J Bras Pneumol. 2023;49(2):e20230137
https://dx.doi.org/10.36416/1806-3756/e20230137

Mixed-effects model: a useful statistical tool for longitudinal and cluster studies

Leda Tomiko Yamada da Silveira[1,2,3], Juliana Carvalho Ferreira[1,4], Cecilia Maria Patino[1,5]

PRACTICAL SCENARIO

A multicenter clustered clinical study involving patients with multimorbidity hospitalized at 82 hospitals in Switzerland evaluated the effectiveness of a hospital discharge planning tool embedded in the electronic medical records in comparison with standard discharge procedures on reducing hospital length of stay.[1] The authors analyzed the results of an interrupted time series of acute medical hospitalizations. They reported that the planning tool vs. standard procedures reduced the length of hospital stay (−0.879 h/month; 95% CI, −1.607 to −0.150 h/month vs. −0.011/month; 95% CI, −0.281 to 0.260 h/month), with no increase in hospital readmissions, in-hospital mortality, or facility discharge.

CHARACTERISTICS AND USEFULNESS OF MIXED-EFFECTS MODELS

Regression models, such as linear and logistic regression, are widely applied to evaluate relationships between an exposure (independent variable) and an outcome (dependent variable). Each type of regression model has underlying assumptions that must be met. If these assumptions are not met, a different statistical approach should be required to prevent bias. In our practical scenario, the outcome was repeatedly assessed over time and across several centers. Thus, to accommodate both repeated measures and results across different centers, the authors used a mixed-effects model to analyze the data.

Mixed-effects models include specific subtypes with different names, such as hierarchical models, multilevel models, and cross-sectional time series. Notably, the key characteristic of mixed models is that they integrate both fixed and random effects in the same analysis.

Fixed effects are exposures with specific values that do not change throughout the study, such as sex, age at baseline, race, and ethnicity, and whose relationship with the outcome is constant or fixed. When we include a variable as a fixed effect in a model, we assume that no relationship exists among the levels of this variable. For example, if we include sex as a fixed effect, the model will provide a separate estimated effect for males and females. Exposures, predictors, and interventions at baseline are usually treated as fixed effects in such studies, because we are interested in the independent effect of each of the variable levels on the outcome.

Random effects are more challenging to define and assimilate. In brief, when a variable is introduced in a statistical model as a random effect, its levels are not entirely independent. On the contrary, there is a relationship across the measurements, which adds variability.

Measurements from the same participant over time and within a cluster (in our scenario, the hospitals) are more similar than are measurements across participants not in a cluster. For this reason, to measure the effect of clusters, statistical models that incorporate random effects should be considered in the analytic approach to account for random variation within and between clustering.

In our practical scenario, the data set contained repeated measurements from the same participant, which precludes using, for example, linear or logistic regression analysis. Additionally, because the data were collected from 82 different hospitals, participants were grouped at the center level. Therefore, the authors used mixed-effects regression analysis with random effects for hospitals and patients to account for individual tendencies for each hospital and each patient. Age, sex, comorbidity index, and frailty score were inserted into the model as fixed effects.

Mixed models are suitable for studies with

- Repeated measurements of the same variable from same participant over time (e.g., longitudinal studies)
- Clustered observations
  (e.g., within a hospital, clinic, or doctor)

Typical random effect variables

- Subjects, neighborhood, hospital, treating doctor

Figure 1. Examples of situations in which mixed effects models should be considered in the analysis plan and examples of variables that are often used as random effects.

1. Methods in Epidemiologic, Clinical, and Operations Research-MECOR-program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.
2. Departamento de Fisioterapia, Fonoaudiologia e Terapia Ocupacional, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
3. Hospital Universitário, Universidade de São Paulo, São Paulo (SP) Brasil.
4. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.
5. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (CA) USA.

Examples of studies that require mixed-effects models are time series, longitudinal studies with multiple measurements taken over time on the same subject, and studies that involve participants from different sites or that are evaluated by different physicians (clustered measurements). In such cases, the effects of participants and centers would be considered a random effect, and they would be added to the model to take into account correlations between measurements[2] that traditional linear or logistic regression cannot address.

Choosing the appropriate statistical model to answer specific research questions considering the data set structure improves the models' fit. In addition, it enhances the interpretation of studies that could potentially be used to guide clinical decision making.

KEY POINTS

- Mixed-effects models combine fixed and random variables.
- Fixed effects do not change over time; their levels have explicit values that are not correlated. E.g., age at baseline, color of eyes, race/ethnicity.
- Random effects introduce sources of variation within or across patients.
- Mixed-effects models account for correlations across measurements; therefore, they may be used in studies with longitudinal and clustered data.
- It is important to align research questions and data configuration with the appropriate statistical model to answer and interpret study results adequately, thus guiding clinical decision making efficiently.

REFERENCES

1. Kutz A, Koch D, Haubitz S, Conca A, Baechli C, Regez K, et al. Association of Interprofessional Discharge Planning Using an Electronic Health Record Tool With Hospital Length of Stay Among Patients with Multimorbidity: A Nonrandomized Controlled Trial. JAMA Netw Open. 2022;5(9):e2233667. https://doi.org/10.1001/jamanetworkopen.2022.33667
2. Schober P, Vetter TR. Linear Mixed-Effects Models in Medical Research. Anesth Analg. 2021;132(6):1592-1593. https://doi.org/10.1213/ANE.0000000000005541