

Published in final edited form as:

Stat Interface. 2009 January 1; 2(4): 403–412.

Using latent outcome trajectory classes in causal inference*

Booil Jo,

Department of Psychiatry & Behavioral Sciences, Stanford University, Stanford, CA 94305-5795

Chen-Pin Wang, and

Department of Epidemiology & Biostatistics, University of Texas Health Science Center

Nicholas S. Ialongo

Department of Mental Health, Johns Hopkins University

Booil Jo: booil@stanford.edu; Chen-Pin Wang: cwang@verdict.uthscsa.edu; Nicholas S. Ialongo: nialongo@jhsph.edu

Abstract

In longitudinal studies, outcome trajectories can provide important information about substantively and clinically meaningful underlying subpopulations who may also respond differently to treatments or interventions. Growth mixture analysis is an efficient way of identifying heterogeneous trajectory classes. However, given its exploratory nature, it is unclear how involvement of latent classes should be handled in the analysis when estimating causal treatment effects. In this paper, we propose a 2-step approach, where formulation of trajectory strata and identification of causal effects are separated. In Step 1, we stratify individuals in one of the assignment conditions (reference condition) into trajectory strata on the basis of growth mixture analysis. In Step 2, we estimate treatment effects for different trajectory strata, treating the stratum membership as partly known (known for individuals assigned to the reference condition and missing for the rest). The results can be interpreted as how subpopulations that differ in terms of outcome prognosis under one treatment condition would change their prognosis differently when exposed to another treatment condition. Causal effect estimation in Step 2 is consistent with that in the principal stratification approach (Frangakis and Rubin, 2002) in the sense that clarified identifying assumptions can be employed and therefore systematic sensitivity analyses are possible. Longitudinal development of attention deficit among children from the Johns Hopkins School Intervention Trial (Ialongo et al., 1999) will be presented as an example.

Keywords and phrases

Causal inference; Latent trajectory class; Longitudinal outcome prognosis; Growth mixture modeling; Principal stratification; Reference stratification

1. INTRODUCTION

Heterogeneity in outcome development or prognosis has been of great interest both in observational and experimental studies. To study heterogeneity in terms of repeated measures, various modeling strategies have been developed. A line of methods, widely known as mixed effects modeling or latent growth modeling (e.g., Diggle, Liang, and Zeger, 1994; Laird and Ware, 1982; Meredith and Tisak, 1990; Raudenbush and Bryk, 2002), views this heterogeneity

*The research of the first and the third author was supported by NIMH (MH066319, MH066247) and the research of the second author was supported by NIDDK (DK075092). We thank participants of the Prevention Science Methodology Group Teleconference for helpful feedback. We also appreciate valuable input from Tihomir Asparouhov.

Correspondence to: Booil Jo, booil@stanford.edu.

as a continuum. Another line of approach is known as latent class growth modeling (e.g., Jones, Nagin, and Roeder, 2001; Nagin, 1999), which focuses on heterogeneity in a discrete sense. A third approach, often referred to as growth mixture modeling (e.g., Muthén, 2004; Muthén and Shedden, 1999), also focuses on heterogeneity in a discrete sense, but allows for variation among individuals (i.e., heterogeneity in the continuous sense) within each trajectory class. The third approach, which we will focus on in this paper, can be seen as a combination of the first and the second line of approaches. Whether we should study heterogeneity as a continuous or a discrete entity is a debatable issue. We take the discrete perspective in this paper mainly because the same perspective is usually taken in principal stratification (Frangakis and Rubin, 2002) models, which we intend to utilize to facilitate causal inference considering heterogeneity in outcome trajectories.

Growth mixture modeling is an efficient way of identifying distributionally distinct latent trajectory classes on the basis of longitudinal outcome information. The underlying subpopulations identified by these latent classes are often substantively and clinically meaningful especially when experts' opinion is incorporated in the model selection process. The idea of discrete subpopulations is particularly appealing in the context of studying intervention or treatment effects because these subpopulations may show different treatment responses. In this context, the results of growth mixture analyses may hold important treatment implications and provide useful information for better strategizing future trials. Once we shift our interest from simply identifying heterogeneous trajectory classes to identifying differential treatment effects for these latent classes, whether and under what conditions we can interpret the results as causal becomes a critical issue.

The usual practice of growth mixture modeling is mostly exploratory in the sense that we decide on a particular solution heavily relying on empirical model fitting, where the role of parametric assumptions such as normality can be crucial. The exploratory nature of growth mixture modeling should be considered an advantage because the resulting data-driven, latent trajectory class solutions are likely to have better empirical fit compared to the solutions obtained by imposing external, theory-driven rules that dictate trajectory classes. However, the same exploratory nature creates a puzzling situation when growth mixture modeling is used to estimate treatment effects for different latent trajectory classes. In principle, growth mixture analyses can simultaneously utilize all available information, including treatment response, to effectively divide individuals into latent trajectory classes (Muthén et al., 2002; Muthén and Brown, 2009). In this one-step approach, both the treatment effects and latent trajectory classes are identified on the basis of empirical fit. Whether we interpret the resulting treatment effect estimates as causal depends on whether we are willing to rely on empirical fit and parametric assumptions as a basis for causal inference.

In contrast, in the principal stratification approach (Frangakis and Rubin, 2002), which we intend to combine with growth mixture modeling, the focus has been given to clarification of assumptions necessary to identify causal effects. Principal stratification refers to cross-classification of individuals on the basis of potential values of post-treatment variables under all treatment conditions that are compared. Since the resulting categories (principal strata) are unaffected by treatment (just like pre-treatment covariates), treatment effects conditioning on the principal stratum membership can be interpreted as causal effects. Consideration of potential outcomes facilitates the use of clear identifying assumptions as a basis for causal inference, which is a strong advantage of the principal stratification approach. Principal stratification is mostly conducted in terms of intermediate post-treatment outcome variables. The two key elements in identifying causal effects in the principal stratification framework are 1) pre-determined (i.e., not based on empirical fitting) rules that classify individuals into principal strata and 2) partially observed potential intermediate variable values. For example, in a classic example of principal stratification presented in Angrist et al., (1996), four potential

compliance types are pre-defined based on the treatment assignment status (0 = control; 1 = treatment) and the potential treatment receipt status (0 = would not receive; 1 = would receive). In this framework, we observe the potential treatment receipt status of individuals in the data at least under the condition they are actually assigned to. If an individual assigned to the treatment condition does not receive the treatment, we know that he or she belongs either to the never-taker or to the defier category. Given this partially observed stratum information and pre-determined classification rules, explicit assumptions can be established to make up for missing information and to identify causal effects. Since principal strata are formulated on the basis of pre-determined rules and partially observed potential intermediate variable values, empirical model fitting is not necessary. However, for the same reason, the resulting causal effect estimation models may not conform with the data well.

Although the principal stratification and the mixture modeling approaches share their common interest in heterogeneous subpopulations and share their discrete perspective in characterizing heterogeneity, the two approaches have been developed in parallel and there has been relatively little formal discussion of analysis strategies that combine the strengths of the two. In this paper, we focus on causal inference in the presence of heterogeneity in longitudinal outcome prognosis, which brings up a motivating setting for the consideration of the two approaches jointly. In the principal stratification approach, identification of causal effects hinges on partially observed stratum membership information. In contrast, in the growth mixture modeling framework, since we are not subject to a certain pre-determined stratification rule, stratum (latent class) membership cannot be predetermined for any individual based on the observed data. Instead, trajectory strata are determined based on the results of empirical fitting (substantive theory or experts opinion may play a role here in making a choice among empirical solutions). Growth mixture modeling is an effective way of estimating trajectory classes. Further, using trajectory strata formulated based on empirical fitting is likely to result in better fitting causal effect estimation models. From the growth mixture modeling approach's perspective, adopting the principle of the principal stratification approach is crucial to gain ability to clarify assumptions necessary to identify causal effects and to examine sensitivity of the causal effect estimates to deviations from these identifying assumptions.

The two-step approach we propose here can be thought of as a way of improving causal modeling in the growth mixture modeling framework by incorporating the principal stratification approach, or as a way of assisting formulation of principal strata and improving empirical fit in the principal stratification framework by incorporating the growth mixture modeling approach. In Step 1, heterogeneous latent trajectory classes are identified through growth mixture analyses in one of the assignment conditions (reference condition). In Step 2, treating the latent classes obtained in Step 1 as known for the reference condition (and missing for the rest), causal treatment effects are defined and identified. The main reason for this separation by step-wise analysis is to not involve exploratory components in the identification of causal treatment effects. Therefore, the way in which causal effects are identified in Step 2 is consistent with that in the standard principal stratification approach. The results can be interpreted as how subpopulations that differ in terms of outcome prognosis under one treatment condition would change their prognosis differently when exposed to another treatment condition.

This paper is organized as follows. In Section 2, the data from the Johns Hopkins School Intervention Trial (Ialongo et al., 1999) will be introduced. The overall intention to treat (ITT) effect will be estimated using a common linear mixed effects model. In Section 3, the proposed two-step approach will be described and demonstrated using the Johns Hopkins data. Section 4 provides conclusions.

2. JOHNS HOPKINS PIRC STUDY

The data that motivated our study are from a school intervention trial conducted by the Johns Hopkins University Preventive Intervention Research Center (hereafter PIRC) in 1993–1994 (Ialongo et al., 1999). The PIRC study was designed to improve academic achievement and to reduce early behavioral problems. Two programs were employed in the study: the Classroom-Centered (CC) intervention and the Family-School Partnership intervention. Teachers and first-grade children were randomly assigned to the control and intervention conditions. In this paper, we will compare the control and the CC intervention groups. The CC intervention consisted of components such as curriculum enhancements, enhanced behavior management strategies, and back-up strategies for children not performing adequately. The intervention was provided over the first-grade school year, following a baseline assessment in the early fall. The control condition did not consist of any intervention activities (i.e., standard curriculum).

We will focus on attention deficit measured by the TOCA-R (Teacher Observation of Classroom Adaptation-Revised; Werthamer-Larsson, Kellam, and Wheeler, 1991) as the outcome. The TOCA-R was designed to assess children's adequacy of performance on core tasks in the classroom as rated by the teacher. The attention deficit scale ranges from 1 to 6 and consists of TOCA-R items that measure hyperactivity, concentration problem, and impulsiveness. The outcome was measured in the early fall (Y_1 : baseline), in the spring of first (Y_2 : 6 months followup) and second (Y_3 : 18 months followup) grades. Among 378 (194 intervention, 184 control) cases analyzed, 95% completed the assessment in the spring of first grade, and 82% in the spring of second grade. Our analyses will be conducted under the missing at random (MAR; Little and Rubin, 2002) assumption using the likelihood of the observed data. However, more sophisticated modeling of missing data mechanisms considering the stratum membership is also possible (e.g., Frangakis and Rubin, 1999; Jo, 2008; Mealli et al., 2004). Covariates included in the analyses are child's gender (1 = male, 0 = female), ethnicity (1 = African-American, 0 = other), parent's education (1 = some education beyond high school, 0 = high school or less), marital status (1 = married, 0 = not married), and employment status (1 = employed, 0 = not employed).

Table 1 shows the sample statistics of the variables used for the analyses in this paper. No significant differences between the control and intervention groups were found in the baseline covariates. However, the baseline attention deficit score shows a significant group difference. Given that classrooms were randomly assigned to the two groups (9 classrooms to the intervention condition, 9 classrooms to the control condition), there was no apparent reason for the significant difference between the groups in attention deficit at baseline. Assuming that this is an experimental error occurred by chance, we will focus on how attention deficit changes over time controlling for this baseline difference. Intervention effects will be assessed by the rate of change in attention deficit using both mixed effects model and latent growth mixture model analyses.

2.1 The overall average treatment effect

To estimate the overall intervention effect, we employed a common growth model, where a single population (i.e., one latent class) is assumed. We used maximum likelihood (ML) estimation implemented in Mplus (Muthén and Muthén, 1998–2009). As mentioned earlier, the analysis was conducted assuming that data are missing at random (MAR; Little and Rubin, 2002) conditional on observed information. A quadratic growth model was used given a nonlinear development of attention deficit on average from the fall of the first grade to the spring of the second grade, as shown in the sample statistics in Table 1. According to the likelihood ratio test, the model fit the data significantly better in the presence of the quadratic growth parameter. The quadratic growth model we employed is described below.

The outcome Y for individual i at time point t , where $i = 1, 2, \dots, N$, and $t = 1, 2, \dots, T$ (in the PIRC example, $T = 3$), can be expressed as

$$y_{it} = \eta_{li} + \eta_{li} W_t + \eta_{qi} W_t^2 + \varepsilon_{it}, \quad (1)$$

$$\eta_{li} = \alpha_l + \lambda_l x_i + \zeta_{li}, \quad (2)$$

$$\eta_{li} = \alpha_l + \lambda_l x_i + \gamma_l Z_i + \zeta_{li}, \quad (3)$$

$$\eta_{qi} = \alpha_q + \lambda_q x_i + \gamma_q Z_i + \zeta_{qi}, \quad (4)$$

where

$$\begin{aligned} \varepsilon_i &\sim MN(0, \Sigma_\varepsilon), \\ \zeta_i &\sim MN(0, \Sigma_\zeta), \end{aligned}$$

and ε_i is independent of ζ_i for all i . Note that this independence assumption is necessary for SUTVA (stable unit treatment value; Rubin, 1978, 1980, 1990) to hold, which is explained below. That is, if SUTVA holds, then independence among $\{Y_i: i = 1, \dots, n\}$ should not be violated. Since $T = 3$, we imposed a restriction that Σ_ε and Σ_ζ are diagonal to identify all parameters in Σ_ε and Σ_ζ .

The linear mixed effects model described in (1)–(4) include three random effects: the initial status (η_{li}), the linear growth (η_{li}), and the quadratic growth (η_{qi}). The set of time scores W_t reflects the linear growth (e.g., 0, 1, 3 for the baseline, 6 months, 18 months). The set of time scores W_t^2 reflects the quadratic growth (0, 1, 9). The residual ε_{it} is assumed to be normally distributed with mean zero, and its variance is allowed to vary over time. All five covariates presented in Table 1 were included in the analysis as predictors of the three random effects. The relationship between the random effects and the vector of covariates x is captured by the vectors of regression coefficients λ_l , λ_L , and λ_Q . The intercepts in (2)–(4) can be interpreted as the mean initial status (α_l), linear growth (α_L), and quadratic growth (α_Q) of the control group if the covariates are centered at their means in the control condition. The additional linear growth γ_L and the additional quadratic growth γ_Q when assigned to the intervention condition ($Z = 1$) instead of to the control condition ($Z = 0$) are interpreted as the treatment assignment effects. Under the assumption of random assignment to the treatment conditions, there is no effect of treatment assignment on the initial status. The residuals ζ_{li} , ζ_{Li} , and ζ_{Qi} are assumed to be normally distributed.

To be able to interpret the estimates of γ_L and γ_Q in (3)–(4) as the overall causal effect (i.e., ITT effect) estimates, a couple of identifying assumptions are necessary. These assumptions make it possible to model potential outcomes using observed variables, and therefore make it possible to causally interpret the results. The following two assumptions are widely used basic assumptions in the potential outcomes approach (e.g., Angrist, Imbens, and Rubin,

1996; Frangakis and Rubin, 2002; Holland, 1986; Neyman, 1923; Rubin, 1974, 1978, 1980, 2005) and are slightly modified here to accommodate the longitudinal setting we consider.

- Ignorable treatment assignment: It assumes that treatment assignment is independent of the potential outcomes. That is, $(Y_{it}(1), Y_{it}(0)) \perp Z_i$, where $Y_{it}(1)$ is the potential outcome for individual i at time t when assigned to the treatment condition ($Z = 1$) and $Y_{it}(0)$ when assigned to the control condition ($Z = 0$). This assumption is satisfied in randomized experiments, such as PIRC, when covariates x are balanced among all treatment conditions. This assumption can be relaxed to $(Y_{it}(1), Y_{it}(0)) \perp Z_i | x_i$.
- Stable unit treatment value (SUTVA): It is assumed that the potential outcomes for each person are unaffected by the treatment assignment of other individuals (Rubin, 1978, 1980; Rubin, 1990). In the PIRC example, the unit of randomization was a classroom. Therefore, it seems reasonable to assume that the level of interaction among individuals across different treatment conditions remains about the same as that observed when the unit of randomization is an individual (Sobel, 2006). However, interaction in the same classroom is likely. Therefore, standard errors of causal effect estimates could be underestimated in our analyses. In principle, it is possible to take into account this interaction within each treatment assignment arm in the principal stratification context (Frangakis, Rubin, and Zhou, 2002; Jo, Asparouhov, and Muthén, 2008), although the quality of standard error estimates tend to decline with small numbers of clusters (e.g., 9 classrooms in each arm in PIRC). In terms of the repeatedly measured outcome, we assume independence among subjects, but allow for the dependence across measures within subjects.

Table 2 shows the overall intervention effect estimates based on the linear mixed effects model described in (1)–(4). These estimates can be interpreted as the average causal effect estimates under the random assignment of treatment and SUTVA. The effect of intervention assignment on the linear growth (γ_L) is significant, although the size of the effect is moderate considering the outcome standard deviation of around one (see Table 1). No significant effect of intervention assignment on the quadratic growth (γ_Q) was found, implying similar rates of acceleration in growth between the intervention and control groups. Marital status was a significant predictor of initial status. Children of single parents showed a higher level of initial attention deficit. None of the covariates were significant predictors of either the linear or the quadratic growth rate.

The estimation results are also presented in Figure 1 in terms of average outcome trajectories. These estimated trajectories show that the difference in outcome prognosis between the intervention and control groups is quite small, although significant. Note that these results are based on an analysis assuming a single population. However, in line with the Simpson's paradox argument, the overall intervention effect estimates based on the standard linear mixed effects model could be misleading or not too informative. One of the main questions in the PIRC trial was if the intervention program had a concentrated effect on certain subpopulations, in particular subpopulations with developmentally problematic trajectory patterns. In the following section, we will examine whether the data indicate the existence of subpopulations with distinct outcome trajectory types, and if so, whether the intervention would work differently for these different trajectory classes.

3. TWO-STEP GROWTH MIXTURE APPROACH

In this study, we propose a 2-step approach, where formulation of trajectory strata and identification of causal effects are separated. The first step utilizes the exploratory nature of growth mixture modeling to formulate trajectory strata that will result in well-fitting causal effect estimation models for the second step. However, exploratory components are eliminated

from the second step to avoid identification of causal treatment effects relying on empirical model fitting.

In most principal stratification models, strata are formulated based on partially observed potential intermediate variable values. Hinging on this observed information, assumptions necessary to identify principal causal effects can be established considering potential outcome values under all treatment conditions that are compared. However, this approach does not directly apply to a situation where outcome trajectory types are the strata of interest because trajectory strata are not readily observable. Growth mixture modeling is an effective way of generating strata information in this situation. In principle, formulating trajectory strata and estimating treatment effects for these strata can be carried out in one step, which is convenient and computationally parsimonious. The drawback of this one-step approach is that identification of treatment effects relies on empirical model fitting and parametric assumptions, which is not desirable from the perspective of causal inference. To avoid this situation, it is critical to exclude estimation of treatment effects from the exploratory analysis process, which can be done by handling each treatment assignment arm data separately when formulating trajectory strata. For this reason, the proposed 2-step approach is based on a stratification scheme that considers the potential values of the outcome under only one assignment condition. We use reference stratification to refer to this stratification strategy.

3.1 Reference stratification

In this paper, we will employ a stratification strategy called reference stratification, where individuals are stratified according to their potential post-treatment variable values under a particular treatment condition of interest (i.e., reference condition). The resulting reference strata are not affected by treatment assignment, and therefore the differences in the outcome across treatment groups within reference strata can be interpreted as causal effects. Given that principal stratification (Frangakis and Rubin, 2002) refers to cross-classification of individuals on the basis of potential values of post-treatment variables under all treatment conditions that are compared, reference stratification can be thought of as a special case of principal stratification.

The idea of stratifying individuals on the basis of potential outcome values under only one treatment condition has been utilized previously in the context of principal stratification modeling (Jin and Rubin, 2008; Joffe, Small, and Hsu, 2007; Roy, Hogan, and Marcus, 2008). The common motivation for using reference stratification in these applications was that the potential intermediate variable values (therefore, principal strata or coarse forms of principal strata) are partly observed. Reference stratification in this context simply means grouping of theoretically agreed-upon principal strata into coarser strata that are observed in one of the assignment conditions. While trajectory strata are of interest as in this study, we do not have pre-determined rules for defining principal strata. Instead, we need to empirically identify reference strata under each treatment condition first, and if necessary, construct principal strata using a pair or multiple sets of reference strata. In other words, reference stratification in this context works as a basis for deriving principal strata, not as a way of regrouping already defined principal strata. In this paper, we focus on causal inference based on a single set of reference strata. Although cross-classification using multiple sets of reference strata is more in line with the idea of principal stratification, it is unclear in practice how this method performs and how the estimation results should be interpreted, especially when there are no pre-determined rules that classify individuals into reference or principal strata.

3.2 Step 1: Exploratory growth mixture analysis to identify reference strata

In Step 1, heterogeneous trajectory classes are identified using growth mixture analyses in one of the assignment conditions (reference condition). Step 1 is an exploratory procedure utilizing

empirical model fitting, and therefore the resulting trajectory strata solutions are more likely to fit the data well than the solutions determined by external (or ad hoc) rules. In principle, it is possible to classify individuals without involving latent class type analyses, for example, by dividing individuals into low and high groups based on the median baseline score. However, this strategy can be arbitrary and inefficient, in particular when classifying individuals based on longitudinal outcome data.

One of the main questions in the PIRC trial was whether there are developmentally distinct subpopulations, and if so, whether the intervention would work differently for these different trajectory classes. To examine this, we first conducted a series of exploratory analyses separately for the control group and the treatment group to see whether the data indicate the existence of subpopulations with distinct outcome trajectory types. Let $C(0)$ denote the potential trajectory strata (i.e., reference strata) membership under the control condition, and $C(1)$ under the intervention condition.

The observed outcome Y for individual i in the control condition ($i = 1, 2, \dots, N_0$ with $N_0 = 184$ in PIRC) and stratum j at time point t ($t = 1, 2, \dots, T$ with $T = 3$ in PIRC) is now expressed as

$$y_{it} = \eta_{i0ij} + \eta_{L0ij} W_t + \eta_{Q0ij} W_t^2 + \varepsilon_{0ijt}, \quad (5)$$

$$\eta_{i0ij} = \alpha_{i0j} + \lambda_{i0} x_i + \zeta_{i0i}, \quad (6)$$

$$\eta_{L0ij} = \alpha_{L0j} + \lambda_{L0} x_i + \zeta_{L0i}, \quad (7)$$

$$\eta_{Q0ij} = \alpha_{Q0j} + \lambda_{Q0} x_i + \zeta_{Q0i}, \quad (8)$$

where there are J_0 possible trajectory strata ($j = 1, 2, \dots, J_0$) and $(\varepsilon_{0ij1}, \dots, \varepsilon_{0ijT}) \sim N(0, \Sigma_j)$. The three random effects (η_{i0ij} , η_{L0ij} , η_{Q0ij}) that capture outcome trajectories under the $Z = 0$ condition (i.e., control) are allowed to vary across J_0 classes.

The probability π_{0i} of belonging to a certain latent class j varies depending on the influence of covariates. The multinomial logit model of π_{0i} with a vector of covariates x is described as

$$\text{logit}(\pi_{0i}|x_i) = \beta_{00} + \beta_{10} x_i, \quad (9)$$

where π_{0i} is a $J_0 - 1$ dimensional vector of $(\pi_{0i1}, \pi_{0i2}, \dots, \pi_{0i(J_0-1)})'$, $\pi_{0ij} = \Pr(C_i(0) = j|x_i)$, and $\text{logit}(\pi_{0i}) = (\log \times [\pi_{0i1}/\pi_{0iJ_0}], \log[\pi_{0i2}/\pi_{0iJ_0}], \dots, \log[\pi_{0i(J_0-1)}/\pi_{0iJ_0}])'$. The vector β_{00} is a $J_0 - 1$ dimensional vector of logit intercepts, and β_{10} is a $J_0 - 1$ dimensional vector of multinomial logit regression coefficients.

To obtain the maximum likelihood estimates (MLE's) for the growth mixture model described in (5)–(9) using subjects in the control group, we employed the EM algorithm (Dempster, Laird, and Rubin, 1977; Little and Rubin, 2002; McLachlan and Krishnan, 1997; Tanner, 1996)

implemented in the *Mplus* program (Muthén and Muthén, 1998–2009). From (5)–(9), the log-likelihood for the observed data $\{y_i; i = 1, \dots, N_0\}$ is

$$\sum_{i=1}^{N_0} \log \left\{ \sum_{j=1}^{J_0} \pi_{0ij}(\beta_j) f(y_i | x_i, \alpha_j, \lambda_j, \Sigma_j) \right\}, \quad (10)$$

where $\pi_{0ij}(\beta_j) = \Pr(C_i(0) = j | x_i, \beta_j)$ denotes the likelihood that y_i arising from mixture component j given x_i , and $\sum_{j=1}^{J_0} \pi_{0ij}(\beta_j) = 1$. And the log-likelihood of the complete-data (i.e., $\{y_i, C_i(0); i = 1, \dots, N_0\}$) can be written as

$$\begin{aligned} \text{Log} L_c &= \sum_{i=1}^{N_0} \text{Log} L_{ci} \\ &= \sum_{i=1}^{N_0} \sum_{j=1}^{J_0} \sum_{C_i(0)=j} [(\log[f(y_i | x_i, \alpha_j, \lambda_j, \Sigma_j)] \\ &\quad + \log[\pi_{0ij}(\beta_j)])]. \end{aligned} \quad (11)$$

To maximize (10), the E step computes the expected values of the log-likelihood (11) given observed data and $(\beta^*, \alpha^*, \lambda^*, \Sigma^*)$, the current current parameter estimates. Latent trajectory class $C(0)$ is considered as missing data in this step. That is, the E step computes

$$\sum_{i=1}^{N_0} \sum_{j=1}^{J_0} \log[\pi_{0ij}(\beta_j^*) f(y_i | x_i, \alpha_j^*, \lambda_j^*, \Sigma_j^*)] p_{0ij}(\beta_j^*),$$

where $p_{0ij}(\beta_j^*) = \pi_{0ij}(\beta_j^*) / [\sum_{j'=1}^{J_0} \pi_{0ij}(\beta_{j'}^*) f(y_i | x_i, \alpha_{j'}^*, \lambda_{j'}^*, \Sigma_{j'}^*)]$ being the posterior class probability. The M step computes the parameter estimates that maximize the quantity obtained from the E step. This maximization procedure continues until it reaches the optimal status.

The number of latent classes and the specific model choice can be decided based on two sources of information. First, we may utilize model fit indices such as Bayesian Information Criteria (BIC; Schwarz, 1978). Second, we may incorporate experts' opinion (theory) to decide on substantively meaningful classes in the sense that it is of substantive/clinical interest to find out differential effects of the treatment for these outcome trajectory types.

From the exploratory growth mixture analyses using the control condition data, a 2-class solution was chosen as the most parsimonious solution in the PIRC trial according to model fit and expert's opinion (BIC: 1-class = 1265; 2-class = 1240; 3-class = 1241). Figure 2 shows the results of the 2-class growth mixture analysis ($N_0 = 184$). The 2-class solution presented in Figure 2 showed a pronounced division of the two classes with 96% classified with the estimated posterior class probability (p_{0ij}) of over 0.9 or below 0.1. We are particularly interested in this solution from the substantive point of view because the two trajectories well represent a normative development of attention deficit among the majority of children at this age range (class 2: normative), and a more problematic development among much fewer children (class 1: problematic). This solution may also carry an important treatment implication if we can find out how the intervention would alter these two trajectory classes.

We also conducted growth mixture analyses for the intervention group. Unlike in the control condition, no clear division of heterogeneous trajectory classes was detected (BIC: 1-class = 1453; 2-class = 1461; 3-class = 1465). Further, in comparing the standard curriculum and the proposed intervention program, what is practically meaningful to find out is how subpopulations that differ in terms of outcome prognosis under the control condition would change their prognosis differently when exposed to the intervention condition (instead of how subpopulations that differ under the intervention condition would change their prognosis in the absence of the intervention). On the basis of empirical fit and possible treatment implications, the control condition was chosen as reference condition. Further analyses in this paper are conducted focusing on the latent trajectory types under the control condition.

3.3 Step 2: Identification of differential average causal effects

Once we determine the latent class solution under each treatment condition, we can formulate theoretical reference strata based on the estimated posterior class probabilities. In the analyses of the PIRC trial, two reference strata were identified under the control condition, and a single reference stratum was identified under the treatment condition. Note that, in Step 1, the latent trajectory class membership $C(0)$ is estimated only for individuals assigned to the control condition, while $C(1)$ is estimated only for individuals assigned to the intervention condition. Since our interest is in the reference strata under the control condition, and the posterior probabilities of $C(0)$ were already obtained from Step 1 among individuals assigned to the control condition (still unknown among those assigned to the intervention condition), causal effect estimation in Step 2 is conducted under a relatively simple missing data situation.

The outcome Y for individual i in the entire sample ($i = 1, 2, \dots, N$ with $N = 378$ in PIRC) at time point t ($t = 1, 2, \dots, T$ with $T = 3$ in PIRC) is expressed as

$$y_{it} = \eta_{10ij} + \eta_{L0ij} W_t + \eta_{Q0ij} W_t^2 + \varepsilon_{0it}, \quad (12)$$

$$\eta_{10ij} = \alpha_{10j} + \lambda_{10} x_i + \zeta_{10i}, \quad (13)$$

$$\eta_{L0ij} = \alpha_{L0j} + \lambda_{L0} x_i + \gamma_{L0j} Z_i + \zeta_{L0i}, \quad (14)$$

$$\eta_{Q0ij} = \alpha_{Q0j} + \lambda_{Q0} x_i + \gamma_{Q0j} Z_i + \zeta_{Q0i}, \quad (15)$$

where J_0 is now the number of reference strata that are fixed at 2 (i.e., $j = 1, 2$) based on the results from Step 1.

The probability π_{0i} of belonging to a certain reference stratum j varies depending on the influence of covariates. The multinomial logit model of π_{0i} with a vector of covariates x is described as

$$\text{logit}(\pi_{0i}) = \beta_{00} + \beta_{10} x_i, \quad (16)$$

where π_{0i} is now known among individuals assigned to the control condition (from Step 1), although is still unknown among individuals assigned to the intervention condition.

The model described in (12)–(16) is different from the model described in (5)–(9) only in terms of the additional treatment effect parameters γ_{L0j} and γ_{Q0j} . In this model, not only the three random effects (η_{I0ij} , η_{L0ij} , η_{Q0ij}), but also the treatment effects (γ_{L0j} , γ_{Q0j}) are allowed to vary across J_0 strata. Note that the growth mixture model described in (12)–(16) is applied to the combined sample (control + treatment) whereas the model described in (5)–(9) in Step 1 is applied only to the control condition sample.

To causally interpret estimates of treatment assignment effects conditional on the potential outcome trajectory strata (i.e., γ_{L0j} and γ_{Q0j}), the assumptions of ignorable treatment assignment and SUTVA are slightly modified as follows.

- Ignorable treatment assignment: It is assumed that treatment assignment is independent of the potential outcomes and potential outcome trajectory types. That is, $(Y_{it}(1), Y_{it}(0), C_i(1), C_i(0)) \perp Z_i$. This assumption can be relaxed to $(Y_{it}(1), Y_{it}(0), C_i(1), C_i(0)) \perp Z_i | x_i$, which is more practical for randomized studies in general (see Lin, Ten Have, and Elliott, 2008).
- Stable unit treatment value (SUTVA): It is assumed that the potential outcomes and potential outcome trajectory types for each person are unaffected by the treatment assignment of other individuals.

The assumptions of ignorable treatment assignment and SUTVA are not sufficient to identify average causal treatment effects that vary across trajectory strata even with observed stratum membership among individuals assigned to the control condition. To achieve identifiability without relying on parametric assumptions, we considered two identifying assumptions described below. We considered both assumptions to check sensitivity of the results, although only one assumption is necessary.

- Exclusion restriction (ER): It is assumed that there is no effect of treatment assignment for those who would follow the normative outcome trajectory when assigned to the control condition (i.e., units with $C_i(0) = 2$). In the PIRC example, the level of attention deficit for this class is already very low and increases very slowly over time. Therefore, it is unlikely that the intervention will have any considerable effect on this trend. This assumption imposes restrictions that $\gamma_{L02} = 0$ and $\gamma_{Q02} = 0$ in (14) and (15).
- Additive treatment assignment effect: It is assumed that the effect of treatment assignment does not vary across different values of covariates (e.g., Jo, 2002). In the PIRC example with two trajectory strata, this assumption allows us to estimate main effects of treatment assignment for both reference strata (i.e., ER can be relaxed). However, the assumption may be violated in terms of some of the covariates.

With the help of identifying assumptions such as the exclusion restriction and additivity, causal treatment effects are identified treating the latent classes estimated in Step 1 as known for individuals assigned to the reference condition (i.e., control condition) and unknown for the rest. The same ML-EM procedure used in Step 1 is used treating the unknown stratum membership as missing data. One problem here is that, from the estimation in Step 1, most individuals are likely to have posterior class probability estimates that are not exactly 0 or 1. Another problem in using these posterior class probabilities is that, by ignoring their uncertainty, standard errors of the causal effect estimates may be underestimated in the subsequent causal effect estimation. As a simple solution to both problems, we employed pseudo class draws (Bandein-Roche et al., 1997; Wang, Bandein-Roche, and Brown, 2005). For each pseudo class draw, control condition individuals are stratified into two reference

strata. On the basis of each stratification, causal treatment effects are identified and estimated. Point estimates and standard error estimates are averaged over 20 pseudo class draws.

Table 3 shows the results of causal effect estimation from Step 2. The results can be interpreted as how subpopulations that differ in terms of outcome prognosis under the control condition would differently change their prognosis when exposed to the intervention condition. The results reported in Table 3 were obtained relying on good predictors of stratum membership and the additivity assumption. The analysis assuming ER yielded very similar results (not reported here). For those who would maintain a quite high level of attention deficit from grade 1 to grade 2 under the control condition (class 1), assignment to the intervention condition had considerable and significant effects on their outcome trajectory, both in terms of the linear (γ_{L01}) and the quadratic growth (γ_{Q01}). As expected, for those who would develop a normative growth (class 2) under the control condition, the intervention had little impact both on the linear (γ_{L02}) and the quadratic growth (γ_{Q02}). Marital status was a significant predictor of initial status. Children of single parents showed a higher level of initial attention deficit. Child's gender was a significant predictor of the trajectory strata. Boys were significantly more likely to develop a problematic outcome trajectory (odds ratio = 3.84).

Figure 3 shows estimated average outcome trajectories from Step 2. The estimated trajectories show that individuals in the problematic trajectory stratum (14% of the sample) benefited the most from the PIRC intervention during the first six months of the study. Although the level of attention deficit did not decrease any further, at least it stayed at the same level for an additional year after the intervention. In contrast, being assigned to the intervention condition had little impact on the outcome development of individuals in the normative trajectory stratum throughout the study.

3.4 Comparison of the one and two-step approaches

To compare the results from one- and two-step approaches, we conducted a one-step growth mixture analysis with the entire sample. As discussed earlier, in the onestep approach, both the treatment effects and latent trajectory classes are identified on the basis of empirical fit and parametric assumptions (Muthén et al., 2002; Muthén and Brown, 2009). The same model employed for Step 2 in the 2-step approach described in (12)–(16) was employed for the one-step approach. The same ML-EM estimation method was used. The difference from Step 2 in the two-step approach is that the estimation is now conducted neither with partly known class membership nor with clear identifying assumptions. Given that, the one-step approach is exploratory, and therefore, the model choice was made by comparing empirical fit across a few different solutions. A 2-class solution was chosen based on the fact that increasing the number of classes no longer improved the fit (BIC: 1-class = 2678; 2-class = 2641; 3-class = 2642).

Figure 4 shows estimated average outcome trajectories from the one-step growth mixture analysis. According to this analysis, the problematic trajectory class (Class 1) had a somewhat higher proportion (19%) compared to the proportion estimated in the 2-step approach (14%). Overall, the 2-class exploratory solution is strikingly similar to that of the 2-step analysis solution (see Figure 3). We find this consistency quite remarkable given that trajectory strata and treatment effects for these strata are identified following considerably different principles in the two approaches. We interpret these results as that both solutions conform well with the empirical data. One explanation for this similar empirical fit would be that only one reference stratum is empirically identified under the treatment condition, and therefore the two trajectory strata identified from the control condition is sufficient to capture heterogeneity in the entire sample. Further investigation is necessary to clarify the connection between the two approaches and to clarify how causal modeling can benefit from that connection.

4. CONCLUSIONS

By combining the strengths of growth mixture modeling and principal stratification in the 2-step approach, this study showed that it is possible to consider both empirical fit and causal interpretability when studying heterogeneity in outcome prognosis. We employed reference stratification to formulate outcome trajectory strata based on empirical fitting without involving treatment response. In particular, we focused on causal inference based on a single set of reference strata, which may be the main interest in some studies. For example, in the JHU PIRC example, it was considered very beneficial to understand how subpopulations that differ in terms of outcome prognosis under the control condition would differently change their prognosis when exposed to the treatment. In other studies, it may be critical to cross-classify individuals based on potential prognosis under both conditions. In principle, each treatment condition can have its own reference strata, and therefore it is possible to cross-classify individuals based on multiple sets of reference strata. However, little is known about how this method performs and how the estimation results should be interpreted under various conditions. It is also unclear how one should evaluate the performance of the method when we do not even have pre-determined classification rules and therefore we do not have any clear picture of the true principal strata. Extensive research is needed to establish well-structured inferential and analytical framework that can accommodate subpopulations derived from empirical model fitting.

In the current paper, we focused on situations where the same longitudinal variable is used both as the basis for stratification and as the outcome, and therefore the resulting strata are likely to well reflect distinct outcome distributions. However, in most principal stratification applications, the intermediate variable on which stratification is based is not the same as the outcome. For example, when principal stratification is conducted in terms of treatment compliance (e.g., Angrist et al., 1996), the resulting principal strata solution may not agree well with the outcome data, resulting in ill-fitting causal effect estimation models. Considering latent class type analyses and empirical fit focusing on one of the two (the outcome or the intermediate outcome) may somewhat improve this situation. For example, in Jo and Muthén (2003), mixture analyses were conducted in terms of the longitudinal outcome to locate cut-points that discretize the compliance variable (which was originally continuous). In Lin, Ten Have, and Elliot (2008), latent class analyses were conducted to summarize longitudinal trends of compliance patterns. In these studies, however, principal strata solutions may not optimally reflect distinct subpopulations in terms of outcome distributions because formulation of strata does not fully consider both the outcomes and intermediate outcomes. Further investigation is necessary to extend the proposed method to handle heterogeneity in both.

References

- Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 1996;91:444–455.
- Bandeem-Roche K, Miglioretti DL, Zeger SL, Rathouz PJ. Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association* 1997;92:1375–1386.
- Dempster A, Laird N, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B* 1977;39:1–38.
- Diggle, PJ.; Liang, KY.; Zeger, SL. *The analysis of Longitudinal Data*. Oxford, England: Oxford University Press; 1994.
- Frangakis CE, Rubin DB. Addressing complications of intention-to-treat analysis in the presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika* 1999;86:365–379.
- Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics* 2002;58:21–29. [PubMed: 11890317]

- Frangakis CE, Rubin DB, Zhou XH. Clustered encouragement design with individual noncompliance: Bayesian inference and application to advance directive forms. *Biostatistics* 2002;3:147–164. [PubMed: 12933609]
- Holland PW. Statistics and causal inference. *Journal of the American Statistical Association* 1986;81:945–960.
- Ialongo NS, Werthamer L, Kellam SG, Brown CH, Wang S, Lin Y. Proximal impact of two first-grade preventive interventions on the early risk behaviors for later substance abuse, depression and antisocial behavior. *American Journal of Community Psychology* 1999;27:599–642. [PubMed: 10676542]
- Jin H, Rubin DB. Principal stratification for causal inference with extended partial compliance. *Journal of the American Statistical Association* 2008;103:101–111.
- Jo B. Bias Mechanisms in intention-to-treat analysis with data subject to treatment noncompliance and missing outcomes. *Journal of Educational and Behavioral Statistics* 2008;33:158–185.
- Jo B, Asparouhov T, Muthén BO, Ialongo NS, Brown CH. Cluster randomized trials with treatment noncompliance. *Psychological Methods* 2008;13:1–18. [PubMed: 18331150]
- Jo, B.; Muthén, BO. Longitudinal Studies with Intervention and Noncompliance: Estimation of Causal Effects in Growth Mixture Modeling. In: Duan, N.; Reise, S., editors. *Multilevel Modeling: Methodological Advances, Issues, and Applications*. Multivariate Applications Book Series, Lawrence Erlbaum Associates; 2003. p. 112–139.
- Jo B. Estimating intervention effects with noncompliance: Alternative model specifications. *Journal of Educational and Behavioral Statistics* 2002;27:385–420.
- Joffe MM, Small D, Hsu CY. Defining and estimating intervention effects for groups that will develop an auxiliary outcome. *Statistical Science* 2007;22:74–97.
- Jones BL, Nagin DS, Roeder K. A SAS procedure based on mixture models for estimating developmental trajectories. *Sociological Methods and Research* 2001;29:374–393.
- Laird NM, Ware JH. Random-effects Models for Longitudinal Data. *Biometrics* 1982;38:963–974. [PubMed: 7168798]
- Lin JY, Ten Have TR, Elliot MR. Longitudinal nested compliance class model in the presence of time-varying non-compliance. *Journal of the American Statistical Association* 2008;103:462–473.
- Little, RJA.; Rubin, DB. *Statistical analysis with missing data*. New York: Wiley; 2002.
- Mclachlan, GJ.; Krishnan, T. *The EM algorithm and extensions*. New York: Wiley; 1997.
- Meredith W, Tisak J. Latent curve analysis. *Psychometrika* 1990;55:107–122.
- Mealli F, Imbens GW, Ferro S, Biggeri A. Analyzing a randomized trial on breast self-examination with noncompliance and missing outcomes. *Biostatistics* 2004;5:207–222. [PubMed: 15054026]
- Muthén, BO. Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In: Kaplan, D., editor. *Handbook of quantitative methodology for the social sciences*. Newbury Park, CA: Sage Publications; 2004. p. 345–368.
- Muthén BO, Brown CH. Estimating drug effects in the presence of placebo response: Causal inference using growth mixture modeling. 2009 Under review.
- Muthén BO, Brown CH, Masyn K, Jo B, Khoo ST, Yang CC, Wang CP, Kellam S, Carlin J, Liao J. General growth mixture modeling for randomized preventive interventions. *Biostatistics* 2002;3:459–475. [PubMed: 12933592]
- Muthén BO, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics* 1999;55:463–469. [PubMed: 11318201]
- Muthén, LK.; Muthén, BO. *Mplus user's guide*. Los Angeles: Muthén & Muthén; 1998–2009.
- Nagin DS. Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods* 1999;4:139–157.
- Neyman J. On the application of probability theory to agricultural experiments. Essay on principles. Section 9 translated. *Statistical Science* 1923;5:465–480. 1990.
- Raudenbush, SW.; Bryk, AS. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, CA: Sage; 2002.
- Roy J, Hogan JW, Marcus BH. Principal stratification with predictors of compliance for randomized trials with 2 active treatments. *Biostatistics* 2008;9:277–289. [PubMed: 17681993]

- Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 1974;66:688–701.
- Rubin DB. Bayesian inference for causal effects: The role of randomization. *Annals of Statistics* 1978;6:34–58.
- Rubin DB. Discussion of “Randomization analysis of experimental data in the Fisher randomization test” by D. Basu. *Journal of the American Statistical Association* 1980;75:591–593.
- Rubin DB. Comment on Neyman (1923) and causal inference in experiments and observational studies. *Statistical Science* 1990;5:472–480.
- Rubin DB. Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association* 2005;100:322–331.
- Schwarz GE. Estimating the dimension of a model. *Annals of Statistics* 1978;6:461–464.
- Sobel ME. What do randomized studies of housing mobility demonstrate: Causal inference in the face of interference. *Journal of the American Statistical Association* 2006;101:1398–1407.
- Tanner, M. *Tools for statistical inference: Methods for the exploration of posterior distributions and likelihood functions*. New York: Springer; 1996.
- Wang CP, Brown CH, Bandeen-Roche K. Residual diagnostics for growth mixture models: Examining the impact of a preventive intervention on multiple trajectories of aggressive behavior. *Journal of the American Statistical Association* 2005;100:1054–1076.
- Werthamer-Larsson L, Kellam SG, Wheeler L. Effect of first-grade classroom environment on shy behavior, aggressive behavior, and concentration problems. *American Journal of Community Psychology* 1991;19:585–602. [PubMed: 1755437]

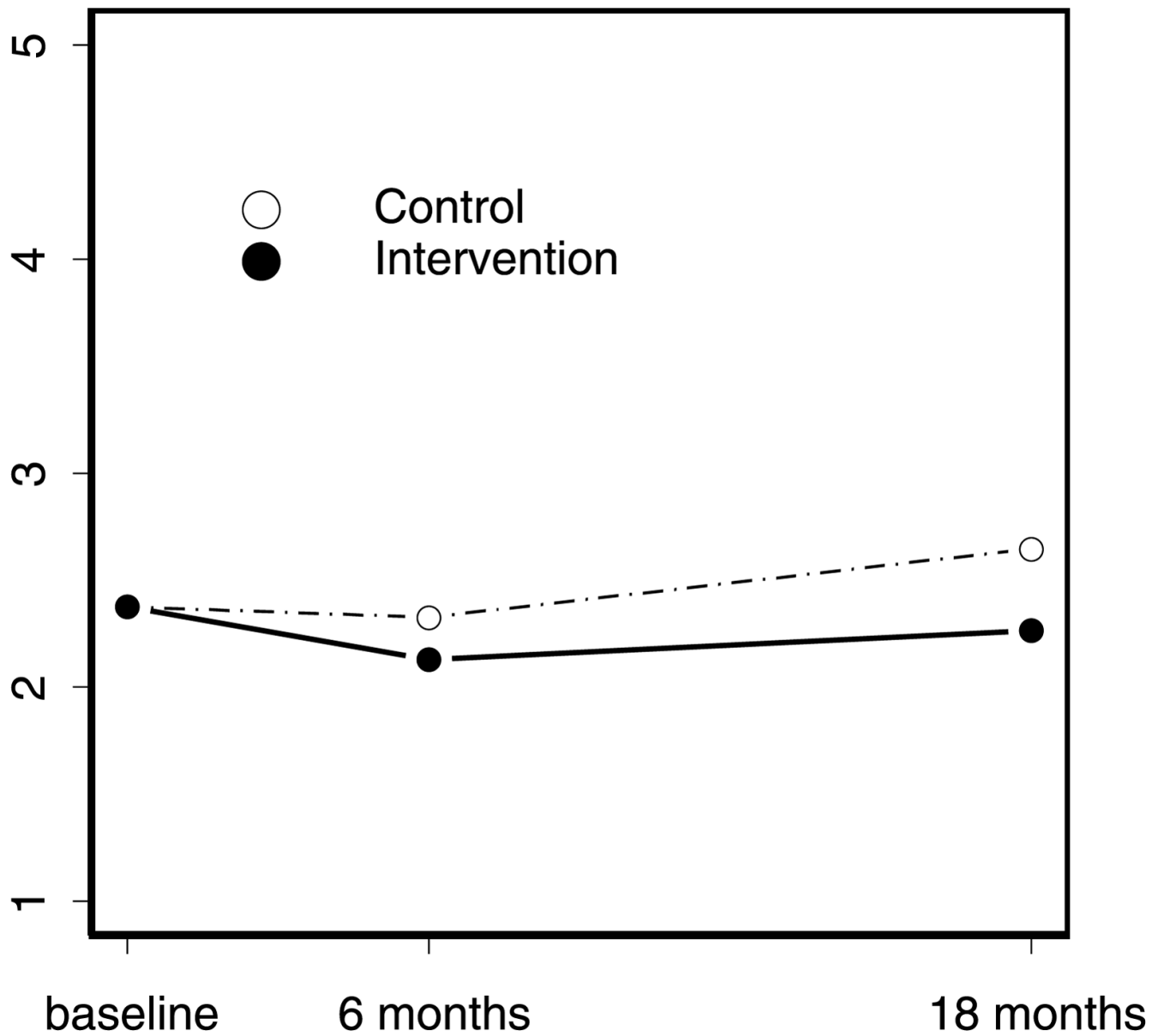


Figure 1.
Estimated overall average growth trajectories (entire sample).

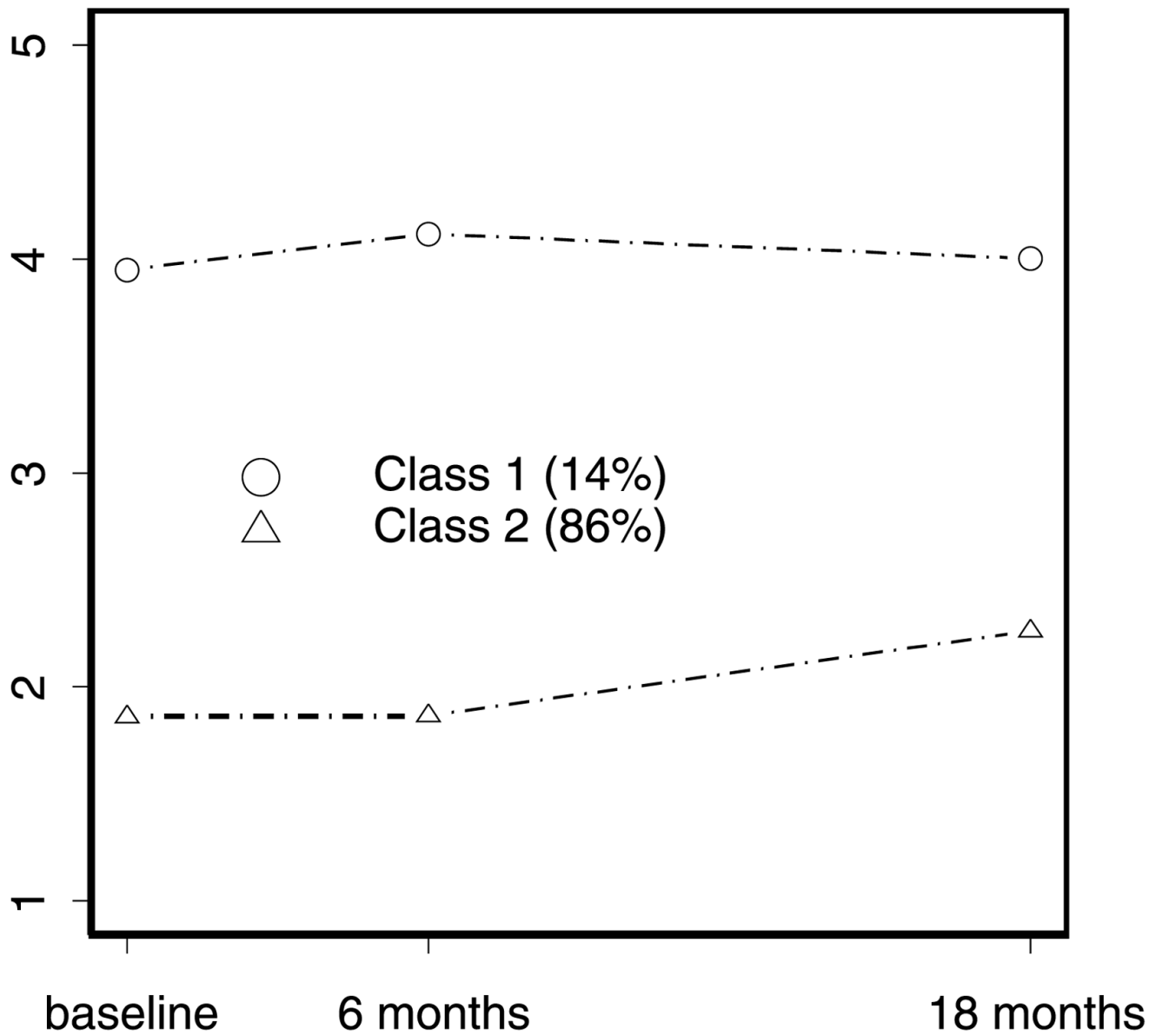


Figure 2.
Estimated mean outcome trajectories from Step 1 (control group only).

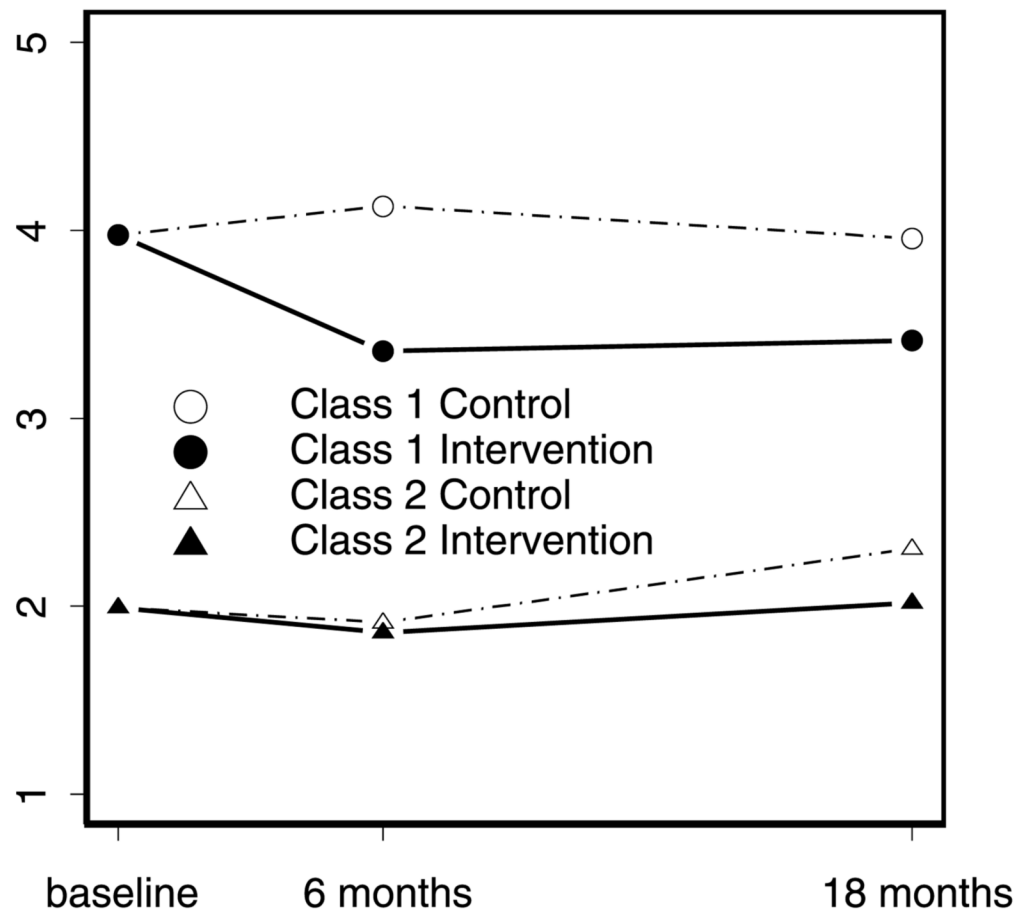


Figure 3.
Estimated mean outcome trajectories from Step 2 (entire sample).

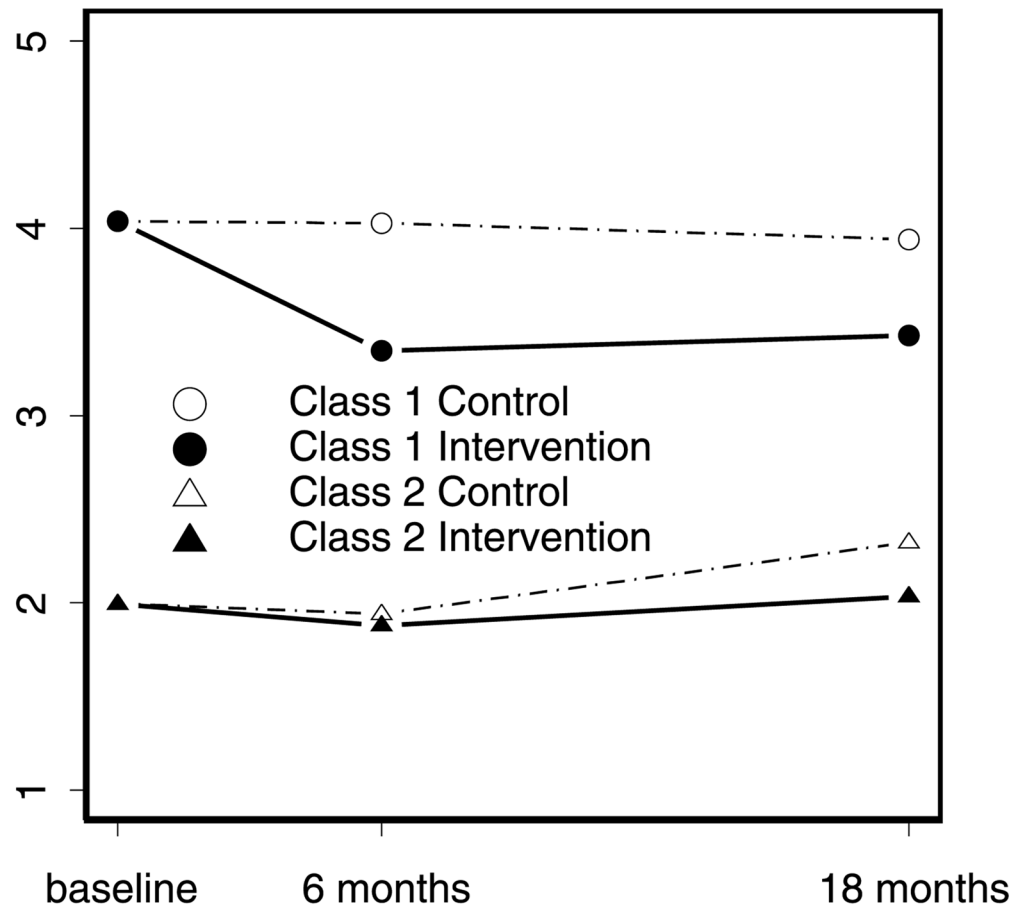


Figure 4.
Estimated mean outcome trajectories using the one-step approach (entire sample).

Table 1

Hopkins PIRC: Sample means (standard deviations in parentheses)

Variable	Description	Control	Intervention
Y_1	Attention deficit at baseline	2.16 (1.02)	2.58 (1.06)
Y_2	Attention deficit at 6 months	2.18 (1.01)	2.26 (0.92)
Y_3	Attention deficit at 18 months	2.50 (1.08)	2.40 (1.06)
X_1	Child's gender (male)	0.51	0.57
X_2	Employed	0.62	0.57
X_3	Married	0.45	0.36
X_4	African-American	0.84	0.84
X_5	Higher education	0.38	0.40

Table 2

Hopkins PIRC: ITT effect estimates from the mixed effects model approach (γ_L stands for the treatment effect on the linear slope and γ_Q on the quadratic slope)

Parameter	Estimate	SE
γ_L	-0.229	0.101
γ_Q	0.034	0.033

Table 3

Hopkins PIRC: Differential Causal Effect Estimates From the Two-Step Growth Mixture Approach (01 stands for the problematic and 02 stands for the normative trajectory stratum under the control condition)

Parameter	Estimate	SE
γ_{L01}	-1.065	0.371
γ_{Q01}	0.295	0.116
γ_{L02}	-0.032	0.099
γ_{Q02}	-0.022	0.035