Biostatistics (2020), **0**, 0, *pp.* 1–6
doi:10.1093/biostatistics/Manuscript-Draft

A tutorial on the "longpower" R package and a web-based App for conducting power analysis and sample size computation for longitudinal data

SAMUEL IDDI[1], MICHAEL C. DONOHUE[1*]

for the Alzheimer's Disease Neuroimaging Initiative[†]

[1] *Data Measurement and Evaluation Unit, African Population and Health Research Center (APHRC), Nairobi, Kenya*

[2] *Alzheimer's Therapeutic Research Institute, Keck School of Medicine, University of Southern California, San Diego, USA*

siddi@aphrc.org

Summary

Longitudinal randomized trials are common in medical and clinical research. These trials often require the collection of data on the same individual at different data collection waves. At the design phase of such studies, sample size computations are critical to ensure that studies are sufficiently powered to provide reliable and valid inference. There are several methodologies for calculating sample sizes for longitudinal studies that depend on many considerations including the type of research design, outcome type, and proposed analytical methods. More advanced procedures take into account several factors leading to more complicated formulas for computing sample sizes.

[*†] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: `http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf`

To provide easy access to common and widely used sample size and power calculation formulas, this tutorial briefly describes some methods for longitudinal data. We also enrich the discussion with real-life examples comparing treatment versus control groups in randomized trials assessing treatment effect on clinical outcomes. Accompanying this tutorial is also a web-based sample size shinyApp developed to help researchers to conduct different sample size and power calculations by allowing user-specified parameters or pilot parameters generated from the Alzheimers Disease Neuroimaging Initiative (ADNI) study.

Key words: Mixed model for repeated measures, linear mixed model, shinyApp, power, sample size, longpower, longitudinal data

1. Introduction

Sample size calculation and power analysis are important components in designing a new trial or study. For many non-statisticians, the processes involved in performing these types of analyses can be daunting. A major hurdle to overcome is the availability of pilot parameters which are required inputs for generating sample size and power outputs. In this paper, we develop a power/sample size App for Alzheimer's Disease (AD) clinical trials. The web-based app implements the linear mixed model and basic mixed method repeated measures (MMRM) allowing user to input own pilot estimates, or use an ADNI-based pilot estimate generator to compute sample size and power.

2. Methods

Clinical trial data collected longitudinally over time are commonly analyzed using the linear mixed model and mixed model for repeated measures methods for continuous outcomes. Prior to such trials, sample size and power analysis are conducted to obtain the required sample size needed to assess treatment effect with optimal power. Various sample size approaches for longitudinal data

have been proposed. We review a few of the most commonly used methods applied in Alzheimer's disease trials focusing on continuous outcome type.

2.1 *Sample size computation based on the linear mixed model (LMM)*

One of the most common sample size computation approaches for correlated data is derived by Liu and Liang(1997). This approach derived sample size from a generalized estimating equation (Liang and Zeger, 1986). Thus, different outcomes types can be handled. A special case is for continuous response measured repeatedly over time and modeled using a linear mixed model (LMM). The model specification involves fixed covariates and random-effects to the outcome for between-subject variability. The error component is assumed to follow a multivariate with a mean vector of zeros and covariance matrix. Different covariance structures such as independence, exchangeable, auto-regressive, and unstructured to estimate the minimum sample size for a given significant level and pre-specified nominal power. A reduced form of the linear mixed model with only the treatment effect and a random-intercept component, the sample size formula generated by this approach is equivalent to that of Diggle et al (1994).

2.2 *Sample size computation based on the mixed model for repeated measures (MMRM)*

An alternative sample size computation method used in many areas of application including AD trials is that based on the mixed model for repeated measures (MMRM Mallinckrodt et al. 2001a; Mallinckrodt et al. 2003; Lane 2008). The MMRM method of model fitting is similar to the mixed-effect model for longitudinal or repeated measures except for the unstructured modelling of time treated as a categorical variable, and the specification of a within-subject error structure. Additionally, the MMRM considers the interaction between time and treatment.

2.3 *The 'longpower' package*

2.4 *A Web-based App*

This ShinyApp dashboard is developed to easily generate sample size and conduct power analysis for a longitudinal study design with two-group comparisons for a continuous outcome. The App implements the sample size formulae of Liu and Liang (1997) and Diggle et al (1994) using functions developed for the R 'longpower' package. The 'longpower' package handles cases where time is treated either as continuous or categorical. The former approach uses the linear mixed model with random intercept and slope while the later leads to the well-known Mixed Model of Repeated Measures (MMRM) used in many clinical trial applications for conducting statistical analysis.

The dashboard is in two parts. The first part accepts user inputs to generate sample size when time is treated as both categorical and continuous. Thus, this part assumes that the user already has pilot parameter estimates including effect size and known variance. Users can generate sample sizes and perform power analysis using different sample size methods (diggle, liuliang, and edland). The basic inputs of the App include an option for whether sample size or power is desired, a slider for sample size and power enabled when appropriate and difference in means (treatment effect) estimate. Additional options include the type I error rate, type of test (one/two sided), estimates of random effect variances, sample size computation method, allocation ratio and time intervals. For the MMRM method, options for the association structure and retention in each group is enable depending on the association structure assumed by the user.

The second part of the dashboard uses the methodology that is similar to the first part except that Alzheimer's Disease Neuroimaging Initiative (ADNI)-based pilot parameters are generated based on user-selected inclusion and exclusion criteria, primary outcome, duration of the study, and covariate options to be included in the linear mixed and MMRM models.

The App interface display plots of power against sample size over range of values, and text

summary of imputed and selected estimates. For the ADNI-based generator, additional summary estimates from fitted model and descriptive summary of baseline participants characteristics are also dynamically generated.

3. Motivating examples

4. Discussion

Acknowledgments

We are grateful to the ADNI study volunteers and their families. This work was supported by Biomarkers Across Neurodegenerative Disease (BAND-14-338179) grant from the Alzheimer's Association, Michael J. Fox Foundation, and Weston Brain Institute; and National Institute on Aging grant R01-AG049750. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions

are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. *Conflict of Interest*: None declared.

References

[*Received August 1, 2010; revised October 1, 2010; accepted for publication November 1, 2010*]