JAMA Guide to Statistics and Methods

Analyzing Repeated Measurements Using Mixed Models

Michelle A. Detry, PhD; Yan Ma, PhD

Longitudinal studies often include multiple, repeated measurements of each patient's status or outcome to assess differences in outcomes or in the rate of recovery or decline over time. Repeated measurements from a particular patient are likely to be more similar to each other than measurements from different patients, and this correlation needs to be considered in the analysis of the resulting data. Many common statistical methods, such as linear regression models, should not be used in this situation because those methods assume measurements to be independent of one another.

It is possible to compare outcomes between treatments using only a final measurement to determine whether there was a difference at the end of the study; however, this approach would not include much of the information captured with repeated measurements and there would be no consideration of the pattern of outcomes each patient experienced in reaching his or her final outcome. When outcomes are measured repeatedly over time, a wide variety of clinically important questions may be addressed.

In the EXACT study recently published in *JAMA*, Moseley et al[1] examined activity limitations and quality of life (QOL) among patients with ankle fractures to determine if a supervised exercise program with rehabilitation advice was more beneficial than advice alone. Activity limitations and QOL were measured at baseline and at 1, 3, and 6 months of follow-up. The authors used mixed models[2] to compare patient outcomes over time between the 2 intervention groups.

Use of the Method

Why Are Mixed Models Used for Repeated Measures Data?

Mixed models are ideally suited to settings in which the individual trajectory of a particular outcome for a study participant over time is influenced both by factors that can be assumed to be the same for many patients (eg, the effect of an intervention) and by characteristics that are likely to vary substantially from patient to patient (eg, the severity of the ankle fracture, baseline level of function, and QOL). Mixed models explicitly account for the correlations between repeated measurements within each patient.

The factors assumed to have the same effect across many patients are called fixed effects and the factors likely to vary substantially from patient to patient are called random effects. For example, the effect of a new treatment may be assumed to be the same for all patients and modeled as a fixed effect, whereas patients may have markedly different baseline function or inherent rates of recovery and these may be best modeled as random effects. Mixed models are called "mixed" because they generally contain both fixed and random effects. The ability to consider both fixed and random effects in the model gives flexibility to determine the effects of multiple factors and to address specific questions of clinical importance. In contrast, repeated measures analysis of variance (ANOVA), often used for analyzing longitudinal data, does not have this flexibility and can yield misleading results if its more rigid assumptions (eg, all effects are considered fixed) are not met.

Furthermore, using a mixed model, data from all assessments contribute to the treatment comparisons, resulting in more precise estimates and a more powerful study. A mixed model can also address if outcomes changed over time (eg, the rate of recovery of function or decline) within each treatment group. Moreover, in addition to population-level comparisons, mixed models can be used to characterize an individual patient's response patterns over time. The specific clinical question motivating the trial determines the structure of the mixed model that is most applicable. For example, if the effect of a treatment on the rate of recovery from a patient-specific baseline is to be determined, then the mixed model is likely to include a random baseline effect and a fixed interaction term between treatment group and time, with the latter term capturing the effect of the treatment on the rate of recovery.

Observations may be correlated with each other in several different ways. These patterns are known as correlation structures and it is important when using mixed models to use the correct structure. For example, if the correlation between each measurement is likely to be the same regardless of the length of time between the measurements, then a "compound symmetry" structure is appropriate. In contrast, if the correlation between measurements decreases as the time between measurements increases, then an "autoregressive" structure should be used. Finally, an "unstructured" correlation can be used if no constraints can be imposed on the correlation pattern, but fitting a model with an unstructured correlation requires a larger data set than the other approaches.

Ideally, the assumed correlation structure should be based on the clinical context in which the repeated measurements were taken. For example, certain longitudinal data (eg, pain scores after joint surgery) at adjacent assessments would tend to be more correlated than those measured farther apart, making an autoregressive structure appropriate. Statistical testing (eg, a likelihood ratio test) may be used when an objective comparison is needed to evaluate competing correlation structures.

Incomplete outcome data, for example, caused by patients missing some visits or dropping out of the study, are common in longitudinal studies.[3] As a result, study participants may have different numbers of available measurements, a situation that cannot be addressed by repeated measures ANOVA. Mixed models can accommodate unbalanced data patterns and use all available observations and patients in the analysis. Mixed models assume that the missingness is independent of unobserved measurements, but dependent on the observed measurements.[4,5] This assumption is called "missing at random" and is often reasonable.[3,5] Repeated measures ANOVA requires a more unlikely assumption

that the missingness is independent of both the observed and unobserved measurements, called "missing completely at random." Using mixed models, reasonably valid estimates of treatment effects can often be obtained even when the missing values are not completely random and additional methods for handling missing data, such as multiple imputation, are generally not required.[3-5]

What Are the Limitations of Mixed Models?

As with any statistical model, a mixed model will have limited validity if its underlying assumptions are not met. For example, if the effect of a treatment varies substantially from patient to patient, for instance, because of genetic differences, then considering the treatment effect as fixed may not be reasonable. Similarly, the assumed correlation structure can adversely impact model results and study conclusions if incorrect. It is important to ensure that the structure of the mixed model matches what is reasonably believed about the clinical setting in which the model is applied.

Because of the larger number of parameters to be estimated from the data, mixed models may be difficult to estimate or "fit" when the available data are limited. This is especially true if an unstructured correlation structure must be used. The precise methods used by different software packages to fit mixed models differ, so the numerical results can vary somewhat based on the statistical software used.

In the presence of missing data, mixed models can provide valid inferences under an assumption that data are missing at random. However, in practice it is often impossible to know that this assumption is met and informative censoring (nonignorable missingness) can never be ruled out. If the investigators suspect deviation from the missing-at-random assumption, sensitivity analyses may be conducted using models appropriate for nonignorable missingness. The models used would depend on the study design, missing data patterns observed, and other study specific considerations.[2]

Why Did the Authors Use Mixed Models in This Particular Study?

The EXACT trial investigators used mixed models in their analyses because they wanted to answer the question of how outcomes changed over time and how they were affected by treatment. The model included fixed effects for treatment group, time of measurement, and baseline score. An interaction term between treatment group and time was also included to determine if the 2 treatment interventions led to different recovery trajectories over time. In addition, the model included a random effect for the baseline value, addressing the variability in the starting point for each patient.

The EXACT trial reported that in each treatment group, 10% to 20% of the patients were lost to follow-up as the study progressed. Thus, it was important for the authors to examine the effects of the missingness. They included a preplanned sensitivity analysis that used multiple imputation[5] to evaluate how sensitive the primary outcome result was to the missing at random data assumption. The results of the main and sensitivity analyses were similar.

Caveats to Consider When Looking at Results From Mixed Models

As with most statistical models, it is important to consider whether the structure of the data obtained and the clinical setting (eg, repeated measures over time) match the model structure. It is often useful to inspect graphical data summaries (eg, "spaghetti" or "string" plots showing the outcome trajectories of individual study participants over time) to determine whether the observed data patterns appear consistent with model assumptions.

When outcome data are missing, the analyst should consider whether the pattern of missingness is likely to be random, meeting the assumptions inherent in mixed models. The rationale for the chosen correlation structure should be clear and based on study design (eg, the pattern of follow-up visits) rather than based on what allows a model to be fit with the available data.

ARTICLE INFORMATION

Author Affiliations: Berry Consultants, LLC, Austin, Texas (Detry); Department of Epidemiology and Biostatistics, Milken Institute School of Public Health, The George Washington University, Washington, DC (Ma).

Corresponding Author: Michelle A. Detry, PhD, Berry Consultants LLC, 4301 Westbank Dr, Bldg B, Ste 140, Austin, TX 78746 (michelle@berryconsultants.com).

Section Editors: Roger J. Lewis, MD, PhD, Department of Emergency Medicine, Harbor-UCLA Medical Center and David Geffen School of Medicine at UCLA; and Edward H. Livingston, MD, Deputy Editor, *JAMA*.

Conflict of Interest Disclosures: The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

Funding/Support: Dr Ma receives funding from the Agency for Healthcare Research and Quality.

REFERENCES

1. Moseley AM, Beckenkamp PR, Haas M, Herbert RD, Lin CW; EXACT Team. Rehabilitation after immobilization for ankle fracture: the EXACT randomized clinical trial. *JAMA*. 2015;314(13):1376-1385.

2. Fitzmaurice GM, Laird NM, Ware JH. *Applied Longitudinal Analysis*. 2nd ed. Hoboken, NJ: Wiley; 2011.

3. Newgard CD, Lewis RJ. Missing data: how to best account for what is not known. *JAMA*. 2015;314(9):940-941. doi:10.1001/jama.2015.10516.

4. Ma Y, Mazumdar M, Memtsoudis SG. Beyond repeated-measures analysis of variance: advanced statistical methods for the analysis of longitudinal data in anesthesia research. *Reg Anesth Pain Med*. 2012;37(1):99-105.

5. Li P, Stuart EA, Allison DB. Multiple imputation: a flexible tool for handling missing data. *JAMA*. 2015;314(18):1966-1967. doi:10.1001/jama.2015.15281.