performance.

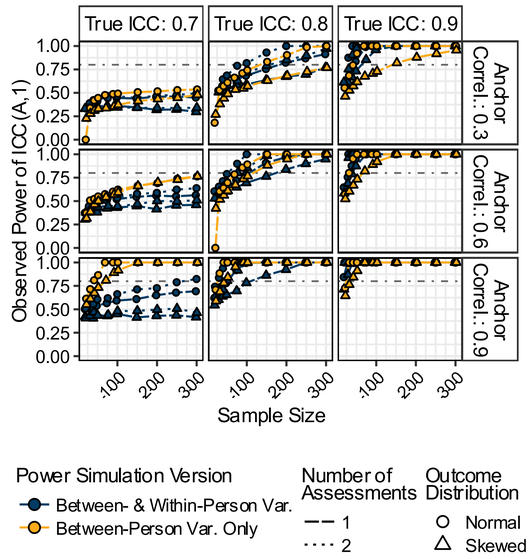
Conditions investigated:

- Normal vs skewed outcome distributions
- One vs two administrations of cognitive assessment (VRFCAT-SLx) per visit
- Varying degrees of strength for population values of:
  - Correlation between known-groups and outcome measure
  - Test-retest reliability
  - Mean change over time
  - Difference between groups

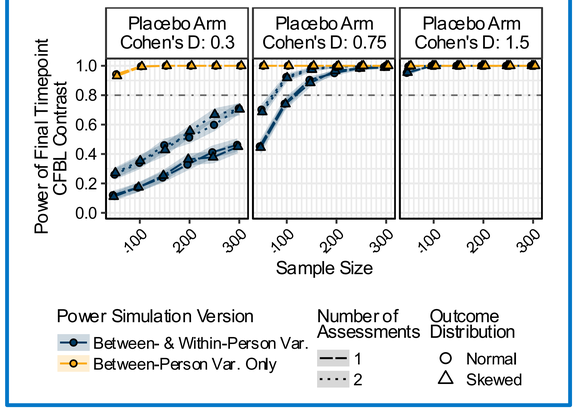
Known-Groups Results



Test-Retest Reliability Results

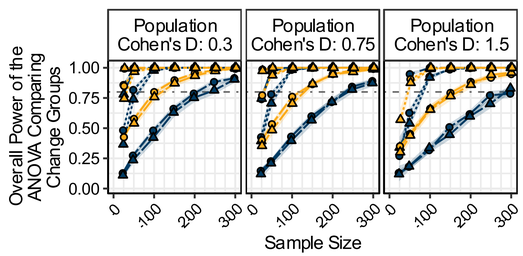


Treatment Comparison Power Results

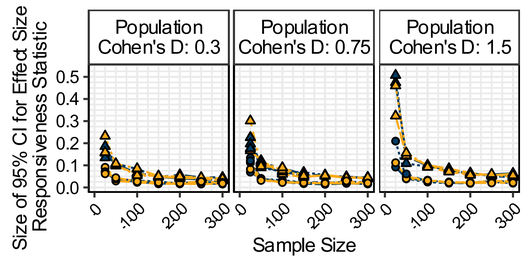


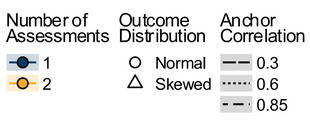
Responsiveness Results

ANOVA Comparing Change Groups



Effect Size Responsiveness Statistic





Conclusions

- Accounting for intra-individual variability in cognition generally resulted in a larger minimum sample size required to achieve power ≥0.80 than when only between-person variability was accounted for.
- A skewed outcome distribution resulted in a larger minimum sample size required to achieve power ≥0.80 than when the outcome was normally distributed.
- Including 2 assessments per clinic visit decreased the minimum required sample size compared to 1 assessment.
- We acknowledge that these simulations did not account for missing data or attrition. We also acknowledge that conducting studies with larger sample sizes is not always feasible, and that conducting adequately-powered studies may at times require more creative solutions.

These results highlight the importance of accounting for characteristics of the data distribution and any unusual properties of the concept of interest when conducting power analyses. In this example, a traditional power analysis that accounts for only between-person variability and assumes the outcome is normally distributed would underestimate the minimum required sample size.

For more information, please contact: **Cole (Nicolai) D. Ayasse, PhD** | *Clinical Outcome Assessment Scientist, Critical Path Institute* | cayasse@c-path.org

Acknowledgments: *Critical Path Institute is supported by the Food and Drug Administration (FDA) of the Department of Health and Human Services (HHS) and is 43% funded by the FDA/HHS, totaling $20,724,703, and 57% funded by non government source(s), totaling $27,346,613. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement by, FDA/HHS or the U.S. Government. This work was supported by FDA DDT Research Grant U01FD008424.*