Group-Based Trajectory Modeling in Clinical Research

Article *in* Annual Review of Clinical Psychology · March 2010

DOI: 10.1146/annurev.clinpsy.121208.131413 · Source: PubMed

2 authors, including:

Daniel S. Nagin
Carnegie Mellon University
172 PUBLICATIONS **24,754** CITATIONS

SEE PROFILE

Group-Based Trajectory Modeling in Clinical Research

Daniel S. Nagin[1] and Candice L. Odgers[2]

[1]Heinz School of Public Policy, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213-3890; [2]School of Social Ecology, University of California, Irvine, California 92697-7050; email: dn03@andrew.cmu.edu

Annu. Rev. Clin. Psychol. 2010. 6:109–38

First published online as a Review in Advance on January 4, 2010

The *Annual Review of Clinical Psychology* is online at clinpsy.annualreviews.org

This article's doi:
10.1146/annurev.clinpsy.121208.131413

1548-5943/10/0427-0109$20.00

Key Words

growth mixture modeling, longitudinal data, dual trajectory models, causal inference

Abstract

Group-based trajectory models are increasingly being applied in clinical research to map the developmental course of symptoms and assess heterogeneity in response to clinical interventions. In this review, we provide a nontechnical overview of group-based trajectory and growth mixture modeling alongside a sampling of how these models have been applied in clinical research. We discuss the challenges associated with the application of both types of group-based models and propose a set of preliminary guidelines for applied researchers to follow when reporting model results. Future directions in group-based modeling applications are discussed, including the use of trajectory models to facilitate causal inference when random assignment to treatment condition is not possible.

Contents

INTRODUCTION

A developmental trajectory describes the course of an outcome over age or time. Over the past decade there has an outpouring of studies of developmental trajectories in psychology, medicine, and criminology that apply a method alternatively called group-based trajectory modeling (GBTM) by Nagin (2005, 1999) or growth mixture modeling (GMM) by Muthén (2001). Analyses have studied outcomes as varied as physical aggression (Nagin & Tremblay 1999), cortisol levels (Van Ryzin et al. 2009), Internet usage (Christ et al. 2002), obesity (Mustillo et al. 2003), anxiety (Côté et al. 2010), and crime trends at the level of local neighborhoods (Weisburd et al. 2004).

The past decade has seen a rapid rise in the application of trajectory-based models in clinical research; a PSYC INFO literature review indicates that between the years 2000 and 2008, application increased from 8 to 80 publications per year in clinically relevant journals such as the *Journal of Clinical and Consulting Psychology*, *Child Development*, *Addiction*, and the *Journal of Abnormal Child Psychology*. Within this area, trajectory models have been applied to understand the etiology and developmental course of a number of different types of disorders, including depression (Dekker et al. 2007, Mora et al. 2009), inattention/hyperactivity (Jester et al. 2008), post-traumatic stress disorder (Orcutt et al. 2004), substance abuse (Hu et al. 2008), and conduct disorder (Odgers et al. 2008b). More recently, group-based models have been extended to capture heterogeneity in treatment responses to clinical and randomized trials (Brown et al. 2008, Peer & Spaulding 2007) and have been leveraged to facilitate causal inference in epidemiological observational studies where randomization to treatment conditions is not possible (Haviland et al. 2007, 2008; Odgers et al. 2008a). In many ways, it is not surprising that GBTMs have been embraced by clinical researchers, as they map closely on to how clinicians conceptualize the growth and development of a wide range of disorders

and associated symptoms; provide an empirical means of identifying clusters of individuals following both typical and atypical courses of development (that is, they do not assume a one-size-fits-all model for characterizing symptom onset and progression); and seem to offer a new set of tools for evaluating individual variation in response to clinical interventions and randomized trials.

While the introduction of GBTM into clinical research has provided new opportunities for discovery, the rapid and continuing evolution of group-based models also means that their underlying assumptions and extensions may not yet be familiar to those working in this area. The purpose of this article is to provide a nontechnical overview of the statistical basis of GBTM and GMM and a sampling of how these models have been applied to inform clinical research. For those interested in applying group-based models in their own work, we discuss several of the key challenges associated with the application of GBTM and GMM and propose some preliminary guidelines for reporting results from group-based trajectory models. The focus of this review is on the methods and their application, not on the several software alternatives available for estimating the models.

HOW CAN GROUP-BASED TRAJECTORY MODELS INFORM CLINICAL RESEARCH?

Although there are important technical differences between GBTM and GMM, which we discuss, conceptually they have much in common.[1] Both methods are designed to identify clusters of individuals, called trajectory groups, who have followed a similar developmental trajectory on an outcome of interest and, for the purposes of this article, both types of models have been applied to answer clinical research questions. There are several reasons why analyzing longitudinal data with a group-based approach may be attractive to clinically focused researchers. One of the most compelling reasons is that there is a long tradition of group-based theorizing about both normal and pathological development in clinical and developmental psychology. Examples include theories of personality development (Caspi 1998), drug use (Kandel 1975), learning (Holyoak & Spellman 1993), language and conceptual development (Markman 1989), depression (Kasen et al. 2001), eating disorders (Tyrka et al. 2000), alcoholism (Cloninger 1987), anxiety (Cloninger 1986), and the development of prosocial behaviors such as conscience (Kochanska 1997) and of antisocial behaviors such as conduct disorder and delinquency (Loeber 1991, Moffitt 1993, Patterson et al. 1989).

To test such taxonomical theories, researchers have commonly resorted to using assignment rules based on subjective categorization criteria to construct categories of developmental trajectories. Although such assignment rules are generally reasonable, there are limitations and pitfalls attendant to their use. One is that the existence of distinct developmental trajectories must be assumed a priori. Thus, the analysis cannot test for their presence, a fundamental shortcoming. Second, ex ante specified rules provide no basis for calibrating the precision of individual classifications to the various groups that comprise the taxonomy. Thus, the uncertainty about an individual's group membership cannot be quantified in the form of probabilities. For a fuller discussion of these issues, see chapter 1 of Nagin (2005).

The data presented in **Figure 1** provide an example of how group-based trajectory models have been applied to empirically test predictions stemming from a widely cited theory in developmental psychopathology. Close to 20 years ago, Moffitt (1993) outlined predictions based on her developmental taxonomy of antisocial behavior. Until recently, evaluations of this taxonomy have relied on the types of clinical algorithms and assignment rules described above. **Figure 1** provides an illustration

Group-based trajectory modeling (GBTM): finite mixture modeling application that uses trajectory groups as a statistical device for approximating unknown trajectories across population members

Growth mixture modeling (GMM): elaboration of Growth Curve Modeling based on finite mixture modeling to identify distinct yet unobservable subpopulations

Trajectory groups: clusters of individuals following similar trajectories on an outcome over time

[1] Throughout this review, we refer to both GBTM and GMM as group-based methods. When the discussion is specific to a method, the method under discussion is specifically referenced as GBTM or GMM.

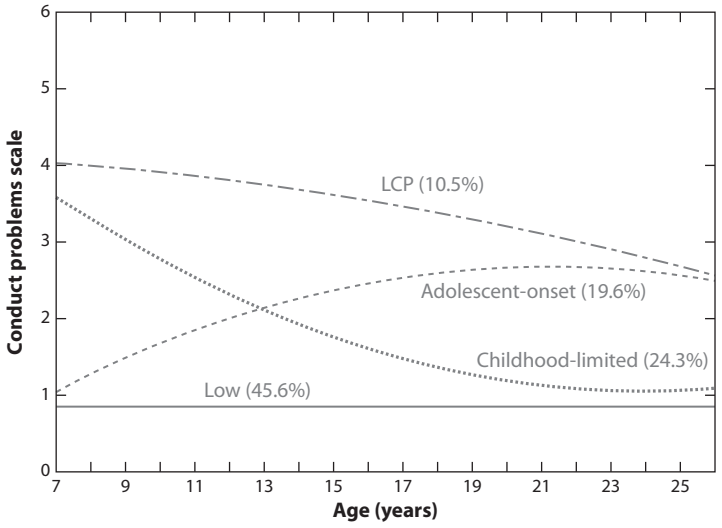

Figure 1

Trajectories of conduct disorder for males from age 7 to 32 in the Dunedin Multidisciplinary Health and Development Study. (Reprinted with permission from Archives of General Psychiatry, April 2007, vol 64, pp. 476–84. Copyright © 2007 American Medical Association. All rights reserved.) LCP, life-course persistent.

of a group- versus classification rule–based approach to testing predictions from this taxonomy. The data were collected as part of Dunedin Multidisciplinary Health and Development Study that tracked 1037 individuals from 3 to 32 years of age. The four trajectories reported in **Figure 1** were derived based on assessments of conduct disorder symptoms from age 7 to 26 (Odgers et al. 2007).[2] As predicted by Moffitt's original theory, approximately 10% of the male population was classified as following a life-course-persistent pathway of antisocial behavior. As shown in **Figure 1**, individuals following this pathway initiated their conduct problems in childhood and persisted into adulthood. The group-based trajectory solution also identified an adolescent-onset and childhood-limited pathway of conduct problems that was not originally anticipated by the taxonomic theory.

The conduct problem trajectories in **Figure 1** illustrate two valuable properties of a group- versus classification rule–based approach. One advantage of the GBTM approach is the capacity to identify qualitatively distinct developmental progressions that are not readily identifiable using ad hoc, ex ante classification rules. In principle, the childhood-limited and adolescent-onset groups shown in **Figure 1** are identifiable ex ante, but given the specific developmental course of each, it would be difficult to identify them without a formal statistical methodology. A second closely related advantage, which also stems from the use of a formal statistical structure, is that the methodology has the capacity for distinguishing chance variation across individuals from real differences. That is, there is an opportunity to calibrate whether individual change represents real versus only random variation in behavior. Because the childhood-limited and adolescent-onset trajectories are the products of a formal statistical model, there is a stronger basis for their reality than if they had been constructed based on subjective classification rules alone.

A group-based methodology is also responsive to calls for the development of person-based approaches to analyzing development (Bergman 1998, Magnusson 1998). Such appeals are motivated by a desire for methods that can provide a statistical snapshot of the key characteristics and behaviors of individuals following distinctive developmental pathways. In the example above, Odgers and colleagues (2008a) went on to distinguish the four groups on childhood antecedents that were hypothesized to characterize each developmental pathway. For example, individuals classified as following the life-course-persistent pathway were characterized by a more compromised constellation of neuro-developmental, familial, and social risk factors as compared to the other members of the population. These findings corresponded with Moffitt's predictions regarding the etiology and origins of individuals following the life-course-persistent pathway and provided an external validation check on the trajectory-group solution.

To further illustrate this point, consider the four trajectories of physical aggression shown

Person-based approaches: study of individuals on the basis of clusters of individual characteristics; often contrasted with the modeling of variables over individuals

[2]These trajectories pertain to the male subjects. A separate trajectory model was estimated for the females.

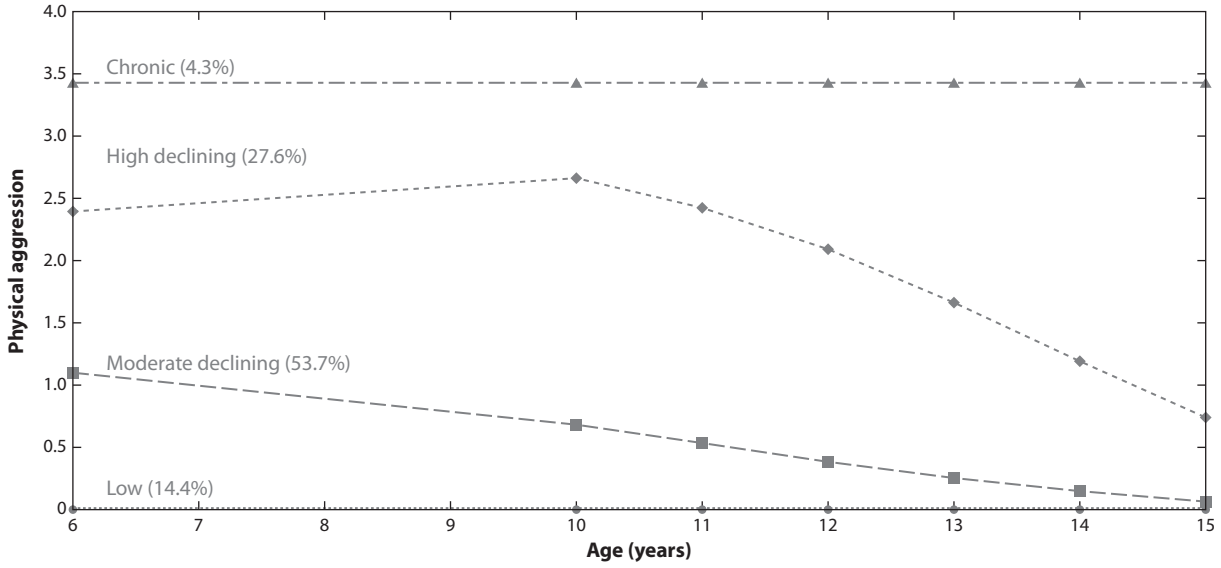

Figure 2

Trajectories of physical aggression from age 6 to 15 for males in the Montreal-based longitudinal study sample. (Data from Nagin & Tremblay 1999.)

in **Figure 2**. These trajectories, which were first reported in Nagin & Tremblay (1999), are based on teacher reports of physical aggression from age 6 to 15 collected as part of a large Montreal-based longitudinal study of more than 1000 males. **Table 1** reports profiles of the characteristics of individuals following the four physical aggression trajectories shown in **Figure 2**. The profiles conform to longstanding findings regarding the predictors and consequences of problem behaviors such as physical aggression. For example, individuals in the chronic-aggression group tended to have the least-educated parents and were most likely to score in the lowest quartile of the sample's IQ distribution. By contrast, individuals in the low-aggression group were least likely to suffer from these risk factors. Further, over 90% of the chronic-aggression group failed to reach the eighth grade on schedule, and 13% had acquired a juvenile record by age 18. By comparison, only 19% of the low-aggression group

Table 1 Physical aggression group profiles in the Montreal-based longitudinal study. (Data from Nagin & Tramblay 1999)

	Group			
Variable	Low	Moderate declining	High declining	Chronic
Years of school: mother	11.1	10.8	9.8	8.4
Years of school: father	11.5	10.7	9.8	9.1
Low IQ (% in lowest quartile)	21.6	26.8	44.5	46.4
Completed eighth grade on time (%)	80.3	64.6	31.8	6.5
Juvenile record (%)	0.0	2.0	6.0	13.3
# of sexual partners age 17[a]	1.2	1.7	2.2	3.5

[a]Number of sexual partners at age 17 within the past year.

had fallen behind grade level by the eighth grade, and none had a juvenile record. Results from the Dunedin and Montreal studies provide examples of how reams of longitudinal data can be summarized into a compact and easily interpretable form using GBTMs, allowing researchers to explore patterns in data and communicate findings in a transparent fashion.

Growth curve modeling (GCM): models that capture the average developmental trend and random variation around the average trend using one set of parameters for the population

A NONTECHNICAL OVERVIEW OF THREE APPROACHES TO TRAJECTORY MODELING

This section provides a nontechnical overview of three forms of trajectory modeling—growth curve modeling, GMM, and GBTM. Citations to far more complete technical developments are provided. Hierarchical modeling (Bryk & Raudenbush 1987, Goldstein 1995) and latent curve analysis (McArdle & Epstein 1987, Meredith & Tisak 1990, Muthén 1989, Willett & Sayer 1994) are two important alternative approaches to GMM and GBTM. For our purpose, we refer to them as growth curve modeling (GCM). GCM, GMM, and GBTM share a common analytical objective—measuring and explaining differences across population members in their developmental trajectories. In other words, analyzing population variability in developmental trajectories forms the statistical raison d'être for each of these methods. What distinguishes GCM, GMM, and GBTM is their approach to modeling individual-level heterogeneity in developmental trajectories, as each approach makes different technical assumptions about the distribution of trajectories in the population. Although not a technical requirement of GCM, GMM, or GBTM, most applications of these three approaches model trajectories with polynomial functions of age or time. For applications based on a polynomial trajectory specification, GCM, GMM, and GBTM are distinguished by their specification of how the parameters defining individual-level trajectories vary within the population.

Researchers interested in modeling individual-level heterogeneity in developmental trajectories are often faced with a choice of whether to apply GCM, GMM, or GBTM. As such, a brief overview of the assumptions underlying each model are provided below to help researchers determine which type of model may provide the best fit to their research question.

Growth Curve Modeling: Does One Type of Curve Fit All?

GCMs allow researchers to map interindividual differences in change over time and provide a means of aggregating repeated measures into relatively few parameters, such as estimates of the average rate of growth and variability in development over time (Bollen & Curran 2006). In doing so, GCMs capture mean trends in development (nomethic aspects of change) as well as individual departures from the average trend (ideographic aspects of change) (Preacher et al. 2008). Although the assumptions underlying hierarchical modeling and latent curve analysis differ in important respects, they also have important commonalities (MacCallum et al. 1997, Raudenbush 2001, Willett & Sayer 1994). For the purposes of this review, one commonality is crucial: Both model the population distribution of individual trajectories based on a continuous distribution function and assume that the random effects are continuously distributed throughout the population, usually according to a multivariate normal distribution. More practically speaking, both types of models assume that individuals are drawn from the same population and that development over time can be mapped using one set of parameters.

The first step in growth curve modeling typically involves fitting an unconditional model that estimates the shape of development over time using two key features of the population distribution of trajectory parameters—their mean and covariance structure. The former defines average growth and the latter calibrates the variance of growth throughout the population. In applications based on polynomial specifications of growth, the parameter estimates of the polynomial describe the average growth in the population, whereas individual-level

variation in development is captured by random effects around the polynomial parameters. The variances and covariances of the random effects describe individual-level variability. Next, conditional models are often fit to test whether variability in trajectory parameters can be predicted by one or more explanatory variables. For example, a conditional GCM has been applied to test whether child maltreatment predicts individual differences in children's initial levels and/or rate of change in depressive symptoms across early adolescence (Kim & Cicchetti 2006).

To summarize, if it is assumed that all individuals in the population follow a similar functional form of development, then GCMs may be sufficient to capture interindividual variability in change across time. However, if one trajectory shape is not assumed to "fit all," then the group-based trajectory models described below will likely provide a better fit to the research question and have the secondary benefit of better fitting the data.

Growth Mixture Modeling

Growth mixture modeling is an innovation of Muthén & Shedden (1999). It is an elaboration of GCM based on a class of statistical models called finite mixture models (McLachlan & Peel 2004, Titterington et al. 1985). One important use of finite mixture models is to analyze data in which the general population is thought to be composed of literally distinct subpopulations that are not identifiable based on measured characteristics ex-ante. An example is a disease study in which two groups comprise the population, one with a genetic vulnerability to a disease and another without the vulnerability. If the two groups were distinguishable based on measured characteristics, they could be analyzed separately. However, if they are not distinguishable, the data will be composed of a mixture of the two groups.

In short, finite mixture models are a class of statistical models designed to analyze data composed of a mixture of two or more groups whose outcomes are generated by distinct statistical processes. For the disease example, one component of the mixture would effectively be committed to modeling the outcome (e.g., symptoms over time) for the genetically vulnerable group, and the second component of the mixture would be committed to modeling the same outcome for the group that is not genetically vulnerable. In general, finite mixture models may have more than two components, but as the adjective "finite" implies, the number of components cannot be infinite.

Finite mixture models: powerful models for analyzing outcomes from a population that contains a finite number of homogenous subpopulations or for approximating unknown distributions

In the context of the analysis of developmental trajectories, such subpopulations may constitute groups with different unconditional (mean) trajectories that cannot be explained by the interindividual variability provided by the random effects. The innovation of Muthén & Shedden (1999) was to apply finite mixture modeling to GCM so that two or more GCMs are used to model population variability in developmental trajectories. The basic outputs of the model are two or more GCMs, each of which is interpretable in the same way as a single group GCM, and estimates of the proportion of the population following each such GCM.[3] Muthén's writings on GMM imply that he considers each GCM as modeling a separate subpopulation following a different growth curve.

Group-Based Trajectory Modeling

Group-based trajectory modeling is also an application of finite mixture, but the motivation for applying this method is fundamentally different from that in GMM. Finite mixture-based trajectory models assume that the population is composed of a mixture of distinct groups defined by their developmental trajectories. For instance, for the model depicted in **Figure 1**, the specification assumes that the population is composed of four distinct conduct problem trajectories. The assumption that the population is composed of distinct groups is not likely literally correct. Unlike biological or physical

[3]The number of GCMs estimated is specified by the analyst rather than estimated directly from the data. Similarly, the number of groups in a GBTM is also specified, not estimated.

phenomena, in which populations may be composed of discrete groups such as different types of animal or plant species, population differences in developmental trajectories of behavior are unlikely to reflect such bright-line differences (although biology has a long tradition of debates concerning classification; see Appel 1987).

To be sure, many taxonomic theories predict different trajectories of development across subpopulations (e.g., Belsky et al. 1991, Kandel 1975, Loeber 1991, Moffitt 1993, Patterson et al. 1989). However, the purpose of such taxonomies is generally to draw attention to differences in the causes and consequences of different developmental trajectories rather than to suggest that the population is composed of literally distinct groups. As already discussed, one purpose of the group-based modeling strategy is to provide a methodological complement to theories that predict differing developmental etiologies and trajectories within the population. Within this framework, the basic elements of such theories can be empirically examined: That is, we can ask, are the developmental trajectories and etiologies predicted by theory actually present in the population?

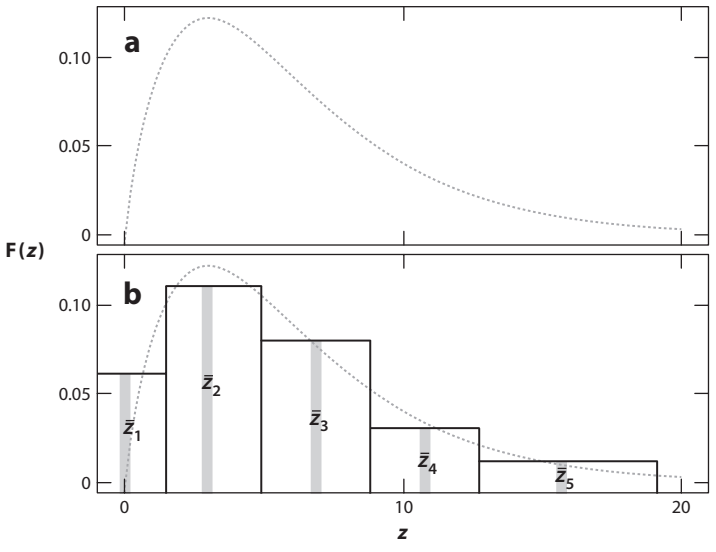

Figure 3

Using groups to approximate an unknown distribution.

It is important to point out, however, that directly testing many of these theories requires making assumptions about the population distribution of developmental trajectories. Yet, developmental theory rarely provides explicit guidance on how unobserved individual differences in development are distributed in the population. This brings us to a fundamental distinction between GBTM and GMM. GMM assumes that the population distribution of trajectories is composed of two or more subpopulations, each following a conventional GCM. In contrast, GBTM takes no stand on the population distribution of trajectories and instead uses the trajectory groups as a statistical device for approximating the unknown distribution of trajectories across population members.

This use of finite mixture models employed in GBTM aligns with the work of Heckman & Singer (1984), who built upon the approximating capability of finite mixture models to construct a nonparametric maximum likelihood estimator for the distribution of unobserved individual differences in duration models. The motivation for this seminal innovation was their observation that social science theory rarely provided theoretical guidance on the population distribution of unobserved individual differences, yet statistical models of duration data were often sensitive to the assumed distributional form of such differences. Their proposed estimator finessed the problem of having to specify a distribution of unobserved individual differences by approximating the distribution with a finite mixture model.

The idea of using a finite number of groups to approximate a continuous distribution is easily illustrated with an example. Suppose that panel *a* in **Figure 3** depicts the population distribution of some behavior *z*. In panel *b*, this same distribution is replicated and overlaid with a histogram that approximates its shape. Panel *b* illustrates that any continuous distribution with finite end points can be approximated by a discrete distribution (i.e., a histogram) or alternatively by a finite number of "points of support" (i.e., the dark shaded "pillars"). A higher number of support points yields

a discrete distribution that more closely approximates the true continuous distribution.

Thus, in GBTM, trajectory groups should be thought of a statistical device for approximating what is in all likelihood a continuous population distribution of trajectories of unknown shape. The trajectory groups are a convenient statistical device for summarizing trajectories in distinctive regions of the distribution. Like contour lines on a topographic map, the groups are not literal entities, but rather approximations of distinctive regions of the surface. Also, just as a topographic map may trace out either a regularly or irregularly varying geographic surface, trajectory groups may suggest a population distribution of trajectories that varies regularly (e.g., all generally rising trajectories) or irregularly (e.g., some rising and others falling). Both are useful pieces of information to establish about the contours of the distribution. From a technical perspective, the difference between trajectory groups in the GMM and GBTM methods is that the former includes random effects in each group's trajectory model, and the latter does not. This difference in the conception of groups can lead to important differences in the way an analysis proceeds. That is, the addition of random effects to a group-based model can result in the use of fewer trajectory groups because their addition allows for more within-group variability in individual-level trajectories. If the groups are thought of as subpopulations following literally different GCMs, the reduction in number of groups results in a more parsimonious model. By contrast, in GBTM where groups are used as a device for approximation, increasing the within-group variability of individual-level trajectories is at odds with the objective of reducing within-group variability in development. Stated differently, in GBTM, a group is conceptually thought of as a collection of individuals who follow approximately the same developmental trajectory, the conceptual equivalent of a contour line on a topographic map. Trajectory groups can also be thought of as latent longitudinal strata where population variability is captured by differences across groups in the shape and level of their trajectories (Haviland et al. 2007, 2008).

The Number of Groups and Extraction of "Fictitious" Groups

One of the key decision points in group-based modeling is a determination of the number of groups or latent classes that best represents the heterogeneity in developmental trajectories. This issue has been discussed extensively in prior work (McLachlan & Peel 2004, Muthén 2004, Nagin 2005, Nylund et al. 2007). As such, only a brief listing of commonly used fit indices used to make this determination is provided here. The most commonly used criteria to evaluate model fit include the Bayesian information criteria (BIC; Raftery 1995, Schwartz 1978), Akaike information criterion (AIC; Akaike 1974), Lo-Mendell-Rubin likelihood ratio test (LMR-LRT; Lo et al. 2001), and entropy. Because comparisons between models with k versus $k + 1$ classes cannot be made via a standard likelihood ratio comparison, indices such as the BIC and the AIC are commonly employed to assess model fit by balancing model complexity (number of parameters) versus goodness of fit to the sample data. More recently, alternative indices such as the LMR-LRT have been applied to evaluate competing models in the GMM framework. The LMR-LRT provides a likelihood-ratio-based method for determining the ideal number of classes; a low p-value indicates that a *k-1* class model should be rejected in favor of a model with at least k classes. Entropy is also used in model selection, as it indexes classification accuracy by averaging the posterior probabilities after individuals have been assigned to their most likely class, with values closer to 1 indexing greater precision (range 0 to 1).

However, it is important to note that model selection based on the mechanical and rigid application of a formal statistical criterion may lead to an inferior choice. The strengths and weaknesses of alternative model specifications depend upon the substantive questions being

BIC: Bayesian information criteria

AIC: Akaike information criterion

LMR-LRT: Lo-Mendell-Rubin likelihood ratio test

asked and the data available for addressing these questions. Thus, the choice of the best model specification cannot be reduced to the application of a single test statistic. To be sure, the application of formal statistical criteria and objective standards to the model selection process serves to discipline and constrain subjective judgment. However, there is no escaping the need for judgment; otherwise, insight and discovery will fall victim to the mechanical application of the method. In the end, the objective of the model selection is not the maximization of some statistic of model fit; rather, it is to summarize the distinctive features of the data in the most parsimonious—and useful—fashion possible.

An often-neglected step in the model selection process is testing the adequacy of the selected model. The most basic test of adequacy is whether the final model adequately addresses the research question under investigation. Beyond this important substantive test, Nagin (2005) lays out several statistically oriented criteria for assessing model adequacy. These include: (*a*) obtaining for each trajectory group a close correspondence between the estimated probability of group membership and the proportion assigned to that group based on the posterior probability of group membership, (*b*) ensuring that the average of the posterior probabilities of group membership for individuals assigned to each group exceeds a minimum threshold of 0.7, (*c*) establishing that the odds of correct classification based on the posterior probabilities of group membership exceed a minimum threshold of 5, and (*d*) observing reasonably tight confidence intervals around estimated group membership probabilities.

Regardless of the criteria used for selection, the choice of the number of groups—and the potential of reifying groups—has been a topic of much discussion and debate in the literature. For example, Bauer & Curran (2003) have demonstrated in simulation analysis that seemingly modest specification error may result in the overextraction of groups in GMM analyses. Their work serves as a useful a caution against the quixotic quest to identify the true number of groups in either GMM or GBTM analyses. Perhaps most importantly, this work reinforces the need to move away from interpretations of trajectory groups as literally distinct entities. As emphasized in the prior section, it is our view that trajectory groups are just approximations of a more complex reality. As William Baumol (1992) observes, "A well-designed model is, after all, a judiciously chosen set of lies, or perhaps more accurately put, partial truths about reality, which have been chosen so as to permit us to reason more effectively about some issue than we otherwise could. The model must be an oversimplification if it is to be tractable analytically. Optimality model constructions must be based on the trade-off between these two desiderata—accuracy of representation of reality and usability in analysis" (p. 55).

Once trajectory groups are understood to be clusters of individuals following similar trajectories, not literally distinct entities, then the issue of the over- (or under-) extraction of groups becomes a red herring—the similarity of the clusters is a reality of the data, not a fiction. The interesting questions become: Are the trajectory groups distinguishable in terms of pre-existing characteristics, subsequent outcomes, their response to treatment, or their relationship to trajectories for other outcomes or behaviors? If they differ on none of these dimensions, then the clusters no longer serve a useful substantive purpose. If, however, differences across the groups are identified, then the clustering has served a useful purpose, and their continued study may have merit by whatever statistical method one may choose. Continued study should, of course, involve replication across datasets to establish that the trajectory groups represent meaningful strata.

Although the strongest applications of group-based modeling will be guided by a priori expectations regarding the number, shape, and size of trajectory groups (Bauer 2007), in the majority of cases, pre-existing theories regarding unobserved individual differences in trajectory distributions do not exist. Thus, applied researchers are left analyzing their data somewhere in between an exploratory and

confirmatory framework, making it crucial to clearly communicate the decision points and justifications employed to select the best trajectory model. With these cautionary notes in mind, the following section touches on the types of clinical research data and questions for which GBTMs may be appropriate.

PRACTICAL ISSUES AND CHALLENGES IN GROUP-BASED TRAJECTORY MODELING

Difficult Distributions

Clinical data often present challenges at the analysis stage owing to the fact that symptoms are rarely normally distributed, and many outcomes are dichotomous (e.g., diagnoses). Fortunately, available software programs allow for the estimation of a wide range of distributions, including censored normal/tobit (e.g., where symptom counts may be clustered at the bottom, top, or both ends of the scale), zero-inflated Poisson (e.g., where a large number of individuals have no symptoms and the remaining individuals experience one or more symptoms), and binary (e.g., classifications of whether or not an individual has met diagnostic criteria for a given disorder). The flexibility of current software programs for accommodating this wide range of distributions is good news for clinical researchers who are trying to make sense of reams of often skewed and categorical data that have been gathered over time. For excellent examples of group-based applications with categorical and skewed diagnostic data, see work by Feldman and colleagues (2009).

We should also note that when group-based models are fit within a structural equation modeling framework, it is possible to simultaneously estimate measurement models that calibrate individual scale items using factor analytic and/or item response theory techniques. This type of extension allows clinical researchers to model the reliable portion of the variance from multi-item symptom scales, link differing item sets across assessment phases, and test for metric invariance of scales over time (McArdle et al. 2009). Again, this type of flexibility is valuable for those working with clinical data, which are typically derived from multi-item scales and often require that item sets be modified over time to be developmentally appropriate. By testing for metric invariance, investigators are able to offer empirical support for the assumption that item sets are tapping into the same latent construct across distinct developmental stages (Meredith 1993, Nesselroade 1995) and perhaps most importantly, ensure that the construct that we are mapping has been calibrated on the same scale over time.

Metric invariance: components of the measurement model relating observed scores to latent variables that are numerically equal in different samples or across time

MAR: missing at random

FIML: full-information maximum likelihood

Missing Data

Clinical researchers are also typically faced with the challenge of analyzing incomplete data as participants drop out of clinical trials and are lost within longitudinal studies. Selective attrition is a common problem across clinical and epidemiological studies and, as in all analyses, needs to be addressed to produce unbiased and efficient estimates. Fuller discussions of missing data issues and strategies can be found elsewhere (Rubin 1976, Schafer & Graham 2002). However, a few key issues related to how missing data are handled in GBTMs are outlined below. Perhaps most importantly, it is imperative for researchers interested in applying GBTMs to document and understand both the type and the amount of missing data across occasions. In situations where data are missing at random (MAR), maximum likelihood estimations will provide parameter estimates that are asymtotically unbiased. When data are MAR, information from the dataset can, in addition, be used to impute missing data prior to input into the trajectory model. Alternatively, as is the case in GMM, full-information maximum likelihood (FIML) estimation can be used to integrate all available information based on MAR assumptions. Although the automated FIML option is attractive, the efficiency gain of any missing data technique will be constrained by the amount and the type of missing data within a particular dataset.

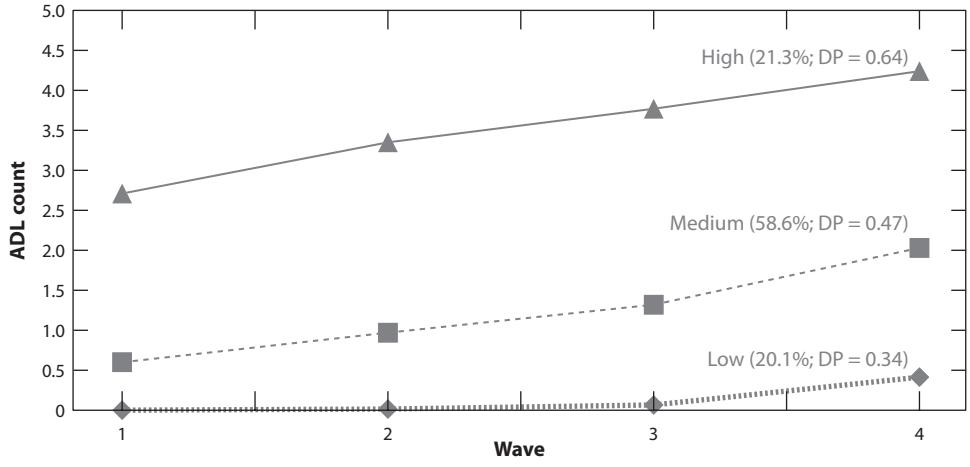

Figure 4

Activities of daily living (ADL) trajectory model with dropout. Note: Disability is measured by a count of the number of basic activities (e.g., bathing oneself), called activities of daily living, that the subjects of the study could not perform without assistance. The baseline data were collected in 1998, and the final wave of data collection was completed in 2005. DP, dropout probability.

Situations where data are missing systematically present the largest challenges to any analysis, and GBTMs are no exception. In trajectory analysis, attrition is of particular concern if rates of attrition are correlated with the outcome of interest, that is, missingness is related to the very thing that is being measured and is therefore considered "nonignorable" (e.g., in **Figure 1**, systematic missingness across occasions would mean that subject attrition is dependent on conduct problem levels). Although there are currently attempts to address nonignorable missing data via statistical techniques such as pattern mixture analyses (Little 1993), the nature of the missingness restricts the ability to test whether proper corrections have been made, and these techniques are not widely implemented. One option that trajectory modeling does allow, however, is the use of GBTM to better understand the predictors and the nature of nonrandom subject attrition. For example, **Figure 4** displays a three-group disability trajectory solution from the Chinese Longitudinal Healthy Longevity Survey (CLHLS) (Nagin et al. 2009). The CLHLS is a four-wave survey conducted in randomly selected counties and cities in 22 Chinese provinces and representing over 85% of the total population of China. As one would expect, mortality rates were closely related to disability trajectory. For the high group, the estimated mortality rate was 64% per assessment occasion; by 2005, only three individuals from this group were still alive. By contrast, the dropout rates due to death for the low- and medium-trajectory groups, although high, are considerably smaller, 34% and 47%, respectively (Nagin et al. 2009). This extension provides the capability to analyze how predictor variables are associated not only with dropout probability but also with trajectory group membership. Extending this example to the clinical context, it is easy to think of how, for example, attrition from a treatment program targeting depression could be predicted based on a study member's pre- or post-treatment symptom trajectory.

To summarize, as in all longitudinal data analyses, missing data challenge researchers to think more carefully about the reasons for subject attrition across development and to devise effective strategies for addressing selective attrition. Practically, researchers applying group-based models are left with the choice of

addressing missing data issues outside of the group-based modeling framework or, in the case of GMM, assuming that the data are MAR and applying the FIML correction. As illustrated in the above example, GBTM may also be helpful in understanding the reasons for subject dropout and, ideally, can serve as a tool for developing strategies to address missing data issues in subsequent analyses.

Defining Time in Trajectory-Based Models

In addition to devising strategies to handle distributional and missing data issues, researchers much also choose how to index time in GBTMs. As illustrated by the previous examples, the most common metric for indexing time in trajectory-based models is age, with the majority of applications involving the analysis of a single age cohort (see, for example, Nagin et al. 1995, Nagin & Land 1993).[4] More recently, group-based models have been extended to evaluate treatment efficacy. In these cases, the metric for counting time is typically time since treatment. For example, investigators have applied group-based models to analyze trajectories of depression following completion of initial treatment (Jones 2001), the course of impairment following outpatient psychotherapy (Stulz et al. 2007), and heterogeneity in the pharmacodynamics of methylphenidate response by children with attention deficit hyperactivity disorder (Sonuga-Barke et al. 2008). Still another possible metric is time since disorder onset or since the occurrence of a life event, with group-based models capturing heterogeneity in behaviors and symptoms following drinking relapse (Witkiewitz & Masyn 2008), remission (Xie et al. 2006), and psychiatric crisis (Halliday-Boykins et al. 2004).

[4]There is no technical obstacle to estimating trajectory models on multiple cohorts. However, the usual concerns regarding cohort effects must be considered. See section 7.6 of Nagin (2005) for a discussion of testing for cohort effects in group-based models.

SAMPLE APPLICATIONS AND EXTENSIONS TO GROUP-BASED TRAJECTORY MODELING IN DEVELOPMENTAL PSYCHOPATHOLOGY

This section reports a sampling of extensions to the basic model. The aim is to provide an overview of the types of analyses that can be conducted with a group-based approach to modeling trajectories. Due to space limitations, discussion of the technical details must be omitted. The majority of the examples provided below focus on capturing heterogeneity in the developmental course of externalizing behaviors across time. However, the logic and procedures underlying these analyses can easily be extended to capture the development of a wide range of mental health symptoms. Although all of the examples are based on GBTM, they could just as well be conducted in the GMM framework.

Adding Predictors of Trajectory Group Membership

The two most important outputs of the basic model are estimates of the shape of each group's trajectory and the size of the group as measured by the proportion of the population following that trajectory. Alternatively, this proportion can be interpreted as the probability of trajectory group membership. This section describes an extension of the basic model that allows the probability of trajectory group membership to depend on psycho-social and other characteristics of study subjects. It is a form of analysis that is designed to identify risk and protective factors associated with membership in a trajectory group. Because such characteristics are used as predictors of trajectory group membership, as a conceptual matter, these factors should be established at baseline or before.

This model extension is demonstrated with an analysis of a classic dataset assembled by Farrington & West (1990), which includes data on convictions from age 10 to 32 in a sample of more than 400 males from a poor neighborhood in London, England. As shown in **Figure 5**, a four-group model, analyzed using the

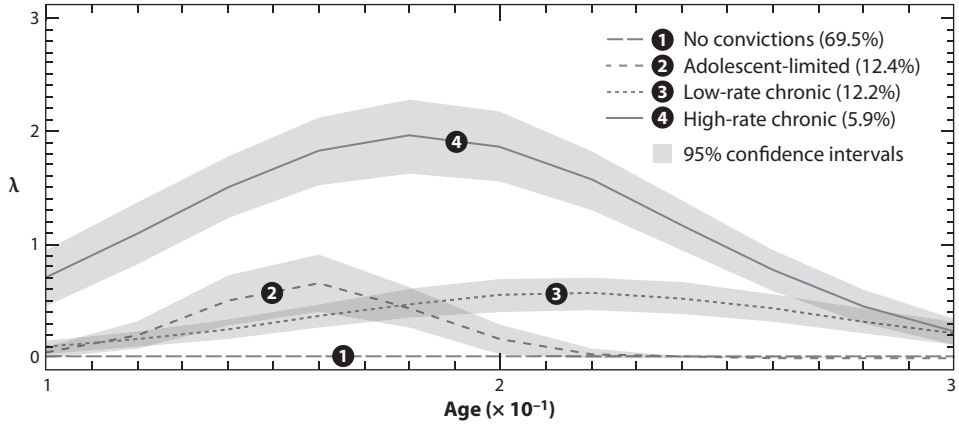

Figure 5

Trajectories of convictions in the Cambridge Study in Delinquent Development (Farrington & West 1990).

zero-inflated Poisson modeling options, was found to best fit the data based on criteria discussed above in The Number of Groups and Extraction of "Fictitious" Groups section. The largest trajectory group accounted for 69.5% of the population and was composed of individuals who generally had no convictions. The three offending trajectories included an adolescent-limited group (12.4% of the population), which peaked sharply in late adolescence and then declined to a near zero rate of offending by age 20, a high-chronic trajectory (5.9% of the population) with a high hump-shaped trajectory, and a low-rate-chronic trajectory that accounted for the remaining 12.2% of the population. Also shown in **Figure 5** are 95% confidence intervals around each trajectory. The fact that the confidence intervals do not overlap indicates that the solution is capturing distinctive features of the population's distribution of trajectories.

Next consider analyses examining how the probability of trajectory group membership varies with four classic psychosocial risk factors for antisocial behavior (low IQ, high risk-taking behavior, poor child-rearing behavior, and parental criminality). The association of these predictor variables to trajectory group membership is examined by specifying the probability of trajectory group membership to follow a multinomial logit model. **Table 2** reports the results of the analysis. The coefficient estimates correspond to the parameters of a multinomial logit function. The coefficients were estimated jointly with the parameter estimates of the trajectories themselves. The

Table 2 Predictors of trajectory group membership in the four-group London model[a]

Variable	Coefficient estimate	Z-score
Adolescent limited		
Intercept	−2.99	−6.31
Low IQ	0.74	1.50
Criminal parents	1.01	2.19
High risk taking	1.44	3.29
Poor child rearing	0.79	1.68
Low chronic		
Intercept	−2.79	−7.48
Low IQ	1.15	2.86
Criminal parents	1.38	3.46
High risk taking	0.72	1.71
Poor child rearing	0.56	1.28
High chronic		
Intercept	−4.97	−8.22
Low IQ	1.33	2.50
Criminal parents	2.08	3.98
High risk taking	2.21	4.03
Poor child rearing	1.16	2.22

[a]No conviction trajectory is the comparison group.
Data from Cambridge Study in Delinquent Development (Farrington & West 1990).

trajectory coefficients are not reported. For the analyses reported here, the comparison group is the no convictions trajectory. However, in general, any group can serve this purpose. For every trajectory group, all coefficient estimates are positive. This implies that each of these psychosocial characteristics is a risk factor for following a trajectory of heightened delinquency. However, not all are statistically significant. For the low-chronic group, the poor child-rearing risk factor has a z-score of only 1.28, which falls short of statistical significance at conventional levels. Still, the results provide strong support for the hypothesis that the majority of these individual-level characteristics are significant risk factors for membership in a trajectory of heightened criminal involvement.

The multinomial logit estimates can also be used to predict the probability of trajectory group membership for different configurations of risk factors. **Table 3** illustrates this capability for six scenarios about the level of these four predictor variables. Scenario 1 assumes that all predictors are equal to 0. This is equivalent to calculating group membership probability for individuals with none of the above risk factors for delinquency. Scenarios 2–5 report these same probabilities for individuals with only one of the four risk factors included in the model. In scenario 6, the group membership probabilities are computed for individuals with all four of the delinquency risks. The calculations illustrate the concept of cumulative risk, whereby an accumulation of risk—versus the presence of a single risk factor—is the most relevant index of an individual's vulnerability to psychopathology (Rutter et al. 1975). That is, the calculations show that each risk factor increases the probability of membership in one or more of the delinquent trajectory groups, but no single factor dramatically shifts the probabilities away from those in the no-risk scenario. For example, consider scenario 3. The model predicts that the probability of membership in the no conviction group is 0.70 for individuals who have at least one parent with a criminal record but who have none of the other risk factors. The counterpart prediction for the high-chronic group for these individuals is 0.039. In contrast, the predicted probabilities of membership in the no conviction and high-chronic groups for individuals with no risks, including criminal parents, are, respectively, 0.89 and 0.006. Thus, parental criminality materially reduces the no

Table 3 Predicted group membership probabilities and 90% confidence intervals for four-group London model

	Group membership probability			
Scenario	No conviction	Adolescent limited	Low chronic	High chronic
No risks	0.89 (0.83, 0.96)	0.05 (0.02, 0.09)	0.06 (0.03, 0.10)	0.01 (0.00, 0.02)
Low IQ only	0.75 (0.62, 0.83)	0.08 (0.04, 0.17)	0.15 (0.08, 0.25)	0.02 (0.01, 0.05)
Criminal parents only	0.70 (0.56, 0.78)	0.10 (0.04, 0.20)	0.17 (0.07, 0.28)	0.04 (0.02, 0.09)
High risk taking only	0.71 (0.58, 0.80)	0.15 (0.08, 0.27)	0.09 (0.05, 0.17)	0.05 (0.02, 0.09)
Poor child rearing only	0.80 (0.69, 0.87)	0.09 (0.04, 0.17)	0.09 (0.04, 0.16)	0.02 (0.01, 0.05)
All four risks	0.08 (0.03, 0.16)	0.21 (0.08, 0.41)	0.22 (0.09, 0.43)	0.48 (0.25, 0.69)
Population base rate	0.70	0.12	0.12	0.06

[a]Predictions based on 90% confidence intervals.
Data from Cambridge Study in Delinquent Development (Farrington & West 1990).

conviction group probability and increases the high-chronic group probability. Still, the basic ordering of the probabilities remains—the rare group is much more likely than the high-chronic group. However, the presence of all four risks (scenario 6) does result in a dramatic shift. The probability of membership in the high-chronic group increases from nearly 0 in the no-risk scenario to 0.48 in the all-four-risks scenario.

Adding Time-Varying Covariates and Estimating Treatment Effects in Group-Based Trajectory Models

The prior section demonstrated an extension of the basic model to allow the probability of trajectory group membership to vary as a function of individual-level characteristics. In this section, the focus shifts to a model generalization designed to analyze whether events that occur during the course of a trajectory alter the trajectory itself. The aim is to provide trajectory group–specific estimates of whether major life transitions such as the birth of a child or clinical interventions such as counseling or drug treatment alter the developmental course of the outcome under study. In a very useful set of analyses, Muthén et al. (2002) have used a similar form of this model extension to analyze whether a randomly assigned intervention has differential effects across trajectory groups. The model generalization can also be used to test for cohort effects in multiple cohort designs and to analyze whether variables that vary over a continuum, such as treatment exposure time, are also associated with changes in the trajectories.

Nagin et al. (2003) first demonstrated the generalized model in the context of an analysis from the Montreal data of whether grade retention or family breakup altered a child's trajectory of violence across adolescence. The results reported here represent a simplified version of that analysis. **Figure 6** shows the resulting trajectories from age 11 to 17 for a five-group model with no covariates other than age in the specification of the trajectories and no predictors of trajectory group membership. Specifically, the trajectories were estimated according

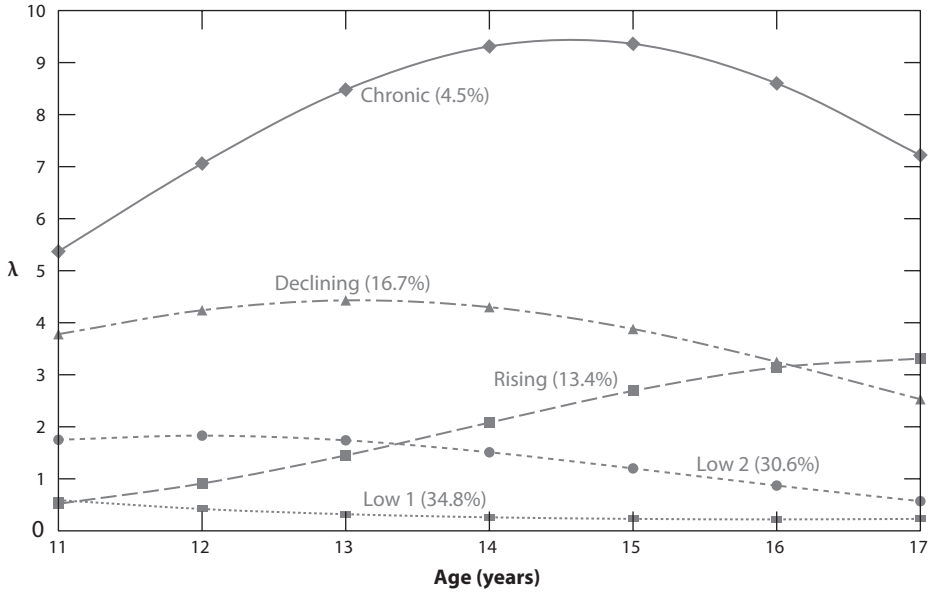

Figure 6

Trajectories of violent delinquency in the Montreal-based longitudinal study. (Data from Nagin & Tremblay 1999.)

to the following Poisson-based trajectory model for analyzing count data:

$$\ln(\lambda_t^j) = \beta_0^j + \beta_1^j Age_t + \beta_2^j Age_t^2$$

where λ_t^j is the Poisson rate parameter with j indexing trajectory group and t indexing time.

Two trajectories that comprised an estimated 65.4% population remained low and also declined throughout adolescence. They are labeled "low 1" and "low 2." The "rising" group, estimated to account for 13.4% of the population, began adolescence with similarly low violent delinquency but subsequently increased steeply. Two groups began with a high rate of violence but subsequently followed very divergent paths, one group declined (and hence is labeled "declining"), whereas the other group (labeled "chronic") remained high through adolescence.

We next consider whether the delinquency trajectories from age 11 to 17 seem to be altered by the experience of first-time grade retention and/or the separation from the subject's biological parents. There is much evidence that both of these stressors are associated with heightened delinquency (Foster et al. 2010, Maguin & Loeber 1996, Nagin et al. 2003, Pagani et al. 2001). To test for such associations, the specification of the group-specific trajectories are expanded as follows:

$$\ln(\lambda_t^j) = \beta_0^j + \beta_1^j Age_t + \beta_2^j Age_t^2 + \alpha_1^j Fail_t + \alpha_2^j Separation_t$$

where Fail_t and Separation_t are indicator variables, measuring each individual's retention and parental breakup status at age t. Specifically, Fail_t equals 1 in all periods subsequent to an individual's first being retained at grade level and 0 in periods prior to retention. This definition of Fail_t tests for whether there is an enduring impact on delinquency of the initial experience of grade retention.[5] Separation is defined to equal 1 in all periods in which the boy is not living with both of his biological parents and equal to 0 in periods in which he is living with them. Thus, if biological parents go through periods of living and not living together, then Separation_t may change multiple times over the observation period. In contrast, Fail_t is defined so that it can only change value once (i.e., when and if the boy is retained at a grade level for the first time after the age of 11).

The model's parameter estimates are reported in **Table 4**. Grade retention was associated with statistically significant increases in violent delinquency for each trajectory group. By contrast, family breakup was only associated with a significant increase in violent delinquency for one group (Low 1). Thus, the findings imply that grade retention in adolescence was associated with subsequent increases in the violent delinquency of each trajectory group, whereas family breakup was not.

Figure 7 graphically depicts the impact of grade retention on the rising trajectory. Within the rising trajectory, subtrajectories for three scenarios based on retention are shown. One scenario (labeled "no retention") uses the parameter estimates for the rising trajectory to predict the expected offending rate from age 11 to 17 under the assumption that the boy is not retained during this time period. The second trajectory (labeled "retention at age 14") depicts an alternative scenario in which the boy is first retained at age 14 and thus from age 15 onward is behind grade level. Observe that these two trajectories are identical through age 14, but at age 15 the expected rate of violent delinquency for the retention scenario increases by about 0.6 acts compared to the continued nonretention scenario. The third trajectory, called the "group average" trajectory, uses the weighted average of the Fail_t variable to trace out the expected trajectory of individuals following the rising trajectory.

[5] Alternative specifications that interacted Fail_t with time from the retention event would allow an examination of whether the effect changed with time (e.g., attenuated). Still further elaboration of the specification could test whether subsequent events of grade retention amplified the effect of the initial event. Such elaborations, although substantively important, are beyond the scope of an illustrative analysis.

Table 4 Influence of grade retention and family breakup on trajectories of violence

Variable	Coefficient estimate	Z-score
Low 1 trajectory		
Intercept	13.26	3.27
Age	−18.95	−3.05
Age^2	5.76	2.47
Grade retention (>10)	0.37	2.28
Family breakup (<10)	0.39	2.30
Low 2 trajectory		
Intercept	−3.45	−2.10
Age	5.99	2.46
Age^2	−2.52	−2.84
Grade retention (>10)	0.36	5.79
Family breakup (<10)	−0.05	−0.60
Rising trajectory		
Intercept	−6.56	−3.38
Age	8.70	3.20
Age^2	−2.42	−2.56
Grade retention (>10)	0.20	2.99
Family breakup (<10)	0.13	1.60
Declining trajectory		
Intercept	−8.52	5.40
Age	15.90	6.66
Age^2	−6.39	−7.15
Grade retention (>10)	0.27	4.78
Family breakup (<10)	0.08	1.28
Chronic trajectory		
Intercept	−6.29	−5.04
Age	11.64	6.42
Age^2	−4.12	−6.35
Grade retention (>10)	0.29	5.40
Family breakup (<10)	−0.01	−0.25

Prior to age 14, the group average trajectory lies above the two other trajectories because a portion of individuals in the group have already been retained, whereas no retention has occurred yet in either of the other two scenarios. After age 14, the group average is sandwiched between the nonretention and retention-at-age-14 trajectories because it reflects a composite trajectory of individuals who have and have not experienced grade retention. **Figure 7** illustrates the capability of the group-based methodology to communicate the findings from a complicated statistical model in a more easily comprehended graphical format.

How can GBTMs be leveraged to facilitate causal inference? Throughout this section, the terms "impact" and "effect on" have been used to describe the statistical association of grade retention and parental separation with trajectories of violent delinquency. Group-based trajectory modeling has no special immunity to the hazards of drawing causal inferences from nonexperimental data. Recent work described in Haviland et al. (2007, 2008; Haviland & Nagin 2005) attempts to reduce this hazard by combining GBTM and propensity score matching (Rosenbaum & Rubin 1983). The goal of combining GBTM and propensity score matching is to recreate the desirable features of experimental designs in nonexperimental longitudinal data by creating balance between "treatment" and "control" groups that have formed naturally over the course of an observational study (Haviland et al. 2007).

Importing key attributes of experiments to the analyses of nonexperimental longitudinal data provides investigators with a way of evaluating the impact of turning-point events such as grade retention or therapeutic interventions on developmental trajectories. It does this in several ways. First, propensity score matching within trajectory group is employed in an effort to create a control group that is comparable to the treated group with respect to the observed covariates. By matching within trajectory, GBTM provides a developmental view of "comparable" by matching treated individuals with individuals who were not treated but who appeared to be on a similar developmental pathway for the behavior under study prior to treatment. For example, individuals identified as following the life-course-persistent pathway of conduct problems in the Dunedin Multidisciplinary Health and Development Study (see **Figure 1**) are likely to serve as the most comparable matches for each other on both observed and unobserved covariates owing to their shared childhood origins and journey along a common developmental pathway. Second, the

trajectory groups provide a means of identifying developmentally meaningful subgroups in the population for whom treatment effects may vary. Identifying these heterogeneous treatment effects is often an important goal or subgoal of developmental research. For example, in our previous work we assessed whether substance exposure in early adolescence had differential effects on health depending on whether children were following a no-conduct problem versus a conduct-problem-prone trajectory (Odgers et al. 2008a). Third, because the adjustment for observed covariates is transparent, it encourages and frames consideration and discussion of the possible biases from unobserved covariates that were not controlled by matching.

The estimation of adjusted treatment effects is also an issue in randomized clinical trials where selective subject dropout can bias treatment effect estimates, and effects may vary across individuals or across groups of individuals. That is, researchers and clinicians want to know, "What is the true impact of an intervention?" and "For who is treatment most effective?" Although it is beyond the scope of this review, a growing body of research is attempting to address these types of questions by conducting intent-to-treat and complier-average causal effect modeling within a group-based modeling framework (see, for example, Brown et al. 2008, Jo et al. 2001, Muthén et al. 2002).

Dual Trajectory Modeling

This section demonstrates the use of the group-based trajectory framework to model the developmental course of two distinct but related outcomes. The resulting dual trajectory model provides a rich, yet easily comprehended, statistical summary of the developmental linkages between the two outcomes of interest. It can be used to analyze the connections between the developmental trajectories of two outcomes that are evolving contemporaneously (e.g., depression and alcohol use) or that evolve over different time periods that may or may not overlap (e.g., prosocial behavior in childhood and school achievement in adolescence).

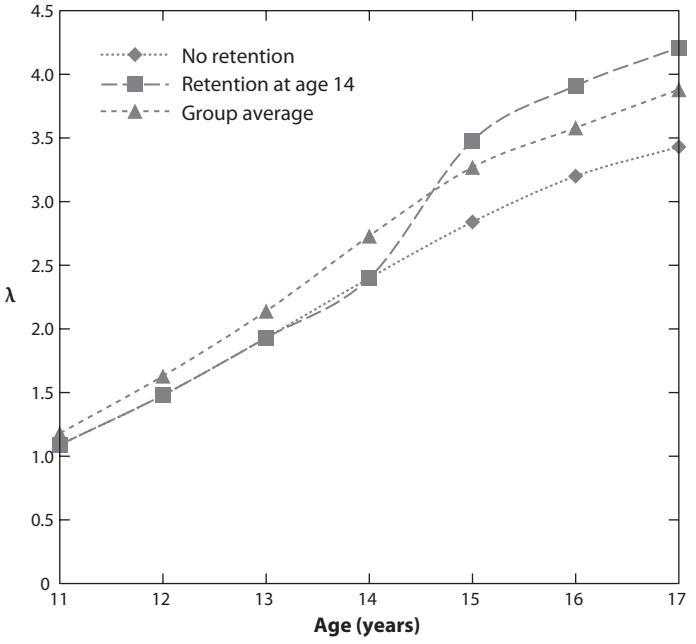

Figure 7

The impact of grade retention on the rising trajectory in the Montreal-based longitudinal study. (Data from Nagin & Tremblay 1999).

The dual model, which was first reported in Nagin & Tremblay (2001) and is described in greater detail in chapter 8 of Nagin (2005), has three key outputs: (*a*) trajectory groups for both measurement series, (*b*) the probability of membership in each trajectory group, and (*c*) probabilities linking membership in trajectory groups across behaviors. The linking probabilities are the key advance of the dual model. Compared to the use of a single summary statistic to measure the association of two outcomes, the linking probabilities provide a far more detailed and varied summary of the developmental connections between the two outcomes under study. The dual model can also be extended to analyze how the joint probabilities vary with individual-level characteristics.

Again using data from the Montreal-based longitudinal study, we provide an illustration in which we test Loeber's (1991) theory that oppositional behavior in childhood is linked to

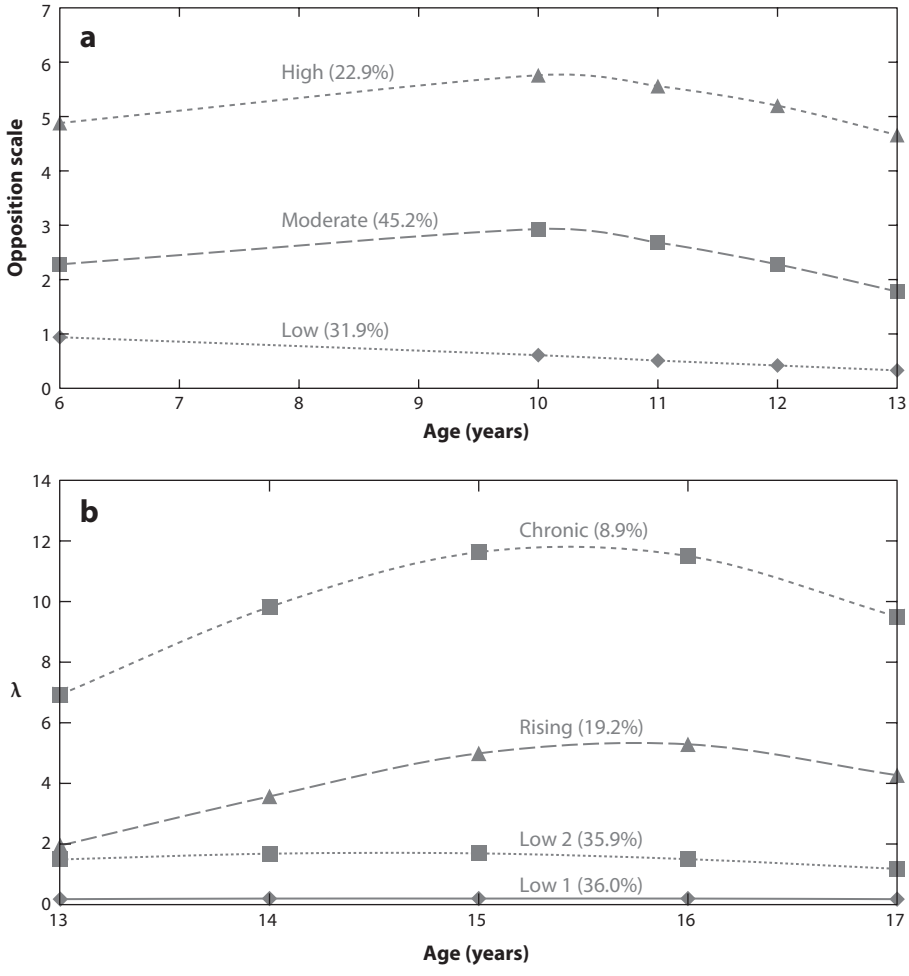

Figure 8

(*Panel a*) Trajectories of opposition from age 6 to 13. (*Panel b*) Trajectories of property delinquency from age 13 to 17.

property delinquency in adolescence. **Figure 8** displays the form of the three trajectories identified for these two behaviors. Panel *a* shows the trajectories of opposition from age 6 to 13, which were a product of the censored normal model. Panel *b* shows the four trajectories for property delinquency from age 13 to 17, which were estimated using the Poisson-based model.

Table 5 reports the key innovation of the joint trajectory model—three alternative representations of the linkage between childhood oppositionality and property delinquency in adolescence. One is the probability of membership in each of the property delinquency trajectories, conditional upon membership in each of the opposition trajectory groups. These probabilities, which are reported in panel A, can be interpreted as the probability of transitioning from each childhood opposition trajectory to each of the property delinquency trajectories. Because the probabilities are conditional upon membership in a given opposition trajectory group, each column of probabilities in panel A sums to 1. Panel B reports the reverse set of conditional probabilities: the probability of membership in each of the opposition trajectories conditional

Table 5 Relationship of oppositional behavior (age 6–13) with property delinquency (age 13–17)

A. Probability of delinquency group conditional on opposition group			
	Opposition trajectory group		
Property delinquency group	Low	Moderate	High
Low 1	0.54	0.29	0.23
Low 2	0.30	0.41	0.34
Rising	0.15	0.19	0.26
Chronic	0.01	0.11	0.17
B. Probability of opposition group conditional on delinquency group			
	Opposition trajectory group		
Property delinquency group	Low	Moderate	High
Low 1	0.48	0.37	0.15
Low 2	0.26	0.52	0.22
Rising	0.25	0.44	0.31
Chronic	0.03	0.55	0.42
C. Joint probability of opposition group and delinquency group			
	Opposition trajectory group		
Property delinquency group	Low	Moderate	High
Low 1	0.17	0.13	0.05
Low 2	0.10	0.19	0.08
Rising	0.05	0.08	0.06
Chronic	0.00	0.05	0.04

upon membership in each of the property delinquency trajectory groups. In this panel, each row of probabilities sums to 1. These probabilities measure the composition of each adolescent trajectory group in terms of the childhood trajectory of origin. The third form of representation, reported in panel C, is the joint probability of membership in a specific property delinquency trajectory and a specific opposition trajectory group. This panel enumerates the probabilities of all the possible combinations of opposition and property delinquency trajectory groups. Thus, the 12 joint probabilities sum to 1.

However represented, the results show a strong relationship between the developmental trajectories for these two behaviors. Panel A shows that the boys who were least oppositional from age 6 to 13 were least likely to be members of the two higher trajectories of property delinquency. Indeed, their probability of membership in the chronic trajectory was nearly zero. By contrast, the probability of transition to the chronic trajectory from the high opposition trajectory was 17%. Notwithstanding these overall tendencies, the transition probabilities also make clear that childhood opposition trajectory is not even close to being a certain predictor of the subsequent trajectory of property delinquency (panel B). With the exception of the chronic trajectory, all of the property trajectories were composed of large contingents of boys from each of the opposition trajectories. Compared to simply correlating the number of acts of property delinquency during each year from age 13 to 17 with opposition at each age from 6 to 13, the dual model provides a far richer, yet still comprehensible, summary of the relationships in the data.

The dual model can also be extended to allow the conditional probabilities linking trajectories across behaviors to vary as a function of individual-level variables. The following extension using the Montreal data tests whether experiencing two potential turning-point events at the time of the transition—not living with

both parents at age 11 or 12 and/or using drugs at age 12—alters the probability of transitioning from each of the childhood opposition trajectories to each of adolescent property delinquency trajectories. A full, technically detailed description of the basic dual trajectory model is described in section 8.7 of Nagin (2005). Here we provide only a brief outline. Let Y_1 and Y_2 denote the two longitudinal series to be modeled in a dual trajectory format, and let j and k index the J and K trajectory groups for Y_1 and Y_2, respectively. In the basic dual model, the direct products of model estimation are the parameters specifying the J and K trajectories of Y_1 and Y_2, the probability of membership in each of Y_1's J trajectories, π_j, and the conditional probability of membership in Y_2's k^{th} trajectory given membership in Y_1's J^{th} trajectory, $\pi_{k|j}$. To insure that $\pi_{k|j}$ is always between 0 and 1, this quantity is not directly estimated. Instead, a set of parameters $\gamma^0_{k|j}$ are estimated in which:

$$\pi_{k|j} = e^{\gamma^0_{k|j}} \Big/ \sum_k e^{\gamma^0_{k|j}} \quad j = 1, \ldots J.$$

Let w_i denote a vector of variables that are thought to be associated with $\pi_{k|j}$. The impact of these variables on $\pi_{k|j}$ is measured by a set of parameter vectors γ'_k whereby

$$\pi_{k|j}(w_i) = e^{\gamma^0_{k|j} + \gamma'_k w_i} \Big/ \sum_k e^{\gamma^0_{k|jk} + \gamma'_k w_i} \quad j = 1, \ldots J.$$

To avoid an unmanageable proliferation of parameters, the model assumes that the effects of the variables included in w_i do not depend upon the trajectory group membership for Y_1. Under this assumption, the various values of $\gamma'_{k|j}$ are equal across the J trajectory groups in Y_1 and therefore they can be denoted by γ'_k. Conceptually, this modification amounts to assuming that the influence of a particular variable on the probability of transition to a specific trajectory group k of Y_2, as measured by γ'_k, does not interact with trajectory membership for Y_1.

Table 6 reports the result of this extension of the dual trajectory model with the Montreal data. Reported in the table are estimates of the γ'_k coefficients and associated z-scores. For each of the trajectory groups, the estimates should be interpreted as the effect of its associated variable on the probability of transition to that delinquency trajectory relative to the low delinquency trajectory. The results indicate that controlling for childhood opposition trajectory, living in a broken home at age 11 or 12 has no effect on the transition probabilities with the possible exception of those related to the chronic trajectory. By contrast, drug use at age 12 has a positive and highly significant effect on the probability of transition to the low 2, rising, and chronic trajectories relative to the low 1 trajectory.

Table 6 Influence of broken homes and drug use on trajectory transition probabilities[a]

Variable	Coefficient	Z-score
Low 2		
Broken home 11–12	0.11	0.48
Drug use 12	0.78	6.00
Rising		
Broken home 11–12	0.12	0.47
Drug use 12	0.96	7.06
Chronic		
Broken home 11–12	0.63	1.81
Drug use 12	1.37	9.23

[a]Low 1 is the comparison group. "Broken home 11–12" refers to children not living with both parents at age 11 or 12, a turning-point event at the time of the transition from each of the childhood opposition trajectories to each of adolescent property delinquency trajectories.

Figure 9 provides a different perspective on these results. It reports the probability of transition to the rising property delinquency trajectory from each of the childhood opposition trajectories for two prototypical individuals—one who does not use drugs at age 12 and one who is at the seventy-fifth percentile of the age 12 drug use distribution. Observe that regardless of the childhood opposition trajectory group, drug use at age 12 greatly increases the risk of transition to the rising delinquency trajectory group. Practically, this type of analysis has the potential to identify putative risk factors that increase the likelihood of transitioning to

a risky developmental trajectory. Ideally, output from dual trajectory models can be used to help predict poor prognosis during key transition periods—an issue that clinicians, and in this case parents and policy makers, care deeply about.

Multitrajectory Modeling

The dual trajectory model is designed to measure the linkages between the trajectories of two distinct but related outcomes. Technically, it is straightforward to extend the dual model to more than two outcomes, but as a practical matter, the addition of more outcomes results in an unmanageable proliferation of probability matrices linking the trajectories for the various outcomes. For example, an extension to three behaviors requires estimation of three matrices of joint probabilities—one for outcome 1 to 2, another for outcome 1 to 3, and still another for outcome 2 to 3. An extension to four outcomes requires the estimation of six matrices of joint probabilities. Still, there are many circumstances where it would be valuable to link trajectories of three or more outcomes of interest. The multitrajectory model is designed to provide this capacity in a model of manageable size.

Figure 10 provides an illustrative application of the multitrajectory model to three distinct but related forms of conduct problems in adolescence that are measured in the Montreal data—violent delinquency from age 11 to 17, drug use from age 11 to 17 (mainly alcohol and marijuana), and precocious sexual behavior as measured by number of annual sexual partners from ages 13 to 17. As can be seen from **Figure 10**, each trajectory group is now defined not by one trajectory but by three trajectories, one for violent delinquency, a second for drug use, and a third for sexual activity. As in the basic model, the size of the group is measured by the probability of group membership. As in the basic model, this probability can be linked to characteristics of the individual. The results reveal a number of interesting patterns. Individuals in trajectory groups 1 and 2 engage in very little violence and are not sexually active. What distinguishes them is their use of drugs. Group 2 remains abstinent, whereas Group 1 is following a trajectory of rising use. Groups 3 and 4 follow similar rising trajectories of drug use and sexual activity. What distinguishes them is their involvement in violent delinquency. Trajectory group 3 is basically nonviolent, whereas trajectory group 4 follows a classic trajectory of rising and then falling violence over the 11 to 17 age range. Finally, the smallest trajectory group, group 5, follows the highest trajectories on all these behaviors. As this example illustrates, multitrajectory modeling provides a compact approach for summarizing the within-individual correspondence of multiple types of longitudinal data.

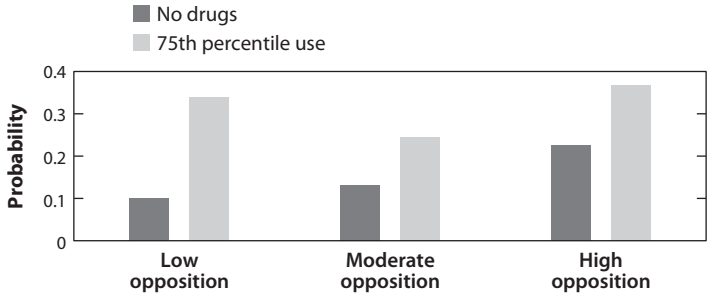

Figure 9

Probability of transition from each childhood opposition trajectory to the rising delinquency trajectory.

RECOMMENDED REPORTING GUIDELINES IN GROUP-BASED TRAJECTORY MODELING

Because GBTMs are relatively new in clinical research, few guidelines exist to assist researchers in making decisions regarding how to report model results. We of course acknowledge that any set of reporting principles will be incomplete and that a consensus on best practices when reporting results from GBTM is still evolving. With these caveats in mind, the following recommendations are provided to encourage the presentation of results from group-based models in a transparent, understandable, and replicable fashion. The list that

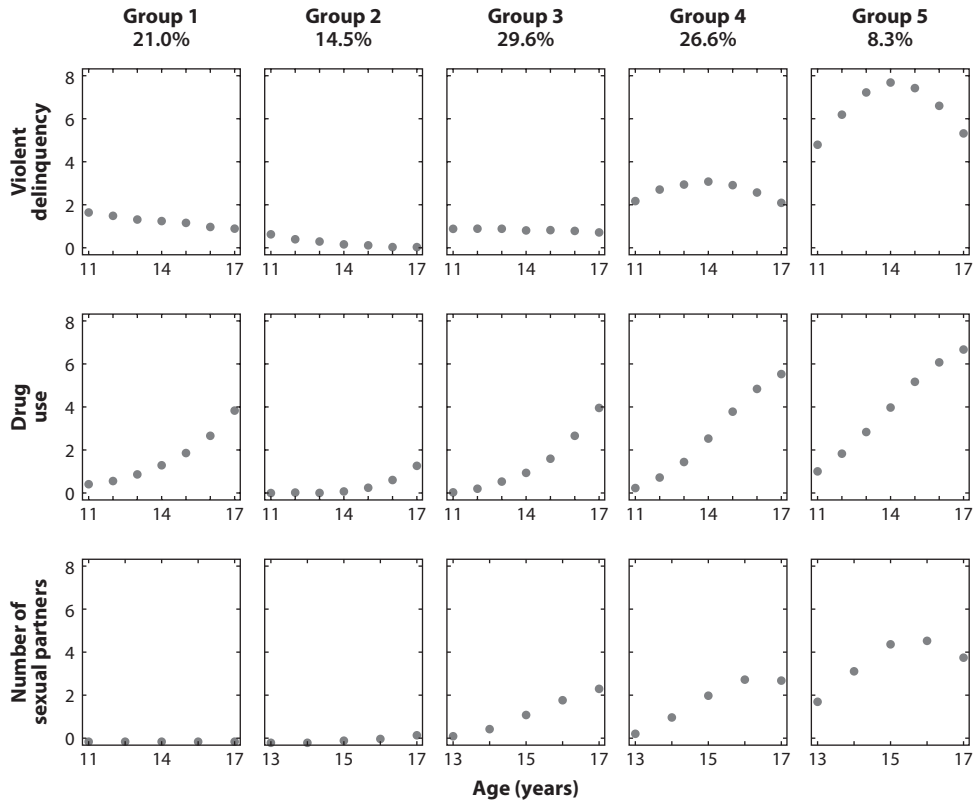

Figure 10

Multi-trajectory modeling in the Montreal-based longitudinal study. (Data from Nagin & Tremblay 1999.)

follows includes a nontechnical summary of basic reporting requirements for these types of models. For those conducting group-based analyses in a structural equation modeling framework, we also recommend a review of reporting guidelines and principles summarized by McDonald & Ho (2002).

1. There is no such thing as an all-purpose statistical method. The chosen method must fit the research question being asked and the data being analyzed. It is important, then, that researchers explicitly articulate their rationale for using GBTM. For example, is the purpose to test a taxonomic theory, or is a GBTM being used as an alternative to a GCM for summarizing population variation in the form of latent developmental strata? Alternatively, GBTM may be used for purely exploratory purposes. Whatever the rationale, it should be articulated.
2. Beyond providing a sound rationale for the use of GBTM for the given research question, it is important that the researcher address important technical issues that specifically pertain to the application of GBTM. These include (*a*) a clear explanation for the choice of the number of groups included in the final model. As discussed above in The Number of Groups and Extraction of "Fictitious" Groups section, this choice should not be reduced to a mechanical choice based on a specific fit statistic. Rather, the justification for the final model should be based on a combination of formal statistical criteria as well as the substantive usefulness and validity of the model as it

relates to the research questions under consideration. (*b*) The shape of each trajectory depends on the order of the polynomial used to model it. Therefore, it is important that the order of the polynomial for each trajectory be high enough to match the shape that theoretically might emerge from the data. This is particularly important in modeling what may be cyclical phenomena, such as symptom patterns exhibited by individuals suffering from bipolar disorder. (*c*) The adequacy of the chosen model should be demonstrated by criteria such as those recommended by Nagin (2005) and outlined in The Number of Groups and Extraction of "Fictitious" Groups section above.

3. Whenever possible, it is good practice to make programming script available by request or online for those interested in replicating the trajectory solution in their own data. In structural equation modeling applications, a correlation matrix and a vector of means and standard deviations could also be provided, which will allow researchers to fit alternative models to the data.
4. Beyond addressing issues that are specific to the application of GBTM, it is important that researchers also adhere to sound reporting guidelines that apply to the use of any statistical method. These include reporting basic univariate information for the variables that will be input into the trajectory model, including indices of central tendency, variability, and skew. This type of full disclosure will allow the reader to evaluate whether appropriate steps were taken to account for deviations from nonnormality in the data. Sound reporting guidelines also include describing the amount and type of missing data across assessments and detailing the type of missing data technique, if any, that was applied in the analysis.

CONCLUSIONS

A hallmark of modern longitudinal studies used in clinical research is the variety and richness of measurements that are obtained about the study members and their circumstances. Less often acknowledged is the fact that this abundance of information is accompanied by a difficult companion—complexity. Commonly, researchers are confronted with the dilemma of how best to explore and communicate the rich set of measurements at their disposal without increasing the analytical complexity to the point where the lessons to be learned from the data are lost on them and their audience. By segmenting the data into trajectory groups, the group-based approach to studying the developmental course of psychopathology provides a powerful statistical tool for summarizing large amounts of data in an easily comprehensible fashion. This approach also provides a vehicle for empirically testing long-standing theories in developmental psychopathology with a taxonomic dimension.

Within applied contexts, clinical researchers have begun to embrace this new set of tools to evaluate treatment effects and explore individual variation in response to clinical interventions. Moving forward, opportunities exist for a mutually beneficial exchange between applied clinical researchers and methodologists who are extending the capabilities of GBTMs. For example, statistical models are constantly evolving in response to clinicians' questions regarding group-specific responses to intervention and how to best address nonrandom attrition from clinical trials. At the same time, the ability to empirically test group-based models is encouraging researchers and clinicians to revise and more clearly articulate predictions from their taxonomic and developmental theories. Ideally, these types of exchanges between applied researchers and methodologists will enhance the potential for discovery in clinical research while at the same time help to drive the continued evolution of trajectory-based models and their extensions.

SUMMARY POINTS

1. Group-based trajectory models are increasingly being applied to address questions related to the development of mental disorders over time and, more recently, are being used to facilitate causal inference in situations where randomization to treatment condition is not possible.
2. GBTM and GMM represent two similar classes of models designed to identify clusters of individuals following a similar developmental trajectory on an outcome of interest; however, the two methods depart with respect to their underlying motivations for applying finite mixture modeling.
3. Selection of the best-fitting trajectory model based only on formal statistical criteria and mechanical application of the method may result in poor model choices. The most useful application of GBTMs will involve a careful weighting of formal criteria against explanatory power and usability in the analyses.
4. There is a need to move away from the interpretation of trajectory groups as distinct entities and instead view trajectory groups as providing approximations for a more complex reality through the identification of clusters of individuals following similar trajectories over time.
5. GBTMs have a wide range of applications in clinical research, including mapping the developmental course of one or more trajectories over time, identifying predictors of trajectory group membership, and evaluating the impact of turning-point events or therapeutic interventions. Recent extensions are also helping investigators to refine their estimation of treatment effects from both observational and randomized control trials.
6. Few established guidelines exist for the reporting of GBTMs. This review provides a set of general principles to encourage transparent and understandable reporting of GBTM results. We hope that developers and users of these models will continue to refine our recommendations with the goal of establishing an agreed upon set of reporting guidelines for the field.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This work was funded in part by the National Science Foundation (SES-99113700; SES-0647576) and the National Institute of Mental Health (RO1 MH65611-01A2). Candice Odgers is a William T. Grant Scholar.

LITERATURE CITED

Akaike H. 1974. New look at statistical-model identification. *IEEE Trans. Automatic Control* 19:716–23

Appel TA. 1987. *The Cuvier-Geoffroy Debate: French Biology in the Decades before Darwin*. New York: Oxford Univ. Press

Bauer DJ. 2007. Observations on the use of growth mixture models in psychological research. *Multivar. Behav. Res.* 42:757–86

Bauer DJ, Curran PJ. 2003. Distributional assumptions of growth mixture models: implications for overextraction of latent trajectory classes. *Psychol. Methods* 8:338–63

Baumol W. 1992. On my attitudes: sociopolitical and methodological. In *Eminent Economists: Their Life Philosophies*, ed. M Szenberg, pp. 51–59. London: Cambridge Univ. Press

Belsky J, Steinberg L, Draper P. 1991. Childhood experience, interpersonal development, and reproductive strategy: an evolutionary-theory of socialization. *Child Dev.* 62:647–70

Bergman LR. 1998. A pattern-oriented approach to studying individual development: snapshots and processes. In *Methods and Models for Studying the Individual*, ed. RB Cairns, LR Bergman, J Kagan, pp. 83–121. Thousand Oaks, CA: Sage

Bollen KA, Curran PJ. 2006. *Latent Curve Models: A Structural Equation Approach*. New York: Wiley

Brown CH, Wang W, Kellam SG, Muthen BO, Petras H, et al. 2008. Methods for testing theory and evaluating impact in randomized field trials: intent-to-treat analyses for integrating the perspectives of person, place, and time. *Drug Alcohol Depend.* 95:S74–104

Bryk AS, Raudenbush SW. 1987. Application of hierarchical linear models to assessing change. *Psychol. Bull.* 101:147–58

Caspi A. 1998. Personality development across the life course. In *Handbook of Child Psychology*, ed. N Eisenberg, W Daom, pp. 311–88. New York: Wiley

Christ M, Krishnan R, Nagin D, Günther O. 2002. *An empirical analysis of Web site stickiness*. Presented at Europ. Conf. Inform. Syst., Gdansk, Poland

Cloninger CR. 1986. A unified biosocial theory of personality and its role in the development of anxiety-states. *Psychiatr. Dev.* 4:167–226

Cloninger CR. 1987. A systematic method for clinical description and classification of personality variants: a proposal. *Arch. Gen. Psychiatry* 44:573–88

Côté S, Boiven M, Liu X, Nagin DS, Zoccolillo M, Tremblay RE. 2009. Depression and anxiety symptoms: onset, developmental course and prediction during early childhood. *J. Child Psychol. Psychiatry* 50:1201–8

Dekker MC, Ferdinand RF, van Lang NDJ, Bongers IL, Van Der Ende JD, Verhulst FC. 2007. Developmental trajectories of depressive symptoms from early childhood to late adolescence: gender differences and adult outcome. *J. Child Psychol. Psychiatry* 48:657–66

Farrington DP, West DJ. 1990. The Cambridge Study in Delinquent Development: a prospective longitudinal study of 411 males. In *Criminality: Personality, Behavior, and Life History*, ed. HJ Kerner, G Kaiser, pp. 115–38. Berlin: Springer-Verlag

Feldman BJ, Masyn KE, Conger RD. 2009. New approaches to studying problem behaviors: a comparison of methods for modeling longitudinal, categorical adolescent drinking data. *Dev. Psychol.* 45:652–76

Foster H, Nagin DS, Hagan JE, Costello J, Angold A. 2010. Specifying criminogenic strains: stress dynamics and conduct disorder trajectories. *Deviant Behav.* In press

Goldstein H. 1995. *Multilevel Statistical Models*. London: Arnold

Halliday-Boykins CA, Henggeler SW, Rowland MD, DeLucia C. 2004. Heterogeneity in youth symptom trajectories following psychiatric crisis: predictors and placement outcomes. *J. Consult. Clin. Psychol.* 72:993–1003

Haviland A, Nagin DS, Rosenbaum PR. 2007. Combining propensity score matching and group-based trajectory analysis in an observational study. *Psychol. Methods* 12:247–67

Haviland A, Rosenbaum PR, Nagin DS, Tremblay RE. 2008. Combining group-based trajectory modeling and propensity score matching for causal inferences in nonexperimental longitudinal data. *Dev. Psychol.* 44:422–36

Haviland AM, Nagin DS. 2005. Causal inferences with group-based trajectory models. *Psychometrika* 70:557–78

Heckman J, Singer B. 1984. A method for minimizing the impact of distributional assumptions in econometric-models for duration data. *Econometrica* 52:271–320

Holyoak KJ, Spellman BA. 1993. Thinking. *Annu. Rev. Psychol.* 44:265–315

Hu MC, Muthen B, Schaffran C, Griesler PC, Kandel DB. 2008. Developmental trajectories of criteria of nicotine dependence in adolescence. *Drug Alcohol Depend.* 98:94–104

Comprehensive review of the strengths and limitations of growth mixture models in psychology.

Describes how combining GBTM and propensity score matching can improve causal inferences in nonexperimental data.

Classic paper detailing the use of finite mixture models to approximate unknown distributions.

Jester JM, Nigg JT, Buu A, Puttler LI, Glass JM, et al. 2008. Trajectories of childhood aggression and inattention/hyperactivity: differential effects on substance abuse in adolescence. *J. Am. Acad. Child. Adolesc. Psychiatry* 47:1158–65

Jo B, Muthén BO, Marcoulides GA, Schumacker RE. 2001. Modeling of intervention effects with noncompliance: a latent variable approach for randomized trials. In *New Developments and Techniques in Structural Equation Modeling*, ed. GA Marcoulides, RE Schumacker, pp. 57–87. New York: Erlbaum

Jones BL. 2001. *Analyzing longitudinal data with mixture models: a trajectory approach*. PhD thesis. Carnegie Mellon Univ., Pittsburgh, PA

Kandel D. 1975. Stages in adolescent involvement in drug-use. *Science* 190:912–14

Kasen S, Cohen P, Skodol AE, Johnson JG, Smailes E, Brook JS. 2001. Childhood depression and adult personality disorder: alternative pathways of continuity. *Arch. Gen. Psychiatry* 58:231–36

Kim J, Cicchetti D. 2006. Longitudinal trajectories of self-system processes and depressive symptoms among maltreated and monmaltreated children. *Child Dev.* 77:624–39

Kochanska G. 1997. Multiple pathways to conscience for children with different temperaments: from toddlerhood to age 5. *Dev. Psychol.* 33:228–40

Little RJA. 1993. Pattern-mixture models for multivariate incomplete data. *J. Am. Stat. Assoc.* 88:125–34

Lo YT, Mendell NR, Rubin DB. 2001. Testing the number of components in a normal mixture. *Biometrika* 88:767–78

Loeber R. 1991. Questions and advances in the study of developmental pathways. In *Rochester Symposium on Developmental Psychopathology*, ed. D Cicchetti, S Toth, pp. 97–116. Rochester, NY: Univ. Rochester Press

MacCallum RC, Kim C, Malarkey WB, Kiecolt-Glaser JK. 1997. Studying multivariate change using multilevel models and latent curve models. *Multivar. Behav. Res.* 32:215–53

Magnusson D. 1998. The logic and implications of a person-oriented approach. In *Methods and Models for Studying the Individual*, ed. RB Cairns, LR Gergman, J Kagan, pp. 33–64. Thousand Oaks, CA: Sage

Maguin E, Loeber R. 1996. Academic performance and delinquency. In *Crime and Justice: A Review of Research*, ed. M Tonry, pp. 145–264. Chicago, IL: Univ. Chicago Press

Markman EM. 1989. *Categorization and Naming in Children: Problems of Induction*. Cambridge, MA: MIT Press

McArdle JJ, Epstein D. 1987. Latent growth-curves within developmental structural equation models. *Child Dev.* 58:110–33

McArdle JJ, Grimm KJ, Hamagami F, Bowles RP, Meredith W. 2009. Modeling life-span growth curves of cognition using longitudinal data with multiple samples and changing scales of measurement. *Psychol. Meth.* 14:126–49

McDonald RP, Ho MHR. 2002. Principles and practice in reporting structural equation analyses. *Psychol. Meth.* 7:64–82

McLachlan G, Peel D. 2004. *Finite Mixture Models*. New York: Wiley-Intersci.

Meredith W. 1993. Measurement invariance, factor analysis and factorial invariance. *Psychometrika* 58:525–43

Meredith W, Tisak J. 1990. Latent curve analysis. *Psychometrika* 55:107–22

Outlines expectations for one of the most widely cited taxonomies in developmental research.

Moffitt TE. 1993. Adolescence-limited and life-course-persistent antisocial-behavior: a developmental taxonomy. *Psychol. Rev.* 100:674–701

Mora PA, Bennett IM, Elo IT, Mathew L, Coyne JC, Culhane JF. 2009. Distinct trajectories of perinatal depressive symptomatology: evidence from growth mixture modeling. *Am. J. Epidemiol.* 169:24–32

Mustillo S, Worthman C, Erkanli A, Keeler G, Angold A, Costello EJ. 2003. Obesity and psychiatric disorder: developmental trajectories. *Pediatrics* 111:851–59

Muthén B. 1989. Latent variable modeling in heterogeneous populations. *Psychometrika* 54:557–85

Muthén B. 2001. Second-generation structural equation modeling with a combination of categorical and continuous latent variables: new opportunities for latent class/latent growth modeling. In *New Methods for the Analysis of Change*, ed. A Sayers, L Collins, pp. 291–322. Washington, DC: Am. Psychol. Assoc.

Review of recent advancements in growth mixture modeling and related techniques.

Muthén B. 2004. Latent variable analysis: growth mixture modeling and related techniques for longitudinal data. In *Handbook of Quantitative Methodology for the Social Sciences*, ed. D. Kaplan, pp. 345–68. Newbury Park, CA: Sage

Muthén B, Brown CH, Masyn K, Jo B, Khoo ST, et al. 2002. General growth mixture modeling for randomized preventive interventions. *Biostatistics* 3:459–75

Muthén B, Shedden K. 1999. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics* 55:463–69

Nagin D. 2005. *Group-Based Modeling of Development*. Cambridge, MA: Harvard Univ. Press

Nagin D, Pagani L, Tremblay RE, Vitaro F. 2003. Life course turning points: a case study of the effect of school failure on interpersonal violence. *Dev. Psychopathol.* 15:343–61

Nagin D, Tremblay RE. 1999. Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child Dev.* 70:1181–96

Nagin DS. 1999. Analyzing developmental trajectories: a semiparametric, group-based approach. *Psychol. Methods* 4:139–57

Nagin DS, Farrington DP, Moffitt TE. 1995. Life-course trajectories of different types of offenders. *Criminology* 33:111–39

Nagin DS, Jones BL, Haviland A. 2009. Group-based trajectory modeling extended to account for non-random subject attrition. Pittsburgh, PA: Carnegie Mellon Univ. Manuscript under review

Nagin DS, Land KC. 1993. Age, criminal careers, and population heterogeneity: specification and estimation of a nonparametric, mixed Poisson model. *Criminology* 31:327–62

Nagin DS, Tremblay RE. 2001. Analyzing developmental trajectories of distinct but related behaviors: a group-based method. *Psychol. Meth.* 6:18–34

Nesselroade JR. 1995. Testing for equivalence of measurement scales: simple structure and metric invariance reconsidered. A comment. *Multivariate. Behav. Res.* 30:119–20

Nylund KL, Asparouhov T, Muthén BO. 2007. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Struct. Equ. Model.* 14:535–69

Odgers CL, Caspi A, Broadbent JM, Dickson N, Hancox RJ, et al. 2007. Prediction of differential adult health burden by conduct problem subtypes in males. *Arch. Gen. Psychiatry* 64:476–84

Odgers CL, Caspi A, Nagin DS, Piquero AR, Slutske WS, et al. 2008a. Is it important to prevent early exposure to drugs and alcohol among adolescents? *Psychol. Sci.* 19:1037–44

Odgers CL, Moffitt TE, Broadbent JM, Dickson N, Hancox RJ, et al. 2008b. Female and male antisocial trajectories: from childhood origins to adult outcomes. *Dev. Psychopathol.* 20:673–716

Orcutt HK, Erickson DJ, Wolfe J. 2004. The course of PTSD symptoms among Gulf War veterans: a growth mixture modeling approach. *J. Trauma Stress* 17:195–202

Pagani L, Tremblay RE, Vitaro F, Boulerice B, McDuff P. 2001. Effects of grade retention on academic performance and behavioral development. *Dev. Psychopathol.* 13:297–315

Patterson GR, Debaryshe BD, Ramsey E. 1989. A developmental perspective on antisocial-behavior. *Am. Psychol.* 44:329–35

Peer JE, Spaulding WD. 2007. Heterogeneity in recovery of psychosocial functioning during psychiatric rehabilitation: an exploratory study using latent growth mixture modeling. *Schizophr. Res.* 93:186–93

Preacher KJ, Wichman AL, Briggs NE, MacCallum RC. 2008. *Latent Growth Curve Modeling*. Thousand Oaks, CA: Sage

Raftery AE. 1995. Bayesian model selection in social research. In *Sociological Methodology*, ed. P Marsden, pp. 111–63. Cambridge, MA: Blackwell Sci.

Raudenbush SW. 2001. Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annu. Rev. Psychol.* 52:501–25

Rosenbaum PR, Rubin DB. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70:41–55

Rubin DB. 1976. Inference and missing data. *Biometrika* 63:581–90

Rutter M, Yule B, Quintin D, Rowlands O, Yule W, Berger W. 1975. Attainment and adjustments in two geographical areas: the prevalence of psychiatric disorders. *Br. J. Psychiatry* 126:493–509

Schafer JL, Graham JW. 2002. Missing data: our view of the state of the art. *Psychol. Meth.* 7:147–77

Schwartz G. 1978. Estimating the dimension of a model. *Ann. Stat.* 6:461–64

Sonuga-Barke EJS, Van Lier P, Swanson JM, Coghill D, Wigal S, et al. 2008. Heterogeneity in the pharmacodynamics of two long-acting methylphenidate formulations for children with attention deficit/hyperactivity disorder: a growth mixture modelling analysis. *Eur. Child Adolesc. Psychiatry* 17:245–54

Important paper introducing growth mixture modeling.

Comprehensive description of the theory and application of GBTM.

Important application of the censored normal GBTM to study the developmental origins of externalizing disorders.

Early comprehensive paper-length review of GBTM.

First application of GBTM featuring the ZIP model form.

Stulz N, Lutz W, Leach C, Lucock M, Barkham M. 2007. Shapes of early change in psychotherapy under routine outpatient conditions. *J. Consult. Clin. Psychol.* 75:864–74

Titterington DM, Smith AFM, Makov UE. 1985. *Statistical Analysis of Finite Mixture Models*. New York: Wiley

Tyrka AR, Graber JA, Brooks-Gunn J. 2000. The development of disordered eating: correlates and predictors of eating problems in the context of adolescence. In *Handbook of Developmental Psychopathology*, ed. JA Sameroff, M Lewis, SM Miller, pp. 607–24. New York: Kluwer Acad./Plenum

Van Ryzin MJ, Chatham M, Kryzer E, Kertes DA, Gunnar MR. 2009. Identifying atypical cortisol patterns in young children: the benefits of group-based trajectory modeling. *Psychoneuroendocrinology* 34:50–61

Weisburd D, Bushway S, Lum C, Yang SM. 2004. Trajectories of crime at places: a longitudinal study of street segments in the city of Seattle. *Criminology* 42:283–320

Willett JB, Sayer AG. 1994. Using covariance structure-analysis to detect correlates and predictors of individual change over time. *Psychol. Bull.* 116:363–81

Witkiewitz K, Masyn KE. 2008. Drinking trajectories following an initial lapse. *Psychol. Addict. Behav.* 22:157–67

Xie HY, Drake R, McHugo G. 2006. Are there distinctive trajectory groups in substance abuse remission over 10 years? An application of the group-based modeling approach. *Admin. Policy Ment. Health* 33:423–32

Annual Review of
Clinical Psychology
Volume 6, 2010

Contents