Journal of Clinical Epidemiology 82 (2017) 158−166

Journal of Clinical Epidemiology

Latent growth models matched to research questions to answer questions about dynamics of change in multiple processes

Graciela Muniz-Terrera[a,*], Annie Robitaille[b], Amanda Kelly[b], Boo Johansson[c], Scott Hofer[b], Andrea Piccinin[b]

[a]*Centre for Dementia Prevention, Kennedy Tower, Royal Edinburgh Hospital, University of Edinburgh, Scotland*
[b]*Department of Psychology, Cornett Building, University of Victoria, Victoria, British Columbia, Canada*
[c]*Department of Psychology, University of Gothenburg, Haraldsgatan 1, Gothenburg 40530, Sweden*

Accepted 2 September 2016; Published online 14 September 2016

Abstract

Objectives: Given theoretical and methodological advances that propose hypothesis about change in one or multiple processes, analytical methods for longitudinal data have been developed that provide researchers with various options for analyzing change over time. In this paper, we revisited several latent growth curve models that may be considered to answer questions about repeated measures of continuous variables, which may be operationalized as time-varying covariates or outcomes.

Study Design and Setting: To illustrate each of the models discussed and how to interpret parameter estimates, we present examples of each method discussed using cognitive and blood pressure measures from a longitudinal study of aging, the Origins of Variance in the Old-Old study.

Results and Conclusion: Although statistical models are helpful tools to test theoretical hypotheses about the dynamics between multiple processes, the choice of model and its specification will influence results and conclusions made.

Keywords: Latent growth model; Time-varying covariates; Bivariate latent growth model; Longitudinal models; Repeated measures, Multilevel models

1. Introduction

The identification of the most appropriate statistical method to answer a research question is an essential step toward obtaining a meaningful answer to the question posed. Concurrent with advances in theoretical models [1,2] that propose dynamic associations between change in one or multiple processes, development of software [3−5] and analytical methods for longitudinal data [6−12] facilitate the implementation of these newly developed models. These advancements provide researchers with the opportunity to re-examine theoretical models of increasing complexity although they have also made the selection of the most appropriate model for answering the various possible questions posed and the interpretation of its results increasingly challenging.

Latent growth curve models (LGMs) are statistical models conceptualized under the structural equation modeling (SEM) framework often used for the analysis of univariate trajectories of longitudinal data [13−15]. LGM permits the estimation of mean and subject-specific curves and the inclusion of covariates that may be time invariant (i.e., variables that do not change over time, e.g., sex) or time varying (i.e., variables that change over time, e.g., blood pressure [BP]). See Appendix A at www.jclinepi.com for a mathematical formulation of an unconditional LGM and pictorial representation of an LGM with time-invariant covariates (Fig. A1 at www.jclinepi.com). In these longitudinal models, time-invariant covariates are included to explain differences between individuals (e.g., to examine hypothesis regarding differences in level and rate of

Conflict of interest: The authors have no conflicts of interest to disclose. All authors have made substantial contributions to the article and have approved this submission.

Funding: Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under award number P01AG043362 for the Integrative Analysis of Longitudinal Studies of Aging research network.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

* Corresponding author. Tel.: 0131 537 6530.
E-mail address: G.Muniz@ed.ac.uk (G. Muniz-Terrera).

http://dx.doi.org/10.1016/j.jclinepi.2016.09.001

What is new?

Key findings

- Advances in theoretical models and analytical methods provide researchers with ample opportunities for research, but mismatches between questions and methods are possible.
- When repeated measures of multiple variables are collected over time, whether longitudinal variables are operationalized as time-varying covariates (TVCs) or outcomes depends on the question of interest and leads to the application of substantially different analytical approaches.

What this adds to what was known?

- We revisited and provide guidance about various latent growth curve models (LGMs) to help researchers identify the best methodology to answer a series of commonly asked questions about the dynamics of change in single or multiple longitudinal variables.

What is the implication and what should change now?

- LGMs with TVCs and multivariate models of change allow for the examination of change over time and the examination of longitudinal relationships between two or more variables.
- Variations of these longitudinal models are also possible and should be considered.
- The choice of longitudinal model should be based on the research question of interest and take into account characteristics of the data used.

cognitive change between men and women). The role of other variables that are measured repeatedly over time is less clear as they may explain change within individuals (how an individual changes compared to their previous level) while also explain differences between individuals (how an individual differs from another individual in the sample). The additional complexity of accounting for the different sources of information conveyed by time-varying variables increases the difficulty in choosing the most appropriate analysis method. Furthermore, they can be operationalized in different ways resulting in either univariate models with time-varying covariates (TVCs) or multivariate models of change.

The purpose of the current paper is to present an overview and discussion of a series of research questions that involve the analysis of information from variables that have been measured repeatedly over time and that may be answered by fitting univariate LGMs with TVCs or multivariate LGMs of change. Because these are related but substantially different models, understanding what each model offers and when they can be best applied is essential to avoid a mismatch between research questions and models used. Moreover, variations of these two models can and should be considered depending on the research question being addressed. We present exemplary questions that researchers, in the field of aging, may ask and to illustrate the application of the different models, we also present an empirical example where we match research questions to methods. We hope this paper will be a useful tool for investigators involved in research using longitudinal designs.

1.1. Matching models to questions

To facilitate the matching of research questions to the most appropriate statistical method, we present and discuss the different models in the context of answers to a series of hypothetical questions. Even though the hypothetical questions we present here examine possible dynamics of cognitive function (CF) and BP, the models discussed are applicable to other research areas where similar questions are of interest. This is different from the current literature on research methods, which tends to describe the mathematical specifics of the statistical modeling approaches with less focus on the actual questions that can be answered by each model. Our approach, which highlights how the nuances between models answer slightly different questions, will be useful to all readers including those who are less mathematically oriented.

1.2. LGMs with a TVCs

1.2.1. Research questions

Question 1: ‘‘What is the trajectory of CF once the effect of BP, at each concurrent time point is accounted for?’’; question 2: ‘‘Is the effect of BP on CF consistent over time, considering the fact that CF changes?’’; question 3: ‘‘Do individuals whose BP is higher than average also have higher than average CF?’’

1.2.2. Analytical approach

A LGM with a TVC (TVC model) can be considered to answer these three questions as described next. Specifically, they can be answered fitting a TVC model to CF measurements with BP measures regarded as TVCs.

The TVC model is an extension of an LGM that permits the incorporation of covariates that, similarly to the outcome variable, are also measured repeatedly over time. It estimates the trajectory of the outcome variable (CF in our example) as a function of the metric of time, but, it does so while simultaneously controlling the outcome at each time point for some measure of the TVC (BP in our example).