Scoping Review Protocol: Statistical Models for Longitudinal Data

Ariel I. Mundo Ortiz

2022-08-06

Table of contents

1 Background

Longitudinal studies are frequently used in the health sciences (biomedical research, epidemiology, public health, among others) as they allow to examine how the temporal effect of a treatment or intervention. Despite their benefits, the inherent time dependency of longitudinal studies creates additional challenges with regard to data missingness, financial costs, and more importantly, the selection statistical methodologies that are appropriate to analyze the collected signal[1,2].

Historically, the repeated measures analysis of variance (rm-ANOVA) has been the preferred method to analyze longitudinal data, despite the fact that frequently, the assumptions required for the application of this method are not satisfied by longitudinal data[2]. Modern statistical methods that overcome the limitations of rm-ANOVA have been developed, such as linear mixed models, generalized additive models, Bayesian models, and generalized estimating equations[3–6]. During the last decade, these modern statistical methods have started to being adopted in the health sciences to analyze longitudinal data[7–10], but so far, it is not know how widespread their use is, and what are the challenges that may limit their adoption by the broad health sciences field.

Surveying the statistical methods used to analyze longitudinal data in the health sciences can provide a better understanding of the extent in the adoption of modern statistical methods, and provide rationale for situations where they are still not widely adopted.

2 Objective

This study aims to summarize the different statistical models for longitudinal data that are used in the health sciences, identify the extent of the adoption of modern statistical methods in the field, and determine if in each case, model assumptions are checked by researchers to ensure congruency between the data and the model.

3 Review Question

Summarize the statistical methods used to analyze longitudinal data in the health sciences to identify which methods are most commonly used, the applicability of such methods in the context of each study, and gaps that might exist that prevent the adoption of modern statistical methods that can be better suited to analyze the data. Additionally, identify if studies check for model assumptions, and how this in turn impacts the reported results.

4 Databases

- PubMed
- Web of Science

5 Search Terms

6 Criteria

6.1 Inclusion Criteria

- methods paper see new methods developed

- application

6.2 Exclusion Criteria

7 Additional Resources

8 Comparison (?)

9 Data Extraction

10 Data Synthesis Strategy

11 References

1. Caruana EJ, Roman M, Hernández-Sánchez J, Solli P. Longitudinal studies. *Journal of Thoracic Disease.* 2015;7(11):E537-40.

2. Mundo AI, Tipton JR, Muldoon TJ. Generalized additive models to analyze nonlinear trends in biomedical longitudinal data using r: Beyond repeated measures ANOVA and linear mixed models. *Statistics in Medicine.* Published online July 2022.

3. Jiang J, Nguyen T. *Linear and Generalized Linear Mixed Models and Their Applications.* 2nd ed. Springer; 2021.

4. Hastie TJ. *Statistical Models in S.* (Chambers JM, Hastie TJ, eds.). Routledge; 2017.

5.   Rosa GJM, Gianola D, Padovani CR. Bayesian longitudinal data analysis with mixed models and thick-tailed distributions using MCMC. *Journal of Applied Statistics.* 2004;31(7):855-873.

6.   Ballinger GA. Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods.* 2004;7(2):127-150.

7.   Mundo AI, Muhammad A, Balza K, Nelson CE, Muldoon TJ. Longitudinal examination of perfusion and angiogenesis markers in primary colorectal tumors shows distinct signatures for metronomic and maximum-tolerated dose strategies. *Neoplasia.* 2022;32:100825. doi:10.1016/j.neo.2022.100825

8.   Wang M. Generalized estimating equations in longitudinal data analysis: A review and recent developments. *Advances in Statistics.* 2014;2014:1-11.

9.   Tian Q, Qin L, Zhu W, Xiong S, Wu B. Analysis of factors contributing to postoperative body weight change in patients with gastric cancer: Based on generalized estimation equation. *PeerJ.* 2020;8(e9390):e9390.

10.   Şevik M, Doğan M. Epidemiological and molecular studies on lumpy skin disease outbreaks in turkey during 2014-2015. *Transboundary and Emerging Diseases.* 2017;64(4):1268-1279.