Pattern-Mixture Models for Addressing Nonignorable Nonresponse in Longitudinal Substance Abuse Treatment Studies

SUSAN M. PADDOCK, MARIA O. EDELEN, SUZANNE L. WENZEL, PATRICIA A. EBENER, WALLACE MANDELL

WR-441-HLTH
November 2006
PATTERN-MIXTURE MODELS FOR ADDRESSING NON-IGNORABLE NONRESPONSE IN LONGITUDINAL SUBSTANCE ABUSE TREATMENT STUDIES

Susan M. Paddock*, Maria O. Edelen*, Suzanne L. Wenzel*, Patricia A. Ebener*, Wallace Mandell§

* From RAND Health and the Drug Policy Research Center, Santa Monica, California.
§ From Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland and The Phoenix House Foundation, New York, New York.

+ Address correspondence to:
Susan M. Paddock, PhD, RAND, P.O. Box 2138, Santa Monica, CA 90407-2138.
Voice: 310-393-0411, ext. 7628; Fax: 310-260-8155; Email: paddock@rand.org

November 10, 2006
ABSTRACT

Missing data is a pervasive problem in longitudinal treatment research studies. Missing data due to study non-completion complicate the task of drawing conclusions about the effect of a treatment or policy on a measure of interest (e.g., a process measure or outcome). Biased estimates of change over time in a measure could result if attrition is related to the constructs that are being measured. Identifying potential biases in estimates is critical for research involving longitudinal assessments. The pattern-mixture model (PMM) provides a way to understand and account for attrition when analyzing data and communicating results to research stakeholders. This paper demonstrates the use of PMMs in a study of the quality of care in therapeutic communities (TCs) using the Dimensions of Change Instrument (DCI) to measure longitudinal client-level change and TC treatment process. The effect of choice of missing data pattern and its effects on conclusions drawn from analyses is highlighted along with the role of clinical expertise in formulating PMMs.

Key Words: attrition; longitudinal data; missing data; non-ignorable non-response; substance abuse treatment; therapeutic community
1. INTRODUCTION

Research on the quality of care in substance abuse treatment has received increased attention in recent years (Etheridge and Hubbard, 2000). The increased interest in understanding the process by which substance abuse treatment produces client outcomes has been fueled by the social and economic forces of health care reform and purchasers’ requests for outcomes and accountability data (Moos, 2000). Donabedian (1988) conceptualizes quality of care as being composed of the three interrelated elements of program structure, treatment process, and client outcomes. Improving quality of care relies upon understanding how components of treatment program structure and process are responsible for producing client outcomes (Simpson et al., 1997; Simpson et al., 2000).

In the current study the treatment process was characterized as part of a larger effort to improve the quality of care in the therapeutic community (TC) modality of substance abuse treatment. The TC is based on the view that substance abuse is a disorder that involves the whole person. The goal of the therapeutic community is to change the client’s lifestyle and for the client to rehabilitate through mutual- and self-help (De Leon and Rosenthal, 1979, 1989; De Leon, 2000). The treatment in the therapeutic community consists of all interactions among staff and clients in formal and informal settings (De Leon, 2000; Hubbard et al., 1989). The therapeutic community modality has been shown to improve long-term outcomes, such as reducing post-treatment drug use and criminal behavior, and in increasing employment and other social functioning (Hubbard et al., 1989; Hubbard et al., 1997; Simpson et al., 1997; Gerstein et al., 1990;
National Treatment Improvement Evaluation Study [NTIES], 1996; Messina et al., 2000). Understanding the process by which positive outcomes are achieved is a relatively new endeavor in TC research, a central component of which was the development of the Dimensions of Change Instrument (DCI) (Orlando et al., in press), which is a client-based measure of treatment process. In the current study, the DCI was administered at several time points during each client’s treatment tenure (at baseline and at 1, 3, 6, and 9 months following treatment enrollment).

Missing data on the client-level treatment process complicates the task of understanding the links among treatment structure, process, and outcomes. Measuring treatment process requires repeated assessments at critical time points during treatment tenure. Missing periodic process assessment data are common, as clients spend varying lengths of time in treatment and thus naturally complete varying numbers of scheduled assessments. Sixty-seven percent of clients in the present study left treatment before the final scheduled in-treatment assessment at 9 months. The occurrence of comparable levels of missing in-treatment data is pervasive in longitudinal substance abuse treatment studies, where attrition is often reported to be as high as 50-70% (Baekeland and Lundwall, 1975; Primm et al., 2000; Yang and Shoptaw, 2005). The client’s length of stay in treatment is a crucial consideration in treatment effectiveness; longer lengths of stay (usually at least three months) have consistently been associated with better post-treatment outcomes (Hubbard et al., 1989; Simpson, 1993). If the achievement of positive post-treatment outcomes is to be explained by treatment process measures, then it is important to correctly characterize the process by incorporating appropriate techniques for accounting for missing treatment process measures data.
The importance of properly addressing and understanding the effect of missing data in longitudinal substance abuse treatment studies is widely acknowledged. For example, Yang and Shoptaw (2005) characterize missing data in a smoking cessation trial as being intermittently missing (i.e., where there is at least one observed value for a study participant following a missing value), or missing data due to attrition (i.e., missing values caused by participant withdrawal from the study). They explore the plausibility of scenarios in which the missing data occurrence could be predicted from observable data.

Cohort attrition has been associated with observed client or treatment characteristics such as severity of drug use at baseline, social stability, and discrepancies between client and therapist expectations (Baekeland and Lundwall, 1975; Gottheil et al., 1997). Simply omitting study participants with missing values from the analysis undoubtedly results in biased estimates, since these clients are unlikely to be a random sample of all clients. Omitting participants with missing data assumes data are missing completely at random (MCAR) under the assumed analysis model. A more sophisticated approach assumes that, conditional on observed client characteristics and outcomes, the clients with missing data are not systematically different from those with fully observed data. Little and Rubin (1987) classify missing data of this type to be ‘missing at random (MAR),’ in which the propensity for data to be missing could be related to covariates or to prior assessments of the outcome measure. Random effects modeling of repeated observations over time (Laird and Ware 1982) is appropriate when the MAR assumption holds. While random effects modeling is an improvement over deleting clients with missing values, it will be inadequate if the reasons for study attrition are related to the client’s underlying treatment process, for example if patient dissatisfaction with treatment
contributed to treatment dropout and patient satisfaction is a key outcome (Fiester & Rudestam, 1975).

Both the MAR and MCAR assumptions imply that the reasons for the missing data are “ignorable” (Little & Rubin, 1987) – in short, that the analyses can proceed without explicitly jointly modeling the missing data and outcomes distributions. An analysis of treatment process in the presence of missing data requires an alternative approach that would account for study attrition due to non-ignorable non-response and requires further model specification. While non-ignorability of missing data cannot be verified or empirically tested, it is important to examine sensitivity to inferences based on data in which it is suspected.

Pattern-mixture models (PMMs; Little, 1993, 1995; Rubin, 1977) have been used in many applications for which the possibility of longitudinal non-ignorable missing data is a concern (Fitzmaurice et al, 2001; Hedeker & Gibbons, 1997; Hogan & Laird, 1997; Pauler et al, 2003) and represent an alternative if non-ignorable non-response is suspected. The idea behind PMMs is to explicitly model the missing data distribution by first identifying different patterns of missing data and then including parameters in the outcomes model that capture this effect. For example, one could simultaneously model whether one is a “study completer” versus a “non-completer” and then include an indicator variable for “study non-completers” as a predictor in a regression model of the outcome measure of interest and examine its interaction with key study covariates. In our study, data from non-completers might have been unavailable due to their departure from treatment or their unwillingness to complete assessments while in treatment.
This paper presents an illustrative analysis of the treatment process from an ongoing study of the quality of care in the therapeutic community. The difficulties of characterizing changes in the treatment process over time when data are non-ignorably missing are examined to illustrate how to use the pattern-mixture model to guide examination of the treatment process. General considerations that substance abuse treatment researchers might make in similarly designed studies are examined and lead to emphasis on the nature of substance abuse treatment and the importance of incorporating knowledge about critical lengths of stay in constructing PMMs.

2. MOTIVATING STUDY: QUALITY OF CARE IN SUBSTANCE ABUSE TREATMENT

Treatment program and subjects. Data were collected on a sample of adults who were enrolled in community residential drug treatment programs that utilize the therapeutic community model. All programs are part of the Phoenix House Foundation, a national therapeutic community that operates 90 programs in New York, California, Texas, Florida, and New England. Data were collected from 519 clients in 12 treatment programs that are located in eight states. The participants in the study are drawn from the population of persons undergoing substance abuse treatment in the Phoenix House network of treatment programs during the data collection periods.

Measures. Client data on several measures were collected as part of an ongoing NIDA-sponsored study (Wenzel et al., 2001). Self-reported client characteristics were obtained as clients entered treatment on gender, age, and race/ethnicity (Hispanic/Latino,
African-American, White, Other), drugs used, type of admission, who referred the client to treatment (i.e., criminal justice system, self/family, other), and health behavior indices (e.g., prior drug treatment, age of first substance use). The Dimensions of Change in TC Treatment Instrument (DCI; Orlando et al., in press), a client-level treatment process measure, was administered at several time points during each client’s treatment tenure (at baseline and at 1, 3, 6, and 9 months following treatment entry). The DCI consists of 54 items assessed with a 5-point Likert-type scale indicating respondents’ extent of agreement with each item (1=not at all; 5=completely). These 54 items are summarized into eight scales representing dimensions that describe aspects of the community environment and personal development. Each of the eight scales has high internal consistency, with Cronbach’s alpha coefficients at or above 0.7 for all of the scales, and with four having alpha coefficients equal to or greater than 0.8. In this paper, we focus on 1 of the 8 DCI scales, the scale reflecting client perceptions of clarity and safety in treatment (CS), to illustrate the PMM application. Cronbach’s alpha for the CS score ranges from 0.83-0.87 across all 5 CS assessments. The CS scale is based on six items pertaining to how clear the client is about the purpose of the treatment program – e.g., “Requirements for program completion are clear” -- and about the client’s feelings of safety in the program – e.g., “This program is a safe place for people in treatment here.”

3. MODELING CHANGE OVER TIME IN THE CLARITY AND SAFETY TREATMENT PROCESS SCORE
The growth curve model was used to estimate the trajectory of the CS score over time in treatment (Laird & Ware, 1982; Bryk & Raudenbush, 1992). The basic growth curve model for individual $i$ in site $j$ ($j=1, \ldots, 12$ treatment sites) over $t$ time points ($t=0, 1, 3, 6, 9$ months following baseline) is:

$$y_{ijt} = \eta_{0ij} + \eta_{1ij}t + \epsilon_{ijt}, \quad (1a)$$

where $y_{ijt}$ is the CS treatment process score, the subscript, $j(t)$, indicates which treatment site, $j$, was utilized at time $t$, $\eta_{0ij}$ represents the baseline CS score (i.e., the growth intercept) for individual $i$ at clinic $j$, $\eta_{1ij}$ represents the change in CS scores over time (i.e., the growth rate), $t$ is the time of observation (months following baseline), and $\epsilon_{ijt}$ is the residual term that is normally distributed with mean 0 and variance $\sigma^2$ and is independent for all $i, j, \text{ and } t$. The growth parameters and their relationship to pre-treatment client characteristics can be expressed by

$$\eta_{0ij} = \alpha_{00j} + \alpha_{01}X_{ij} + \zeta_{0ij} \quad (1b)$$

$$\eta_{1ij} = \alpha_{10j} + \alpha_{11}W_{ij} + \zeta_{1ij} \quad (1c)$$

where $X_{ij}$ represents the $q_1$ pre-treatment client characteristics for person $i$ in clinic $j$ that are related to the growth intercept, and $W_{ij}$ represents the $q_2$ characteristics related to the growth rate. Both $X_{ij}$ and $W_{ij}$ are centered about their grand means prior to analysis to facilitate interpretation of the model parameters. $\alpha_{00j}$ is the mean baseline CS process score for clinic $j$, $\alpha_{01}$ is a vector of length $q_1$ of the average effects of client characteristics, $X_{ij}$, on the baseline CS score, $\alpha_{10j}$ is the mean change in CS score per unit time at clinic $j$, $\alpha_{11}$ is a vector of length $q_2$ that represents the effect of pre-treatment client characteristics. $X_{ij}$, on the baseline CS score, $\alpha_{10j}$ is the mean change in CS score per unit time at clinic $j$, $\alpha_{11}$ is a vector of length $q_2$ that represents the effect of pre-treatment client characteristics.
characteristics, $W_{ij}$, on the CS growth rate, and $\zeta_{0j}$ and $\zeta_{1j}$ are residual terms that are bivariately normally distributed and allowed to be correlated. In turn, the random site terms are modeled as follows:

$$\alpha_{00j} = \gamma_0 + \tau_{0j} \quad (1d)$$
$$\alpha_{10j} = \gamma_1 + \tau_{1j} \quad (1e)$$

where $\gamma_0$ and $\gamma_1$ are the mean baseline clarity and safety score and rate of change in clarity and safety, respectively, for the entire sample, and $\tau_{0j}$ and $\tau_{1j}$ are residual terms that are bivariately normally distributed and allowed to be correlated. Equations 1a-1e can be collapsed into a single equation as:

$$y_{ijt} = \gamma_0 + \alpha_{01}X_{ij} + \gamma_1t + \alpha_{11}W_{ij}t + \zeta_{0ij} + \zeta_{1ij}t + \tau_{0j} + \tau_{1j}t + \epsilon_{ijt} \quad (2)$$

The formulation of the growth curve model above assumes that everyone enrolled in the study will either remain in treatment long enough to complete all DCI assessments or that the missing assessments are ignorable. Like most analyses of in-treatment processes, this model will yield biased estimates if the DCI assessments are missing for non-ignorable reasons (Little, 1995). It is plausible that individuals who leave treatment before completing all of the DCI assessments would have had lower CS scores had they remained in treatment than those who complete all DCI assessments; in other words, the probability that a CS score is observed might be related to the unobserved value of that score or the underlying trajectory of change in CS, and would thus present a non-ignorable missing data problem. Table 1 shows the observed data patterns in this data set through the last available in-treatment assessment. Ninety-two clients (18%) were observed only at baseline, and 131 clients were observed at all assessment points. The
454 clients in patterns 1-5 who have consecutively measured CS scores starting at baseline represent 88% of the sample. Aside from the first five data patterns, the other patterns contain relatively few or no clients.

4. PATTERN-MIXTURE MODEL (PMM)

The primary means by which the effect of possibly non-ignorably missing data when modeling CS was accounted for in our analysis by using a pattern-mixture model (PMM). The PMM is a joint model of the outcomes of interest, Y, and missing data indicator variable, R, which indexes patterns of missing data – e.g., R$_i$=1 if client $i$ leaves treatment early so that some of his/her data are unobserved and R$_i$=0 otherwise. This joint model of Y and R can be re-expressed as a model of the missing data pattern, R, multiplied by a model of Y conditional on R. In contrast, the growth curve model (Equation 1) under the MAR assumption does not require jointly modeling Y and R.

In this example, the 519 clients have 15 response patterns (Table 1). Most of these patterns contain relatively few clients, so conditioning was done on relatively coarse patterns, e.g., examining completers versus non-completers (e.g., Hedeker and Gibbons, 1997; Pauler et al., 2003). Intermittently missing DCI assessments were assumed to be missing at random and so constructing patterns to account for this type of missing data was not pursued. Given the relatively modest level of intermittently missing data (64 clients, or 12%), conclusions should be robust with regard to the missing at random assumption (Yang and Shoptaw, 2005). Thus we initially assumed two missing data patterns based on whether clients completed the 9-month DCI assessment such that
the probability of being a non-completer of the 9-month assessment is $\pi$ and is $1 - \pi$ for completers. We fit a growth curve model for the distribution of CS scores given $R$ that is similar to Equation 1 but with Equations 1b and 1c modified as:

$$\eta_{0ij} = \alpha_{00j} + \alpha_{01}X_{ij} + R_i\alpha_{02}X_{ij} + \xi_{0ij}$$  \hspace{1cm} (3a)$$

$$\eta_{1ij} = \alpha_{10j} + \alpha_{11}W_{ij} + R_i\alpha_{12}W_{ij} + \xi_{1ij}$$  \hspace{1cm} (3b)$$

where $R_i$ indicates whether client $i$ is a non-completer, and $\alpha_{02}$ and $\alpha_{12}$ represent the differences in the fixed effects estimates of $X_{ij}$ and $W_{ij}$, respectively, for non-completers relative to completers. The probability of being a non-completer of the 9-month assessment, $R_i$, was estimated using a binomial distribution with probability $\pi$ (Little, 1995).

The parameterization of the PMM shown in Equation (3) will yield fixed effects regression parameter estimates, $\alpha_{01}$ and $\alpha_{11}$, for those who fail to complete the 9-month assessment ($R=1$) and the change in those coefficients for those who are completers ($R=0$) will be given by $\alpha_{02}$ and $\alpha_{12}$, respectively. The overall estimates for completers and for all persons averaged over missing data pattern are $\hat{\alpha}_c = \hat{\alpha}_{c1} + \hat{\alpha}_{c2}$ and $\hat{\alpha}_a = \hat{\alpha}_{a1} + \hat{\pi}\hat{\alpha}_{a2}$, respectively, where the asterisk (*) equals 0 for the growth intercept equation (Equation 3a) and 1 for the growth rate equation (Equation 3b). The additional variation in $\hat{\alpha}_c$ and $\hat{\alpha}_a$ due to the fact that these estimators are themselves functions of estimates must be accounted for when estimating the standard errors of these estimates by using the delta method (Hedeker and Gibbons, 1997; Hogan and Laird, 1997; Pauler et al, 2003; see Appendix for further details).
All statistical analyses were conducted in R, a freely available statistical software package (R Development Core Team, 2005). The “nlme” mixed effects modeling package (Pinheiro et al., 2005) was used for growth curve modeling and the “assist” package (Wang and Ke, 2005) for matrix manipulation routines that are required to obtain estimates that average over patterns.

5. EXAMPLE: TREATMENT PROCESS IN THE THERAPEUTIC COMMUNITY

Results were examined under several plausible missing data scenarios for specifying the patterns of the PMM and to assess the sensitivity of the results for those different scenarios in order to highlight potential biases resulting from non-ignorable study attrition (Little, 1993). Treatment site- and individual-level random effects were included in all the PMMs, but for clarity only the fixed-effects parameter estimates are reported in Tables 2-4. Table 2 shows the results of fitting the PMM when the patterns reflect grouping the 347 clients who do not complete the final 9-month assessment versus the 172 who do. Based on variable selection that was conducted as part of exploratory analyses that involved all 8 DCI scores (not shown here), the covariates listed in the table under “Predictors of CS at baseline” were used to predict the baseline CS value, while age and prior drug treatment predicted both the baseline CS score and its growth rate. Section (a) in Table 2 shows the regression coefficients for those who completed the 9-month assessment. The coefficient estimates for predictors of the baseline CS score correspond to $\alpha_{01}$ in Equation 3a and the estimates for the predictors of the growth rate
correspond to $\alpha_{11}$ in Equation 3b. Section (b) in Table 2 shows the analogous estimates for those who did not complete the 9-month assessment. Statistically significant differences for the predictor variables across the patterns are denoted on Table 2 as superscripts in the leftmost column and are derived from the $\alpha_{02}$ and $\alpha_{12}$ parameters in Equations 3a-b (estimates of these from the PMM are not shown in the table) along with $\hat{\pi}$, $\alpha_{00}$ and $\alpha_{10}$. The only significant differences between completers and non-completers were for those referred by the criminal justice system to treatment – the non-completers in this group had, on average, lower CS scores at baseline ($p<0.01$) -- and age was a significant predictor of change in CS over time among non-completers but not among completers ($p<0.05$). Comparison of columns (a) and (b) on Table 2 does not suggest major differences between completers and non-completers – most notably, CS is significantly changing over time in both groups at a similar rate. Section (c) on Table 2 shows the results of the PMM for all clients when averaging over the two patterns, and section (d) on Table 2 shows the parameter estimates when fitting a standard random effects model under the missing at random assumption. These models have very similar results, but the standard errors of estimate from the MAR model tend to be smaller. In the model assuming MAR, CS is increasing significantly over time, as indicated by the time coefficient, 0.0405 (s.e. 0.0076, $p<0.0001$). Children of substance abusers had lower CS scores at baseline (coefficient = -0.0984 (s.e. 0.0511), $p<0.0546$), as did Whites (coefficient = -0.3056 (s.e. 0.0538), $p<0.0001$) and those with a lower age of first drug use (coefficient = 0.6712 (s.e. 0.3327), $p<0.0440$). These MAR results are very similar to those in section (c) of Table 2, suggesting that any possible non-ignorable differences in
non-response for completers versus non-completers are not large enough to affect inferences drawn from the data analysis.

In the context of this substance abuse treatment study, defining patterns of observed data according to completion versus non-completion might be inappropriate, because members from each group may not be fundamentally different– i.e., those who complete the 6-month assessment might not be very different from those who complete the 9-month assessment. It may be more useful to group clients by the type of “dose” of treatment they receive such as a minimally effective treatment dose. Phoenix House has suggested that a client should stay longer than 3 months to receive some benefit from treatment. This is in accord with the findings of other researchers that lengths of stay of at least 3 months are associated with better post-treatment outcomes (Hubbard et al., 1989; Simpson 1993). Thus, patterns were created based on the completion of either the 6- or 9-month assessment versus earlier assessments. In addition, we also examined results when the longer-stay pattern was composed of those with CS assessments at 3, 6 or 9 months. This exploration is important because the reasons for missing data vary according to the time it occurs during the course of substance abuse treatment (Howard et al., 1990).

A missing data pattern for those who have only provided baseline CS scores (Pattern 1 on Table 1) could be specified in theory. The problem with doing this is the client-specific parameters that characterize the change in CS over time (i.e., the $\eta_{ij}$) cannot be estimated from the data for clients with only one observed CS score. Thus, the CS change parameter can only be estimated from the data by imposing special constraints (Rubin 1977; Little 1993, 1995). For example, these parameters can be identified by
combining patterns with only baseline observations with other patterns (e.g., Pauler et al., 2002).

The structure of Tables 3 and 4 mirrors that of Table 2. The results given in Table 3 are for a PMM in which the non-response patterns being modeled are for those who complete at least the 6-month assessment versus those who complete only earlier assessments (i.e., distinguishing clients according to receipt of adequate dose, longer than 3 months). The CS score changes significantly over time, regardless of which subgroup is being examined. Shorter stay clients have lower CS scores at baseline than do longer stay clients and are less likely to have been referred to treatment by the criminal justice system (section (a) of Table 3). The results are quite different, however, if the patterns are selected to be those who complete the baseline and/or 1-month assessments versus others. Section (a) of Table 4 shows that CS changes over time for longer stay clients, while section (b) of Table 4 shows that shorter stay clients do not exhibit change over time. Longer stay clients who were children of substance abusers and those who were white had lower CS scores at baseline. When estimates from these two patterns are averaged the overall effect is that CS is not changing over time (section (c) of Table 4). This is in contrast to the MAR model (section (d) of Table 2) and the PMM model for the other patterns examined (section (c) of Tables 2 and 3). Together, section (d) of Table 2 and section (c) of Tables 2-4 show a range of change experienced over time in the CS score by a client who enters the therapeutic community. The basis for selecting the best model among these choices cannot be ascertained from the data, as both non-ignorability and missing at random premises are a priori assumptions. In this situation expert knowledge is needed to guide the selection of a model. The currently available evidence
suggests that clients need to stay in treatment for 90 days or more before beginning to experience benefits. Based on this assumption, the model depicted on Table 3 would be the most appropriate for examining the change over time in CS scores in light of possible departures from the missing at random-based findings of section (d) of Table 2.

6. DISCUSSION

With the exception of the assumption of missing data being completely at random – i.e., subjects with missing data are a random sample of those with observed data – assumptions about the nature of the biases introduced by missing data are unavoidably subjective. Pattern-mixture models help clarify the gaps between missing-at-random-based analyses and the range of plausible biases induced by possibly non-ignorable missing data among clients who enter treatment. Based on our analyses and clinical judgment, we found that the random effects and PMM (90 days) yielded similar results, thus suggesting robustness of the random effects model to the possibility that persons who completed the 180- and/or 270-day surveys were non-ignorably different from others. Using that assumption when examining the overall PMM results shown here illustrate the findings for the CS score trajectory.

The PMM is limited by the available degrees of freedom as to how elaborate and specific the PMM can be. In the present case there could have been up to 15 patterns in the CS pattern-mixture model, but clearly the paucity of cases in many of the patterns relative to the complexity of the model made that number unfeasible. Recent methodological developments address this problem by proposing latent class dropout
classes (Roy, 2005) or latent pattern effects (Guo et al., 2004). While subjective model building choices are a limitation of the PMM, the more standard MAR-based analyses require similar assumptions.

The present analyses do not exhaust the many possibilities for constructing PMMs to address the questions raised by non-identified biases introduced by non-ignorable missing data. Other modeling strategies can be used such as setting the unidentified growth rate parameters equal to zero or to the expected value of the parameter for the fully-observed group (Rubin, 1977) or modeling the growth parameters across patterns as functions of the drop-out times (Little 1995; Mori et al., 1992; Wu & Bailey, 1988). A disadvantage of these approaches is that they are not easily implemented using standard statistical software packages.

Alternatives to the PMM exist for accommodating non-ignorable non-response, most notably selection models (Hogan & Laird, 1997; Little, 1995). Selection models involve jointly modeling the probability of non-response, R, and outcome, Y, by constructing first a model for R given Y and then a model for Y. Selection models require a detailed mathematical specification of the missing data mechanism (Little, 1995), which makes them appealing in situations in which that information is available. Very often, however, the exact nature of attrition is unknown, in which case PMMs are advantageous. Further, it is easier to utilize conventional statistical modeling software to fit PMMs versus selection models (Hogan and Laird, 1997).

We incorporated subject-matter expertise into our analyses when comparing the effects of using various patterns on analytical conclusions. Scharfstein et al. (2003) construct a Bayesian selection model to incorporate such expertise into their analysis.
Formalizing the sensitivity analyses described in this paper by adopting a Bayesian statistical approach is an intriguing future extension of this work. Under a Bayesian approach, prior probability distributions would be specified for model parameters such as the non-identified CS change parameters and the patterns of the PMM themselves using opinions elicited from substance abuse treatment experts. Accounting for model uncertainty in the selection of patterns in this way could address prior findings of under-coverage of interval estimates based on using a single PMM (Demertas and Schafer, 2003). While Bayesian approaches to PMMs have been proposed (e.g., Rubin 1977), we are unaware of serious applications in which expert opinion has been incorporated into the pattern-mixture modeling process in substance abuse treatment research.

ACKNOWLEDGEMENTS
This research was supported by the National Institute of Drug Abuse grant number R01-DA 14969 and the Agency for Healthcare Research and Quality grant number R03-HS 014805.

APPENDIX: DELTA METHOD TO DERIVE VARIANCE ESTIMATES FOR PMM PARAMETERS

The delta method approximation of the variance of $\hat{c}_a$ is $\hat{V}_a = \hat{\Delta}'\hat{V}\hat{\Delta}$. $\hat{V}$ is the estimated variance-covariance matrix of $(\gamma_0, \alpha_{01}, \alpha_1, \alpha_{02}, \alpha_{11}, \alpha_{12}, \pi)$, evaluated at $(\hat{\gamma}_0, \hat{\alpha}_{01}, \hat{\alpha}_1, \hat{\alpha}_{02}, \hat{\alpha}_{11}, \hat{\alpha}_{12}, \hat{\pi})$. When using standard statistical software packages to fit growth curve PMMs, $\hat{V}$ is typically available from the growth curve output; it is saved as
a list item called varFix in nlme, and in SAS Proc Mixed it is available using the statement, “ods output covb.” \(\hat{\Delta}\) is a \(k=q_1 + q_2 + 2\) by \(2(k+1)\) matrix of partial derivatives of \(\alpha_u\) with respect to the parameters \((\gamma_0, \alpha_{01}, \gamma_1, \alpha_{02}, \alpha_{11}, \alpha_{12}, \pi)\), evaluated at \((\hat{\gamma}_0, \hat{\alpha}_{01}, \hat{\gamma}_1, \hat{\alpha}_{02}, \hat{\alpha}_{11}, \hat{\alpha}_{12}, \hat{\pi})\). In our case, \(\hat{\Delta}' = [I_{k+1}, \hat{\pi} \times I_{k+1}, A]\), where \(A' = (\alpha_{02}', \alpha_{12}')\), a vector of regression parameters that reflect the difference between those in pattern \(R=1\) versus others.

REFERENCES


Table 1. Patterns of observed data on CS score

<table>
<thead>
<tr>
<th>Time (months)</th>
<th>0 (baseline)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>3</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>6</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>9</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Number of clients 92 86 74 71 131 13 12 10 10 6 4 4 3 2 1
Table 2. Growth Curve Analyses of Non-Completers (did not complete 9-month assessment) versus Completers.

Columns (a)-(c) summarize results from one model fit for completers (a), non-completers (b), and overall (c).

<table>
<thead>
<tr>
<th></th>
<th>(a) Completers of 9-month assessment</th>
<th>(b) Non-completers of 9-month assessment</th>
<th>(c) PMM averaged over patterns</th>
<th>(d) Random-effects longitudinal model (under MAR assumption)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (months)</td>
<td>Value 0.0375 Std.Error 0.0082 p-value 0.0000</td>
<td>Value 0.0479 Std.Error 0.0131 p-value 0.0003</td>
<td>Value 0.0445 Std.Error 0.0098 p-value 0.0000</td>
<td>Value 0.0405 Std.Error 0.0076 p-value 0.0000</td>
</tr>
<tr>
<td>Predictors of CS at baseline:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>3.8879 Std.Error 0.0781 p-value 0.0000</td>
<td>3.8304 Std.Error 0.0738 p-value 0.0000</td>
<td>3.8495 Std.Error 0.0693 p-value 0.0000</td>
<td>3.8619 Std.Error 0.0698 p-value 0.0000</td>
</tr>
<tr>
<td>Prior drug treatment (yes/no)</td>
<td>0.1697 Std.Error 0.1053 p-value 0.1075</td>
<td>0.0191 Std.Error 0.0848 p-value 0.8217</td>
<td>0.0690 Std.Error 0.0675 p-value 0.3071</td>
<td>0.0528 Std.Error 0.0659 p-value 0.4233</td>
</tr>
<tr>
<td>Age (years/100)</td>
<td>0.2818 Std.Error 0.5875 p-value 0.6316</td>
<td>-0.4944 Std.Error 0.4473 p-value 0.2694</td>
<td>-0.2372 Std.Error 0.3582 p-value 0.5082</td>
<td>-0.1325 Std.Error 0.3472 p-value 0.7029</td>
</tr>
<tr>
<td>History of sexual abuse (yes/no)</td>
<td>0.0039 Std.Error 0.1430 p-value 0.9781</td>
<td>0.1715 Std.Error 0.0974 p-value 0.7878</td>
<td>0.1159 Std.Error 0.0806 p-value 0.1506</td>
<td>0.1240 Std.Error 0.0799 p-value 0.1210</td>
</tr>
<tr>
<td>Child of a substance abuser (yes/no)</td>
<td>-0.1480 Std.Error 0.0808 p-value 0.0675</td>
<td>-0.0792 Std.Error 0.0659 p-value 0.2297</td>
<td>-0.1020 Std.Error 0.0520 p-value 0.0501</td>
<td>-0.0984 Std.Error 0.0511 p-value 0.0546</td>
</tr>
<tr>
<td>Female</td>
<td>-0.1091 Std.Error 0.1163 p-value 0.3488</td>
<td>0.1032 Std.Error 0.0892 p-value 0.2476</td>
<td>0.0328 Std.Error 0.0717 p-value 0.6469</td>
<td>0.0238 Std.Error 0.0708 p-value 0.7363</td>
</tr>
<tr>
<td>Age of first use (years/100)</td>
<td>0.9092 Std.Error 0.4832 p-value 0.0604</td>
<td>0.3458 Std.Error 0.4471 p-value 0.4396</td>
<td>0.5325 Std.Error 0.3441 p-value 0.1223</td>
<td>0.6712 Std.Error 0.3327 p-value 0.0440</td>
</tr>
<tr>
<td>Not a high school graduate</td>
<td>0.1355 Std.Error 0.0870 p-value 0.1199</td>
<td>0.0732 Std.Error 0.0707 p-value 0.3009</td>
<td>0.0939 Std.Error 0.0554 p-value 0.0905</td>
<td>0.0905 Std.Error 0.0543 p-value 0.0956</td>
</tr>
<tr>
<td>Race (White)</td>
<td>-0.3737 Std.Error 0.0880 p-value 0.0000</td>
<td>-0.2650 Std.Error 0.0683 p-value 0.0001</td>
<td>-0.3010 Std.Error 0.0546 p-value 0.0000</td>
<td>-0.3056 Std.Error 0.0538 p-value 0.0000</td>
</tr>
<tr>
<td>Referred to treatment by criminal justice+</td>
<td>0.1677 Std.Error 0.0872 p-value 0.0549</td>
<td>-0.1203 Std.Error 0.0715 p-value 0.0933</td>
<td>-0.0248 Std.Error 0.0580 p-value 0.6689</td>
<td>0.0021 Std.Error 0.0572 p-value 0.9702</td>
</tr>
<tr>
<td>Predictors of CS change:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior drug treatment (yes/no)*time</td>
<td>-0.0002 Std.Error 0.0141 p-value 0.9895</td>
<td>-0.0182 Std.Error 0.0244 p-value 0.4562</td>
<td>-0.0122 Std.Error 0.0171 p-value 0.4738</td>
<td>0.0024 Std.Error 0.0119 p-value 0.8428</td>
</tr>
<tr>
<td>(Age/100)<em>time</em></td>
<td>-0.0914 Std.Error 0.0773 p-value 0.2373</td>
<td>0.2621 Std.Error 0.1299 p-value 0.0438</td>
<td>0.1449 Std.Error 0.0909 p-value 0.1113</td>
<td>0.0021 Std.Error 0.0645 p-value 0.9744</td>
</tr>
</tbody>
</table>

Completers and non-completers differ: p<0.05 (^), p<0.01 (+)
Table 3. Growth Curve Analyses of 3-month or earlier ("Shorter stay" pattern) versus others. Columns (a)-(c) summarize results from one model fit for 6- and/or 9-month completers (a), non-completers (b), and overall (c).

<table>
<thead>
<tr>
<th>Predictor of CS at baseline:</th>
<th>(a) Completers of 6 and/or 9-month assessments</th>
<th>(b) Completers of Baseline, 1, and/or 3-month assessments only</th>
<th>(c) PMM averaged over patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (months)</td>
<td>Value 0.0360 Std.Error 0.0081 p-value 0.0000</td>
<td>Value 0.0588 Std.Error 0.0268 p-value 0.0287</td>
<td>Value 0.0475 Std.Error 0.0147 p-value 0.0013</td>
</tr>
<tr>
<td>Prior drug treatment (yes/no)</td>
<td>3.8979 0.0771 0.0000</td>
<td>3.7642 0.0842 0.0000</td>
<td>3.8304 0.0745 0.0000</td>
</tr>
<tr>
<td>Age (years/100)</td>
<td>0.0727 0.0854 0.3946</td>
<td>0.1208 0.1062 0.2554</td>
<td>0.0970 0.0693 0.1617</td>
</tr>
<tr>
<td>History of sexual abuse (yes/no)</td>
<td>0.3535 0.4684 0.5348</td>
<td>-0.7231 0.5546 0.1928</td>
<td>-0.1900 0.3659 0.6038</td>
</tr>
<tr>
<td>Prior drug treatment*time</td>
<td>0.0052 0.0127 0.6821</td>
<td>-0.0489 0.0519 0.3468</td>
<td>-0.0221 0.0271 0.4142</td>
</tr>
<tr>
<td>(Age/100)*time</td>
<td>-0.0362 0.0689 0.5990</td>
<td>0.0685 0.2743 0.8027</td>
<td>0.0167 0.1427 0.9071</td>
</tr>
</tbody>
</table>

^Completers of 6-9 month assessments differ from non-completers of those assessments (p<0.05)
Table 4. Growth Curve Analyses of 1 month or earlier ("Shorter stay" pattern) versus others. Columns (a)-(c) summarize results from one model fit for completers of the 3-month or later assessment (a), non-completers (b), and overall (c).

<table>
<thead>
<tr>
<th>Prediction</th>
<th>(a) Completers of 3, 6 and/or 9-month assessments</th>
<th>(b) Completers of Baseline and/or 1-month assessments</th>
<th>(c) PMM averaged over patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (months)</td>
<td>Value 0.0382, Std.Error 0.0078, p-value 0.0000</td>
<td>Value -0.0717, Std.Error 0.0810, p-value 0.3765</td>
<td>Value 0.0005, Std.Error 0.0285, p-value 0.9851</td>
</tr>
<tr>
<td>Predictors of CS at baseline:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>3.8944, 0.0762, 0.0000</td>
<td>3.7815, 0.0966, 0.0000</td>
<td>3.8557, 0.0758, 0.0000</td>
</tr>
<tr>
<td>Prior drug treatment (yes/no)</td>
<td>0.0788, 0.0748, 0.2928</td>
<td>-0.0651, 0.1536, 0.6720</td>
<td>0.0294, 0.0728, 0.6860</td>
</tr>
<tr>
<td>Age (years/100)</td>
<td>0.0388, 0.4011, 0.9229</td>
<td>-1.3214, 0.7932, 0.0962</td>
<td>-0.4277, 0.3813, 0.2625</td>
</tr>
<tr>
<td>History of sexual abuse (yes/no)</td>
<td>0.1029, 0.0973, 0.2907</td>
<td>0.1902, 0.1452, 0.1905</td>
<td>0.1329, 0.0811, 0.1018</td>
</tr>
<tr>
<td>Child of a substance abuser (yes/no)</td>
<td>-0.1494, 0.0580, 0.0102</td>
<td>0.0247, 0.1080, 0.8191</td>
<td>-0.0897, 0.0535, 0.0943</td>
</tr>
<tr>
<td>Female</td>
<td>0.0088, 0.0841, 0.9170</td>
<td>0.1005, 0.1377, 0.4660</td>
<td>0.0402, 0.0728, 0.5805</td>
</tr>
<tr>
<td>Age of first use (years/100)</td>
<td>0.7146, 0.3661, 0.0514</td>
<td>0.1246, 0.7795, 0.8730</td>
<td>0.5123, 0.3644, 0.1603</td>
</tr>
<tr>
<td>Not a high school graduate</td>
<td>0.0679, 0.0622, 0.2753</td>
<td>0.2223, 0.1134, 0.0505</td>
<td>0.1209, 0.0564, 0.0326</td>
</tr>
<tr>
<td>Race (White)</td>
<td>-0.3537, 0.0615, 0.0000</td>
<td>-0.1596, 0.1129, 0.1581</td>
<td>-0.2871, 0.0565, 0.0000</td>
</tr>
<tr>
<td>Referred to treatment by criminal justice</td>
<td>0.0093, 0.0651, 0.8867</td>
<td>-0.0899, 0.1119, 0.4216</td>
<td>-0.0248, 0.0596, 0.6783</td>
</tr>
<tr>
<td>Predictors of CS change::</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior drug treatment (yes/no)*time</td>
<td>0.0003, 0.0123, 0.9831</td>
<td>0.2831, 0.1890, 0.1344</td>
<td>0.0973, 0.0656, 0.1384</td>
</tr>
<tr>
<td>(Age/100)*time</td>
<td>-0.0233, 0.0663, 0.7251</td>
<td>1.5327, 0.9374, 0.1024</td>
<td>0.5104, 0.3260, 0.1178</td>
</tr>
</tbody>
</table>

Note: Completers of 3, 6, and/or 9-month assessments do not significantly differ from others.