An Introduction and Practical Guide to Strategies for Analyzing Longitudinal Data in Clinical Trials of Smoking Cessation Treatment: Beyond Dichotomous Point-Prevalence Outcomes

George Kypriotakis PhD[1], Steven L. Bernstein MD[2,3], [iD], Krysten W. Bold PhD[4], [iD],
James D. Dziura PhD[2,3,†], Donald Hedeker PhD[5,†], Robin J. Mermelstein PhD[6,†], [iD],
Andrea H. Weinberger PhD[7,8,9], [iD]

[1]Department of Behavioral Science, UT MD Anderson Cancer Center, Houston, TX, USA
[2]Department of Emergency Medicine, Yale University School of Medicine, New Haven, CT, USA
[3]Department of Public Health, Yale University School of Public Health, New Haven, CT, USA
[4]Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA
[5]Department of Public Health Sciences, University of Chicago, Chicago, IL, USA
[6]Department of Psychology and Institute for Health Research and Policy, University of Illinois Chicago, Chicago, IL, USA
[7]Ferkauf Graduate School of Psychology, Yeshiva University, Bronx, NY, USA
[8]Department of Psychiatry and Behavioral Sciences, Albert Einstein College of Medicine, Bronx, NY, USA
[9]Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA

Corresponding Author: Andrea H. Weinberger, PhD, Ferkauf Graduate School of Psychology, Yeshiva University, 1165 Morris Park Ave, Bronx, NY 10461, USA.
Telephone: 646-592-4474; Fax: 718-430-3960; E-mail: andrea.weinberger@yu.edu
†These authors made equal contributions to the paper and are listed in alphabetical order.

Abstract

Conceptualizing tobacco dependence as a chronic relapsing condition suggests the need to use analytic strategies that reflect that premise. However, clinical trials for smoking cessation typically define the primary endpoint as a measure of abstinence at a single timepoint distal to the intervention, typically 3–12 months. This reinforces the concept of tobacco outcomes as a dichotomous state—one is, or is not, abstinent. Fortunately, there are several approaches available to handle longitudinal data that reflect the relapsing and remitting nature of tobacco use during treatment studies. In this paper, sponsored by the Society for Research on Nicotine and Tobacco's Treatment Research Network, we present an introductory overview of these techniques and their application in smoking cessation clinical trials. Topics discussed include models to examine abstinence outcomes (eg, trajectory models of abstinence, models for transitions in smoking behavior, models for time to event), models that examine reductions in tobacco use, and models to examine joint outcomes (eg, examining changes in the use of more than one tobacco product). Finally, we discuss three additional relevant topics (ie, heterogeneity of effects, handling missing data, and power and sample size) and provide summary information about the type of model that can be used based on the type of data collected and the focus of the study. We encourage investigators to familiarize themselves with these techniques and use them in the analysis of data from clinical trials of smoking cessation treatment.

Implications: Clinical trials of tobacco dependence treatment typically measure abstinence 3–12 months after participant enrollment. However, because smoking is a chronic relapsing condition, these measures of intervention success may not accurately reflect the common trajectories of tobacco abstinence and relapse. Several analytical techniques facilitate this type of outcome modeling. This paper is meant to be an introduction to these concepts and techniques to the global nicotine and tobacco research community including which techniques can be used for different research questions with visual summaries of which types of models can be used for different types of data and research questions.

Introduction

Researchers and health care professionals have argued that tobacco dependence, like dependence on other drugs, should be classified and treated as a chronic disease.[1–5] The majority of adults who use commercial tobacco, such as those who smoke cigarettes, state they want to stop smoking and report making quit attempts, but very few of these individuals are successful at long-term abstinence.[6] Relapse rates after smoking cessation attempts are high, and even the most effec-

tive and efficacious pharmacological and behavioral smoking cessation treatments do not help the majority of people who use cigarettes and/or other tobacco products achieve long-term cessation success.[7,8]

This paper reviews methodologic and analytic considerations for clinical trials of smoking cessation treatment. Our premise is that, although tobacco dependence is increasingly framed as a chronic relapsing disease,[5] the primary outcome in many smoking cessation clinical trials is

the measurement of smoking abstinence (abstinent vs. not abstinent) and the standard empirical practice is to evaluate abstinence through self-report and/or biochemical confirmation at a single timepoint (usually 3- or 6-months following treatment initiation or end of treatment (EOT)). The implicit assumption is that abstinence measured at this single point adequately captures smoking status. However, there is no reason to assume that smoking status over timepoints averages out over the study period. For example, if transitions between abstinent and not abstinent status occur over the study period, then the timepoint at which abstinence is measured is important.

Further, a simple timepoint as a proxy of abstinence cannot capture changes over different timepoints and may ignore critical periods in the abstinence process.[9,10] The importance of timing in clinical trials is not limited to smoking abstinence, but also changes in other important variables implicated in quit success (eg, dependence, withdrawal).[10–13] Together these limitations suggest that the standard evaluation of abstinence status in clinical trials through the use of a single timepoint may miss dynamic changes in abstinence and important mechanisms at work in achieving and maintaining abstinence.[14] In this time of value-based care and quality outcomes one needs to show that a treatment is effective but also when (ie, at what timepoints and for how long) the treatment is likely to be most effective to have the best clinical outcomes and to reduce personal and societal negative consequences (eg, side effects, unnecessary cost).[3,15,16] That is, the ability to evaluate and predict the timing of change becomes important. There are numerous statistical methods to deal with longitudinal measurements in smoking cessation trials. Our purpose here is to introduce the most relevant ones appropriate to explore the richness of the time-dependent measurements. Through the systematic assessment of longitudinal data, we can enhance our understanding of smoking behaviors, facilitating the development of more robust theoretical frameworks that incorporate temporal dynamics as essential components.

Topics discussed below include models to examine abstinence outcomes (eg, trajectory models of abstinence, models for transitions in smoking behavior, models for time to event), models that examine reductions in tobacco use, and models to examine joint outcomes (ie, more than one outcome variable simultaneously, as is the case among those who concurrently use more than one tobacco product). Finally, we will discuss three additional relevant topics; heterogeneity of effects, handling missing data, and power and sample size; and provide summary information about the type of models that can be used based on the type of data collected and the focus of the study.

This paper is meant to be a useful introduction of the topic to the global nicotine and tobacco research community with an emphasis on researchers conducting smoking cessation treatment clinical trials in a range of settings (eg, smoking cessation clinics, medical settings, psychiatric settings). As such, we present information, including visual summaries, of what types of models can be used for different types of data and research questions. This paper does not present an exhaustive discussion of these techniques. Interested readers who would like to investigate specific topics in more depth are encouraged to do so by using the ideas presented here when consulting with research team members with more specialized knowledge of statistical techniques. An Appendix with a list of references related to each modeling approach is also provided for interested readers.

In this paper, "tobacco" refers to commercial, not ceremonial, tobacco. In addition, while we refer to clinical trials for smoking cessation treatment, these treatment studies may have a range of inclusion criteria (eg, tobacco dependence, current smoking, daily cigarette use) or labels for the type of study (eg, smoking cessation treatment, tobacco dependence treatment, tobacco use treatment) and we use "smoking cessation treatment" as an umbrella term for these types of studies for consistency.

Examining Abstinence Outcomes

Trajectory Models for Abstinence

In the trajectory models of abstinence, we shift the focus from measuring abstinence at a specific timepoint or points to modeling change in abstinence over time in the form of curves (Figure 1). In doing so, we approach time-specific observations of smoking status as snapshots of an underlying curve (see green and red curves in Figure 1) that help to
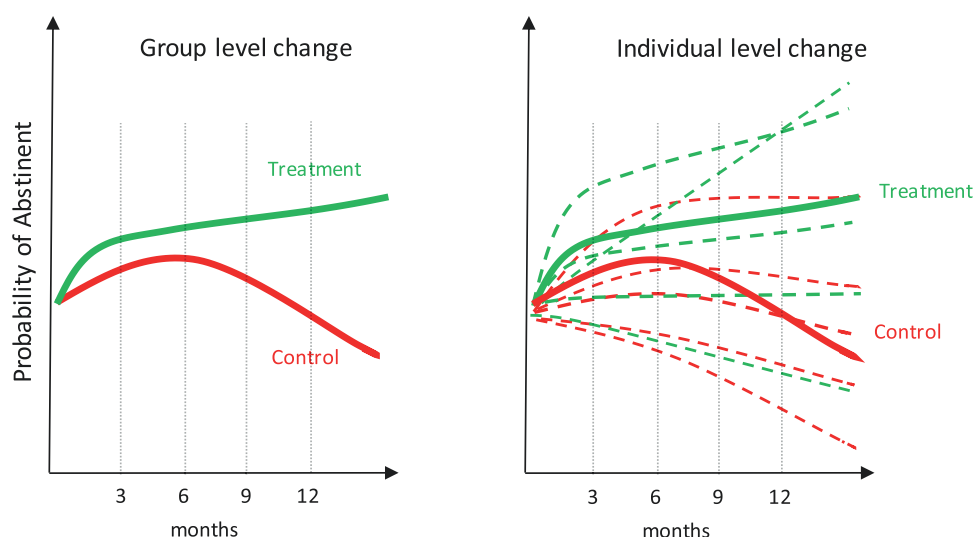

Figure 1. Hypothetical trajectories in the probability of abstinence in group level and individual level change.

uncover trends in the abstinence/non-abstinence continuum that may reveal important information on the timing of smoking cessation treatment (eg, early or delayed treatment effects), as well as the impact of changes of other smoking-related attributes (such as affect and withdrawal symptoms). In general, the trajectory model of abstinence captures a richer and more nuanced characterization of the abstinence process than assessing abstinence at specific points in time.

Questions of Clinical Importance
Questions that a clinician can ask when using trajectory models include: (a) When, on average, do we observe the most rapid change in the probability of abstinence or relapse? (b) How big is the change in the probability of abstinence or relapse? (c) How do abstinence trajectories differ between treatment groups? and (d) How do abstinence trajectories differ for individuals within treatment groups?

Analytic Technique
A general analytic approach for modeling group, as well as individual-specific, abstinence trajectories is growth curve models (GCM)[17–21] (other names for these models are multilevel models, hierarchical linear models, or mixed effects models). GCM has several desirable features that characterize cessation trials. They can: (a) handle situations where abstinence is measured at different timepoints, (b) incorporate both time-varying and time-invariant treatments eg, adjust to treatment dose, augmentation, or switching pharmacotherapies during treatment) as well as both time-varying (eg, changes in depression during treatment) and static (eg, level of tobacco dependence at baseline) covariates, (c) produce unbiased estimates as long as missingness is missing-at-random (MAR; see the section below on missing data) (note, however, that this is only valid in full maximum likelihood estimation procedures), (d) provide an error structure that can be modeled instead of making arbitrary assumptions (eg, assuming that within-group variation is homogeneous across treatment groups and across time), and (e) produce individual-specific abstinence trajectories as a function of cessation treatment.

The model parameters of GCM allow for the estimation of a variety of quantities of interest that can be compared across treatment groups, including abstinence rates at any point in time or averaged over time (eg, at treatment initiation, EOT, or any follow-up point, or across specific time-intervals), the shape of the abstinence curve (linear, quadratic, step, nonlinear), and rates of change at time intervals of interest captured by piecewise growth models (eg, from treatment initiation to treatment augmentation). The initial level and the rate of change (also known as intercept and slope) are the most used metrics that characterize trajectories. In binary endpoints, such as abstinence status, the initial level and rate of change are most intuitive when interpreted as probabilities of abstinence. Another way to summarize the treatment's overall impact is by capturing the cumulative effects across the treatment timeline by integrating the entire trajectory in the form of an area under the curve. This approach is illustrated in Figure S1. This method offers a comprehensive summary by considering the total impact over time, with the flexibility to prioritize key moments in the treatment timeline (by applying weights to specific timepoints).

Models for Transitions
Models of transition allow researchers to model multiple sequential transitions between abstinence and smoking. Unlike trajectory models which often use a continuous dependent variable eg, smoking rate), transition models often use a dichotomous on/off state-based variable.

Questions of Clinical Importance
In modeling abstinence/non-abstinence transitions, the salient research questions change from "Who is most likely to be abstinent at a specific timepoint?" to questions such as: (1) Who is most likely to remain abstinent and who is most likely to relapse? (2) How do smoking cessation interventions affect the probabilities of switching between smoking states? and (3) How do past abstinence states affect future abstinence states and how does this relationship vary as a function of treatment?

Illustrative Example
Figure 2 depicts hypothetical scenarios of abstinence dynamics. The y-axis is divided by abstinence status which separates the abstinent from the non-abstinent. The horizontal axis represents the study time with hypothetical trajectories of several study participants. Person B, for instance, was abstinent for most of the period observed but experienced a short transitory non-abstinence period, so intensified interventions during this time may be more effective in reversing the trend during the specific time interval. In contrast, Person C experienced a transition from abstinence to non-abstinence
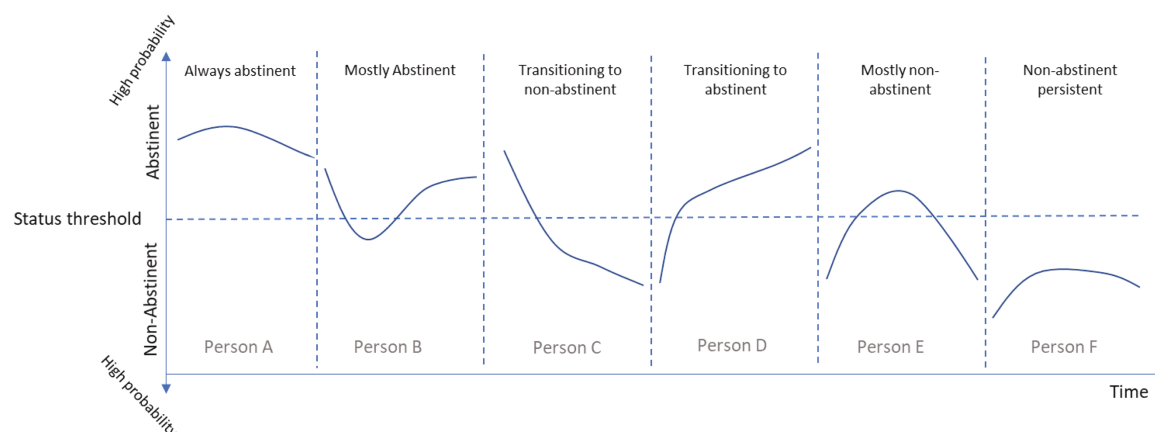

Figure 2. Trajectories and transitions of six hypothetical study participants.

and remained in a state of non-abstinence for the rest of the study. This graphical representation illustrates the probability of being in a specific state (abstinent or not abstinent) over time, despite the outcome itself being binary. This approach captures the dynamic process of transition, reflecting the evolving likelihood of an individual's smoking status at different points. Here, intervening with higher intensity near the tipping timepoints may be needed to delay or completely prevent a return into the not abstinent state.

Analytic Technique

The static analysis assumes that the abstinent/not abstinent categories are fixed groups with members of the groups displaying persistent characteristics that do not change over time (ie, based on a measurement at one timepoint, a person is classified as either "abstinent" or "not abstinent"). However, people within each of the two groups may display a range of potential smoking status patterns that can be captured over time. The assumption of persistence of smoking states is not merely inadequate for causal analysis, it can also have distorting effects when evaluating the efficacy of smoking cessation interventions.[22] Single point analysis misses important information about when people initiate abstinence, for how long they are abstinent, and when they lapse or relapse back to smoking.[23] Further, the trajectories of participants in clinical trials can exhibit abstinent and non-abstinent periods that may reveal processes central to smoking cessation and reduction. In terms of actionable knowledge, information on trajectories of abstinence and transition patterns between abstinence states bring much added value to clinicians compared to information from studies using single fixed timepoint analyses, and accommodate modeling of harm reduction, insofar as the endpoint modeled does not remain dichotomous (ie, either abstinent or not abstinent).

One approach for the analysis of abstinence dynamics is based on Markov models[22] that estimate the probability of transition (ie, the probability of someone who smokes transitioning to abstinence, and the probability of someone who abstains transitioning to relapse) as a function of an intervention as well as other characteristics eg, tobacco dependence, effect). Technically, a Markov chain is a discrete dynamical system of transition probabilities (conditional on previous states and other participant characteristics) from one smoking status to the other.[24] Markov chains can be of first-order (ie, the future smoking state depends only on the present smoking state and is independent of past states) or higher-order (ie, the future smoking state depends on the present smoking state as well as on past states).

Models of transition between smoking states shift the focus of interventions from those who currently smoke toward identifying routes to quitting and preventing relapse. These models explicitly address the underlying processes which lead people to abstain and to relapse, critical processes that can inform treatment to improve individual smoking outcomes.

Model for Time to Event
Questions of Clinical Importance

In research settings where the primary endpoint is the point when a change in abstinence status occurs, the time until that change (often called survival time, failure time, or event time) is the outcome of interest. Important questions in this setting include: (1) What is the probability that a study participant will stay abstinent for 6 months (or any other length of time)? (2) Are there differences in time to relapse between the treatment and the control groups? and (3) Does the risk of relapse change as the number of previous relapse events increases?

Illustrative Example

Figure S2 presents a hypothetical example of the abstinence experience of two groups showing the percentage remaining abstinent at any point in time throughout the study. For any point on the curves, we can estimate the percentage of those who have not relapsed up to this point for any group. Statistical comparisons (that can be made for the whole curve or any point on the curve) can evaluate whether these different percentages are statistically significant.

Analytic Technique

Survival analysis is one strategy to consider when the interest is in examining the time until an event occurs (eg, abstinence or relapse).[25,26] For example, survival analysis can be used when the research question is to identify the treatment group that is more likely to relapse soonest, or the treatment group that remains abstinent the longest before relapsing. There are three main analytic elements in survival analysis[19]: (1) a well-defined dichotomous event (eg, from "not abstinent" to "abstinent"), (2) a clearly defined beginning of time, and (3) a meaningful metric for measuring time. Like models of transitions discussed above, an event in survival analysis is understood as a transition from one state to another (eg, from not abstinent to abstinent) and the timing of this transition is the key point of interest. A clearly defined beginning of time refers to a timepoint where all study participants can experience the event (ie, abstinence) but have not yet done so, such as a prespecified quit date or start date of the intervention. Finally, measurement of time should be done in the smallest possible unit that is relevant to the study. For example, weekly measurements may be the most meaningful time metric for pharmacotherapy smoking cessation studies as it is common to report outcomes such as 7-day-point-prevalence. One can also perform survival analysis on typical longitudinal data (such as from monthly or yearly questionnaires).[27–29] In general, survival analysis should be used when the research question asks whether a predefined smoking event occurs and when it occurs.

Consider a hypothetical example of a clinical study comparing the abstinence rates between two treatment groups who have been followed for 12 months. After 12 months a certain proportion of each group will be abstinent, while the rest are not abstinent (ie, still smoking). What is the most efficient way to evaluate the effects of the treatment on abstinence? Usually, what is reported is the percentage of each group who is abstinent, ignoring important additional information such as whether the treatment group reached abstinence sooner than the control group. Single timepoint analysis also does not consider the arbitrariness of the timing of the endpoint. There are almost always study participants who are not abstinent by the endpoint, but they may eventually be abstinent if followed long enough. This lack of complete information on the event of interest due to the arbitrary time limitations of a study is called censoring.[25] Censoring is the most important characteristic that distinguishes the analysis of time-to-event data from other methods for time-dependent data. For censored study participants, we do not have information on when the event will occur, if ever. The advantage

of survival analysis in the presence of censored data is that it produces unbiased estimates at each unique event time by suitably adjusting the number of participants at risk (ie, the participants who have not experienced the event (eg, abstinence) and are not censored).[26] The most commonly used method to model time to event data is the Cox model,[30] also known as the relative risk or relative hazard model. The effect of cessation treatment on the time to event denotes the difference in hazard between treatment groups of experiencing the event of interest (eg, abstinence) at any fixed timepoint during the observation period. The basic Cox model can be extended to facilitate time-dependent covariates, repeated events, and competing risks.[19,31]

Examining Reduction Outcomes

Models for Reduction in Use

Alongside the collection of abstinence data in clinical trials, longitudinal information on reduction in the use of tobacco products is frequently obtained as an additional endpoint. In these endpoints, changes in cigarettes per day (CPD) are often the outcome of interest and other outcomes may include the number of relapses to smoking or the number of days abstinent from cigarettes.

Questions of Clinical Importance

One clinical question of interest for these data would be: (1) Is the smoking or abstinence outcome of interest (eg, CPD, number of relapses) observed more frequently in one of the treatment groups, and if observed, at what level of intensity? and (2) Does there exist a subpopulation of people who smoke that will never experience a relapse during treatment (eg, will always have zero CPD throughout the study)? An investigator can also evaluate the effects of treatment in more detail by examining both processes simultaneously. Hence, in longitudinal measurements of count-related outcomes (eg, number of CPD, number of relapses, number of days abstinent) one can assume two populations, one characterized by never displaying the behavior under study and one characterized by always or sometimes displaying the behavior.

Illustrative Example

As an illustration, Figure S3 shows a hypothetical total number of CPD counts for a treatment and a control group. In this hypothetical scenario, the time by CPD displayed is characterized by heterogeneity in treatment group trajectories. The red line representing the control group is an example of slightly rising CPD counts early in the study with a slight reduction later in the study. In contrast, the intervention group (green line) presents a dramatic initial decrease in CPD counts followed by an increase thereafter.

Analytic Technique

Here, to model these changes in reduction, similar to the earlier section of trajectories of abstinence, growth curve modeling can be used to estimate differences in initial level and rate of change between treatment groups[32] using the Poisson or negative binomial model. Smoking reduction outcomes (eg, CPD, number of relapses, number of days abstinent) may frequently have a large number of data points at zero indicating that the behavior represented by the outcome was not realized (eg, zero number of relapses). This is seen in Figure S3, which shows differences in counts of CPD across time for the control and the intervention groups as well as differences in the number of zeros (ie, the number who did not smoke at all). Two different statistical models can be used to model the zero vs. any CPD, and the total number of CPD: (1) a logistic model for the zero versus any CPD and (2) a count regression model for non-zero CPD. Models for these types of outcomes are generally known as Two-Part models[33,34] and are usually estimated using logistic regression for the binary part (zero vs. any CPD) and Poisson or negative binomial regression for the continuous part (non-zero CPD).

In addition, by simultaneously examining both processes, the likelihood of engaging in a behavior and, the frequency of the behavior for those who do engage, these models can provide a more nuanced and detailed evaluation of the intervention effect.[35,36]

Examining Multiple Outcomes Jointly

Simultaneous modeling of longitudinal and time-to-event data can become a valuable tool in the analysis of data in smoking cessation studies. This is especially important when certain processes, such as craving or withdrawal, are time-dependent and co-evolve with abstinence or any other event of interest. Moreover, there may be interest in the association of longitudinal measurements of two or more different tobacco products (eg, cigarettes and e-cigarettes). Joint models can accommodate both multiple categorical or continuous longitudinal measurements and the examine interdependence of poly-tobacco use.

Questions of Clinical Importance

Questions that a clinician can ask when using joint modeling include: (1) What is the interdependence of the trajectory of craving or withdrawal with abstinence status for different treatment groups? (2) What is the time-dependent association of use among different tobacco products? and (3) Are these types of time-dependent associations different between or among treatment groups?

Illustrative Example

Figure S4 provides an illustration of a joint model for abstinence and craving. The red and green line in the top panel shows how the probability of abstinence changes in time for the control and treatment groups respectively. In the bottom panel, the lines represent the underlying longitudinal trajectory of craving for the control and treatment groups. This joint model in its simple form assumes that the probability of abstinence at any point in time (denoted by the dotted black line) is directly associated with craving at the same timepoint, but only for the control group. For the treatment group, the evolution in abstinence is not correlated with the evolution in craving.

Figure S5 presents an illustration of two longitudinal processes representing the dual use of two products (cigarette and e-cigarette use) and hypothetical differences in the relationship between the two products. This example illustrates that at any point in time, the association of use between two (or more) tobacco products can be derived and the strength of this association can be compared between different treatment groups. In this hypothetical example, the data from the control group suggests switching from cigarettes to e-cigarettes,

followed by a return to greater cigarette use and lesser use of e-cigarettes. In this case, and assuming the use of nicotine-containing e-cigarettes, the total daily consumption of nicotine may remain constant. In the treatment group, in contrast, there is an overall decrease in the use of both products.

Analytic Technique

Smoking cessation trials often collect data on daily cigarette consumption and other smoking behaviors in a longitudinal fashion. Behavioral data such as craving, dependence, affect, depression, or withdrawal, can be important predictors of abstinence or relapse.[37,38] This dependency of the longitudinal outcomes and abstinence status is not considered when analyzing each outcome separately. Models that consider the dependency of the longitudinal measurements and abstinence status simultaneously, and assess the degree of dependence between them, can provide a more accurate evaluation of treatment efficacy. Knowledge of treatment response of a trajectory of withdrawal or dependence can be an indication of a clinical benefit and provide prognostic information on future abstinence status. Moreover, to the degree that the treatment effect is mediated via a longitudinal surrogate, joint modeling can clarify the mechanism of treatment response of clinical outcomes of interest. Most importantly, when examining outcomes separately, the presence of random error in the outcome of interest biases the estimates of the effect of the treatment towards the null and produces less precise treatment effects.[39] That translates into potentially characterizing interventions as non-efficacious, when they truly are.

A joint model can be constructed by bringing together two or more longitudinal outcomes and performing an analysis accounting for their dependence and estimating parameters that evaluate their relationship. That includes: (1) the estimation of the true trajectory of a behavior such as craving and (2) the estimation of a survival model for abstinence. The trajectory and survival models are linked through a shared parameter, and treatment effects can be evaluated separately for each outcome accounting for their dependence. Extensions of this basic joint model include estimating the effect of the slope of a trajectory of craving before a specific timepoint on the probability of abstinence at that timepoint, cumulative effects of craving on abstinence at future timepoints (usually represented as area under the trajectory before the timepoint of interest), or lagged effects. In addition, the survival analysis part of the joint model can examine abstinence as a single event, recurrent event, or as a competing event to other behaviors such as e-cigarette use. In clinical studies, the direct effects of cessation treatment on craving and abstinence can be estimated, controlling for the association between the two, as well as the indirect effect of treatment on abstinence via craving (which can be defined as the overall treatment effect). Moreover, in the process of discovery of mechanisms on how smoking cessation treatment cause changes in clinical outcomes, an investigator can examine the effects of treatment on the association of two time-dependent processes (eg, the relationship between craving and abstinence), as well as the change of this association as a function of treatment. Although we have discussed in more detail joint models of a longitudinal outcome and an event of interest, the extension to handling other longitudinal outcomes (eg, multiple tobacco products) is straightforward with the same implementation characteristics as the ones discussed.[40]

Additional Considerations

Heterogeneity of Effects

In randomized controlled trials for smoking cessation, we can estimate the average causal effect of the treatment by experimentally controlling for confounding through randomization. When clinicians translate that effect to individuals, they assume that the treatment effect of the specific study is constant across individuals. This assumption has the advantage that the effect identified in the study is relevant to each study participant. To verify this assumption, we suggest researchers test for heterogeneity of treatment effects.

Questions of Clinical Importance

Relevant clinical questions relative to heterogeneity of effects include: (1) Is the effect identified in the trial constant across individuals? (2) If not, how much do individual effects deviate from the average effect identified? and (3) If heterogeneity of effect is identified, what are predictors of differential response?

Illustrative Example

Examination of heterogeneity of effects provides evidence for whether smoking cessation treatments should be refined at the group or individual level by searching for predictors of differential response to smoking cessation treatment.

We illustrate three different types of treatment effects in clinical trials in Figure S6. In Scenario A, the treatment effect is constant implying that smoking cessation treatment affects all participants similarly. Scenarios B and C, however, indicate heterogeneity by treatment group (Scenario B) and by individual (Scenario C) and would need personalized approaches to treatment. Scenario B should be followed with analyses to identify baseline predictors of group level variation in treatment effects. Scenario C is more complicated, however, and it does not predict well the effects of treatment on future patients so there is the need to further personalize the clinical interventions. Finally, examination of unobserved heterogeneity can demonstrate treatment efficacy in the presence of placebo effects (a response observed in the placebo group that is not related to treatment) that can confound the true effect of the treatment.[41]

Analytic Technique

Testing this assumption by examination of the variability of the effects of smoking cessation treatments is important because it conveys clinically relevant information about whether to adapt treatments based on patients' profiles. Variability of treatment effects can be attributed to either patient variability (addressed by subgroup analyses, eg, by race or gender) or differential treatment effects that require more personalized smoking cessation treatments. In longitudinal analysis, variability of treatment effects can be explored with mixture modeling approaches that can identify observed or unobserved heterogeneity of temporal development in smoking outcomes: growth mixture modeling,[42] latent class growth analysis,[43] latent transition analysis also known as hidden Markov models,[44,45] and survival mixture analysis.[45,46] These methods accommodate analyses with outcomes that are binary, continuous, counts, ordered categorical, and censored. Moreover, these methods can be applied to test for the presence of both observed heterogeneity (eg, due to demographic groups) in

treatment effects, but also for unobserved heterogeneity of treatment effects.

Missing Data and Dropout

Missing data in longitudinal studies on tobacco use are often caused by dropout, where participants' responses are lost to follow-up. Analysis of such data is complicated by the fact that the responder's probability of dropout may relate to smoking status or other longitudinal outcomes that would have been observed if the participant had remained in the study.[47] Below we present the taxonomy of missing data introduced by Rubin and Little[48] and discuss three ways to deal with missing data.

When dropout can be predicted from the observed data (ie, dropout at time t can be predicted by the measurement of the outcome at time $t - 1$, especially in data with measurements of close proximity) and/or if the dropout process is assumed to be independent of the outcome process given observed information, a MAR assumption (MAR), then missingness can be ignored[49] when using full maximum likelihood estimation.[50,51] When data are MAR, maximum likelihood or multiple imputation can provide unbiased estimates, and the effects can be estimated as if no one has dropped out.[52] In rare cases (such as lost questionnaires or participants skipping survey pages accidentally), missing data can be assumed to be independent of both observed and unobserved information, missing completely at random, and point estimates are not likely to be affected by missingness. However, if dropout is caused by specific values in the outcome (eg, those who are non-abstinent or those with higher withdrawal symptoms are more likely to drop out) then the missingness is not-missing at random and should be accounted for. The researcher must explicitly specify a plausible model for the mechanism of missingness and integrate it into the analysis.

The most common modeling approaches for handling dropout can be generally classified either as pattern-mixture modeling (PMM),[53] selection modeling (SM),[54] or shared-parameter model (SPM).[39,55] In PMM, the sample is stratified by dropout time, and the outcome is modeled separately for each stratum, and potentially averaged across patterns.[56,57] This method is more intuitive to researchers and clinicians because it separates clinical outcomes between those who drop out from the study early and those who remain in the study.[58] In contrast, SM describes how the probability of dropout in a smoking cessation clinical trial may depend on the smoking status (or any other outcome) of the individuals. Hence, in SM missingness is predicted by the observed data (current and previous responses), usually by a logistic model, and these predictions are integrated into the analyses. Finally, in SPM, the outcome and missingness are assumed to be described by latent variables (such as random effects) that when estimated explain the dependence between the outcome and the dropout process.

Pattern-mixture models may be more suitable for situations where it is not meaningful to consider non-response as missing data (eg, missing data due to death), and it may be preferable to estimate the trajectories of subgroups defined by their dropout patterns. Selection models may be preferred when we have a good understanding the mechanisms of dropout due to their intuitive appeal (ie, using current earlier measurements of the outcome to predict missingness). Although SPM has not received a lot of attention, it is an efficient method for handling dropout, especially when other types of missing data (such as intermittent) are present in the data.[59]

Power and Sample Size

The utilization of longitudinal models in smoking cessation research can provide a substantial increase in statistical power, primarily due to their capacity to capture within-subject variability across multiple timepoints.[60–62] Unlike single timepoint analyses, which offer a static glimpse into the cessation process, longitudinal models can discern and quantify subtle changes and trends that might be obscured in a time-fixed framework. The characteristic of longitudinal models is further magnified by their ability to account for the correlated structure of repeated observations, which reduces random error. This increased efficiency translates to an improved sensitivity in detecting treatment effects, even when these effects are modest or evolve over time. Consequently, studies can achieve statistically significant results with fewer participants and more timepoints,[63,64] thus mitigating costs without compromising the rigor or reliability of the research.

Determining the optimal number of timepoints for longitudinal analysis can vary based on the specific goals and contexts of the study. In GCM, for instance, even with only three timepoints, we can estimate an initial trajectory of change. However, to capture more complex growth patterns such as non-linear trends, or to understand the variability in individual growth trajectories, additional timepoints are usually necessary. Similarly, in transition models, having more than two timepoints allows for a more nuanced understanding of the transition dynamics. With only two timepoints, these models can identify a change in state, but they are limited in their ability to characterize the process of transition or the factors influencing it. Survival analysis can be effective with a minimal number of timepoints, as it often focuses on the occurrence and timing of a single event. However, more timepoints allow for a more detailed exploration of the hazard function over time, particularly in cases where the risk of the event may change at different stages of the follow-up period. While these time-dynamic models can be applied with a limited number of assessments, the depth of the analysis is proportional to the number of timepoints. For instance, with only two follow-ups, the models might not fully exploit their capacity to describe intricate temporal patterns or detailed transitions over time. We do not think that there is a strict minimum number of timepoints for these models, but the complexity and detail of the insights obtained can be significantly enhanced with more follow-up assessments. The choice of the number of timepoints should be considered in the context of the specific objectives of the study. In smoking cessation studies, where individual trajectories can vary widely, the adoption of longitudinal models represents a methodologically sound, but also a cost-effective strategy to enhance the power of clinical trials.[65]

Discussion

Throughout this paper, we have presented an overview of the most relevant longitudinal models that can be used in the evaluation of smoking cessation treatments including how these methods reflect the type of questions researchers ask in the context of treatment trials. Figure 3 provides a schema summarizing the different types of data, study focus,
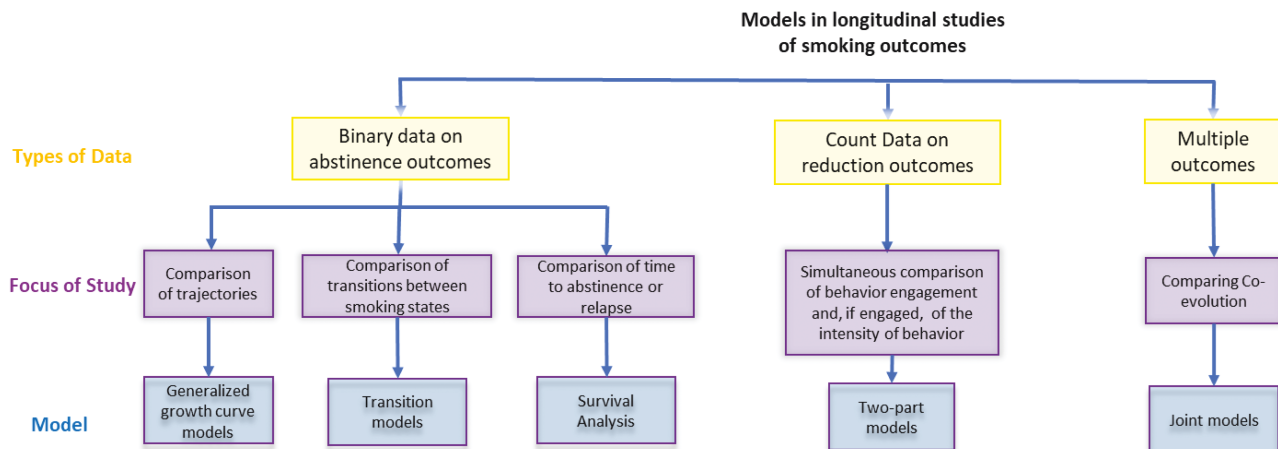
Figure 3. Schema of the different types of data, study focus, and models discussed.
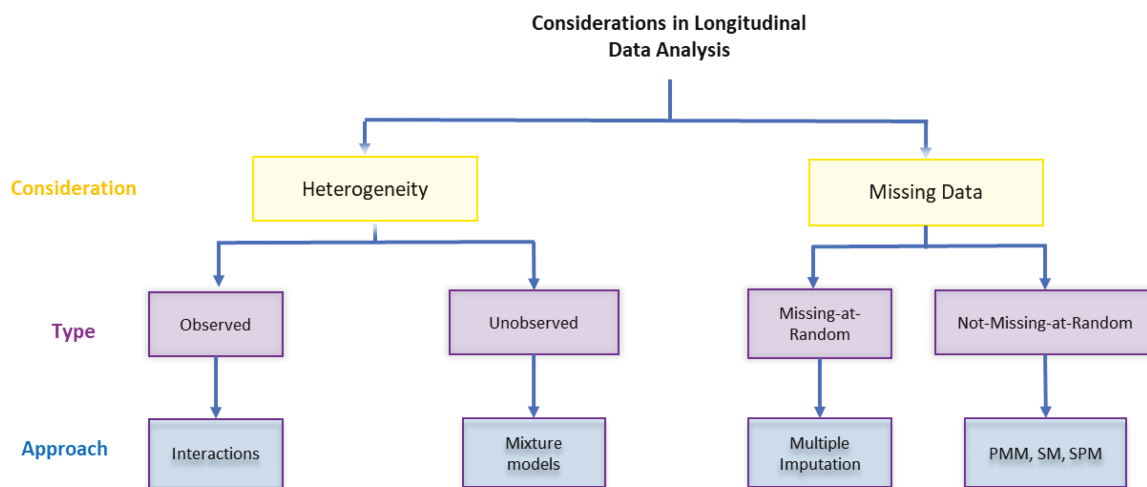
Figure 4. Schema of considerations in longitudinal models.

and models discussed while Figure 4 provides a schema of considerations to aid researchers in the potential options, choices, and types of questions that can arise in the process of analyzing longitudinal data from clinical trials of smoking cessation. Several widely used software packages have default routines that can estimate the models discussed including R, Stata, SAS, SPSS, and Mplus.

In addition to what is discussed here, there are additional approaches to analyzing longitudinal data that allow investigators to expand on the type of temporal questions they can ask. These include but are not limited to generalized forms of modeling, including generalized estimating equations,[66] dynamic structural equation model,[67] and time-varying effects models.[68,69] Many of these techniques can handle missing data, without the need for additional imputation when data are MAR. If imputation is desired, modern methods of multiple imputation allow investigators to avoid the traditional approach of imputing smoking at specified endpoints.[70] In addition, new techniques like time-varying effect modeling[68] and time-series,[71] allow investigators to analyze the kinds of intensive longitudinal data obtained from ecological momentary assessment (or EMA) assessments. We encourage investigators to familiarize themselves with these techniques and use them in the analysis of clinical trials of treatment. They are methodologically

rigorous, clinically sensible, and respect the framing of tobacco dependence as a chronic relapsing disease. Note, however, that while our manuscript emphasizes the benefits of analyzing the entire trajectory of treatment effects, we acknowledge the continued importance of individual timepoint assessments. These single timepoint assessments offer vital information about the immediate effectiveness of treatments, potential short-term benefits, or time-specific effects.

Conclusions

Modern computer-intensive statistical methods permit investigators to model longitudinal trajectories of tobacco use and quitting as the complex, relapsing/remitting behaviors they are. Longitudinal models therefore provide more realistic representations of the process of tobacco cessation, albeit at the cost of increased statistical complexity. The final choice of a model will depend on the research questions, the data at hand, the data structure, and the nature of the outcomes.

Supplementary material

Supplementary material is available at *Nicotine and Tobacco Research* online.

Declaration of Interest

None declared.

Author contributions

George Kypriotakis (Conceptualization [lead], Writing—original draft [lead], Writing—review & editing [lead]), Steven Bernstein (Conceptualization [lead], Writing—original draft [equal], Writing—review & editing [equal]), Krysten Bold (Conceptualization [equal], Project administration [equal], Writing—review & editing [equal]), James Dziura (Conceptualization [equal], Writing—review & editing [equal]), Donald Hedeker (Conceptualization [equal], Writing—review & editing [equal]), Robin Mermelstein (Conceptualization [equal], Writing—review & editing [equal]), and Andrea Weinberger (Conceptualization [lead], Project administration [lead], Writing—original draft [equal], Writing—review & editing [equal])

Previous Presentations

Work related to this paper was presented at the 2022 meeting of the Society for Research on Nicotine and Tobacco, Baltimore, Maryland.

References

1. Palmer AM, Toll BA, Carpenter MJ, *et al*. Reappraising choice in addiction: novel conceptualizations and treatments for tobacco use disorder. *Nicotine Tob Res*. 2022;24(1):3–9.
2. McLellan AT, Lewis DC, O'Brien CP, Kleber HD. Drug dependence, a chronic medical illness: implications for treatment, insurance and outcomes evaluation. *JAMA*. 2000;284(13):1689–1695.
3. Joseph AM, Fu SS, Lindgren B, *et al*. Chronic disease management for tobacco dependence: a randomized, controlled trial. *Arch Intern Med*. 2011;171(21):1894–1900.
4. Leone FT, Evers-Casey S. Developing a rational approach to tobacco use treatment in pulmonary practice: a review of the biological basis of nicotine addiction. *Clin Pulm Med*. 2012;19(2):53–61.
5. Steinberg MB, Schmelzer AC, Richardson DL, Foulds J. The case for treating tobacco dependence as a chronic disease. *Ann Intern Med*. 2008;148(7):554–556.
6. Creamer MR, Wang TW, Babb S, *et al*. Tobacco product use and Cessation indicators among Adults—United States, 2018. *MMWR Morb Mortal Wkly Rep*. 2019;68(45):1013–1019.
7. Fiore MC, Jaén CR, Baker TB, *et al. Treating Tobacco Use and Dependence: 2008 Update*. Rockvill, MD: U.S. Department of Health and Human Services. Public Health Service; 2008.
8. Piasecki TM. Relapse to smoking. *Clin Psychol Rev*. 2006;26(2):196–215.
9. Velicer WF, Prochaska JO, Rossi JS, Snow MG. Assessing outcome in smoking cessation studies. *Psychol Bull*. 1992;111(1):23–41.
10. Carroll KM. Methodological issues and problems in the assessment of substance use. *Psychol Assess*. 1995;7(3):349–358.
11. Patten CA, Martin JE. Does nicotine withdrawal affect smoking cessation? Clinical and theoretical issues. *Ann Behav Med*. 1996;18(3):190–200.
12. Robinson JD, Kypriotakis G, Al'absi M, *et al*. Very low nicotine content cigarettes disrupt the feedback loop of affective states and smoking behavior. *Nicotine Tob Res*. 2020;22(8):1294–1300.
13. Robinson JD, Kypriotakis G, Karam-Hage M, *et al*. Cigarette nicotine content as a moderator of the relationship between negative affect and smoking. *Nicotine Tob Res*. 2017;19(9):1080–1086.
14. Piasecki TM, Fiore MC, McCarthy DE, Baker TB. Have we lost our way? The need for dynamic formulations of smoking relapse proneness. *Addiction*. 2002;97(9):1093–1108.
15. Hughes JR, Cummings KM, Hyland A. Ability of smokers to reduce their smoking and its association with future smoking cessation. *Addiction*. 1999;94(1):109–114.
16. Baker TB, Mermelstein R, Collins LM, *et al*. New methods for tobacco dependence treatment research. *Ann Behav Med*. 2011;41(2):192–207.
17. Bollen K, Curran JP. *Latent curve models: A structural equation perspective*. New Jersey: John Wiley & Sons; 2006.
18. Raudenbush SW, Bryk AS. *Hierarchical Linear Models: Application and Data Analysis Methods. 2nd ed*. Thousand Oaks, CA: Sage Publication, Inc; 2002.
19. Singer JD, Willett JB. *Applied Longitudinal Data Analysis*. London, U.K.: Oxford University Press; 2003.
20. Hedeker D, Mermelstein RJ. Application of random-effects regression models in relapse research. *Addiction*. 1996;91(Suppl):S211–S229.
21. Hedeker D, Mermelstein RJ. Analysis of longitudinal substance use outcomes using ordinal random-effects regression models. *Addiction*. 2000;95(Suppl 3):S381–S394.
22. Killeen PR. Markov model of smoking cessation. *Proc Natl Acad Sci USA*. 2011;108(Suppl 3):15549–15556.
23. Mantey DS, Harrell MB, Chen B, *et al*. A longitudinal examination of behavioral transitions among young adult menthol and non-menthol cigarette smokers using a three-state Markov model. *Nicotine Tob Res*. 2021;23(6):1047–1054.
24. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Med Decis Making*. 1993;13(4):322–338.
25. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part I: basic concepts and first analyses. *Br J Cancer*. 2003;89(2):232–238.
26. Singer JD, Willett JB. Modeling the days of our lives: Using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. *Psychol Bull*. 1991;110(2):268–290.
27. Hedeker D, Siddiqui O, Hu FB. Random-effects regression analysis of correlated grouped-time survival data. *Stat Methods Med Res*. 2000;9(2):161–179.
28. Hox JJ, Robert JK, eds. *Handbook of Advanced Multilevel Analysis*. New York: Routledge; 2011.

29. Hedeker D, Mermelstein RJ. Multilevel analysis of ordinal outcomes related to survival data. In: Hox JJ RJ, eds., *Handbook of Advanced Multilevel Analysis*. New York: Routledge; 2011:115–136.

30. Su PF, Lin CK, Hung JY, Lee JS. The proper use and reporting of survival analysis and cox regression. *World Neurosurg*. 2022;161:303–309. doi:10.1016/j.wneu.2021.06.132.

31. Lin H, Zelterman, D. Modeling survival data: extending the cox model. *Technometrics*. 2002;44(1):85–86.

32. Xie H, Tao J, McHugo GJ, Drake RE. Comparing statistical methods for analyzing skewed longitudinal count data with many zeros: An example of smoking cessation. *J Subst Abuse Treat*. 2013;45(1):99–108.

33. Smith VA, Neelon B, Maciejewski ML, Preisser JS. Two parts are better than one: modeling marginal means of semicontinuous data. *Health Serv Outcomes Res Methodol*. 2017;17(3):198–218.

34. Farewell VT, Long DL, Tom BDM, Yiu S, Su L. Two-part and related regression models for longitudinal data. *Annu Rev Stat Appl*. 2017;4:283–315. doi:10.1146/annurev-statistics-060116-054131.

35. Liu H. Growth curve models for zero-inflated count data: An application to smoking behavior. *Struct Equ Modeling*. 2007;14(2):247–279.

36. Pittman B, Buta E, Krishnan-Sarin S, *et al*. Models for analyzing zero-inflated and overdispersed count data: an application to cigarette and marijuana use. *Nicotine Tob Res*. 2018;22(8):1390–1398.

37. Chandra S, Scharf D, Shiffman S. Within-day temporal patterns of smoking, withdrawal symptoms, and craving. *Drug Alcohol Depend*. 2011;117(2-3):118–125.

38. Robinson JD, Li L, Chen M, *et al*. Evaluating the temporal relationships between withdrawal symptoms and smoking relapse. *Psychol Addict Behav*. 2019;33(2):105–116.

39. Vonesh EF, Greene T, Schluchter MD. Shared parameter models for the joint analysis of longitudinal data and event times. *Stat Med*. 2006;25(1):143–163.

40. Lin X, Mermelstein R, Hedeker D. Analysis of multivariate longitudinal substance use outcomes using multivariate mixed cumulative logit model. *BMC Med Res Methodol*. 2021;21(1):239.

41. Muthén B, Brown HC. Estimating drug effects in the presence of placebo response: causal inference using growth mixture modeling. *Stat Med*. 2009;28(27):3363–3385.

42. Jung T, Wickrama KAS. An introduction to latent class growth analysis and growth mixture modeling. *Soc Personal Psychol Compass*. 2008;2(1):302–317.

43. Andruff H, Carraro N, Thompson A, Gaudreau P, Louvet B. Latent class growth modelling: a tutorial. *Tutorials in Quantitative Methods for Psychology*. 2009;5(1):11–24.

44. Collins LM, Lanza ST. *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. Hoboken, NJ: Wiley; 2009.

45. Muthén B, Asparouhov T, Boye M, Hackshaw M, Naegeli A. *Applications of continuous-time survival in latent variable models for the analysis of oncology randomized clinical trial data using Mplus*. Los Angeles, CA: Muthén & Muthén;2009.

46. Muthén B, Masyn K. Discrete-time survival mixture analysis. *J Educ Behav Stat*. 2005;30(1):27–58.

47. Aigner CJ, Cinciripini PM, Anderson KO, *et al*. The association of pain with smoking and quit attempts in an electronic diary study of cancer patients trying to quit. *Nicotine Tob Res*. 2016;18(6):1449–1455.

48. Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. Hoboken, NJ: John Wiley & Sons, Inc; 2002.

49. Laird NM. Missing data in longitudinal studies. *Stat Med*. 1988;7(1-2):305–315.

50. Schneider KL, Hedeker D, Bailey KC, Cook JW, Spring B. A comment on analyzing addictive behaviors over time. *Nicotine Tob Res*. 2010;12(4):445–448.

51. Hedeker D, du Toit SHC, Demirtas H, Gibbons RD. A note on marginalization of regression parameters from mixed models of binary outcomes. *Biometrics*. 2018;74(1):354–361.

52. Schafer JL, Olsen MK. Multiple imputation for multivariate missing-data problems: a data analyst's perspective. *Multivariate Behav Res*. 1998;33(4):545–571.

53. Little JAR. A class of pattern-mixture models for normal incomplete data. *Biometrika*. 1994;81(3):471–483.

54. Glynn RJ, Laird NM, Rubin DB. Selection modeling versus mixture modeling with nonignorable nonresponse. In: Wainer H, ed. *Drawing Inferences from Self-Selected Samples*. New York, NY: Springer New York; 1986:115–142.

55. Ten Have TR, Kunselman AR, Pulkstenis EP, Landis JR. Mixed effects logistic regression models for longitudinal binary response data with informative drop-out. *Biometrics*. 1998;54(1):367–383.

56. Niaura R, Borreli B, Hedeker D, *et al*. Multicenter trial of fluoxetine as an adjunct to behavioral smoking cessation treatment. *J Consult Clin Psychol*. 2002;70(5):887–896.

57. Hedecker D, Rose JS. The natural history of smoking: a pattern-mixture random-effects regression model. In: Rose JS, Chassin L, Presson CC, Sherman SJ, eds. *Multivariate applications in substance use research: New methods for new questions*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 2000:79–112.

58. Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. New York: Wiley; 2014.

59. Troxel AB, Harrington DP, Lipsitz SR. Analysis of longitudinal data with non-ignorable non-monotone missing values. *J R Stat Soc Ser C Appl Stat*. 1998;47(3):425–438.

60. Muthén BO, Curran PJ. General longitudinal modeling of individual differences in modeling of individual differences in experimental designs: a latent variable framework for analysis and power estimation. *Psychol Methods*. 1997;2(4):371–402.

61. Hedeker D, Gibbons RD, Waternaux C. Sample size estimation for longitudinal designs with attrition: comparing time-related contrasts between two groups. *J Educ Behav Stat*. 1999;24(1):70–93.

62. Maxwell SE, Kelley K, Rausch JR. Sample size planning for statistical power and accuracy in parameter estimation. *Annu Rev Psychol*. 2008;59:537–563.

63. Winkens B, Schouten HJA, van Breukelen GJP, Berger MPF. Optimal number of repeated measures and group sizes in clinical trials with linearly divergent treatment effects. *Contemp Clin Trials*. 2006;27(1):57–69.

64. Candel MJ, van Breukelen GJ. Best (but oft forgotten) practices: efficient sample sizes for commonly used trial designs. *Am J Clin Nutr*. 2023;117(6):1063–1085.

65. Moerbeek M. Powerful and cost-efficient designs for longitudinal intervention studies with two treatment groups. *J Educ Behav Stat*. 2008;33(1):41–61.

66. Wang M. Generalized estimating equations in longitudinal data analysis: a review and recent developments. *Adv Stat*. 2014;2014(Article ID 303728). https://www.hindawi.com/journals/as/2014/303728

67. Asparouhov T, Hamaker EL, Muthén B. Dynamic structural equation models. *Struct Equ Modeling*. 2018;25(3):359–388.

68. Tan X, Shiyko MP, Li R, Li Y, Dierker L. A time-varying effect model for intensive longitudinal data. *Psychol Methods*. 2012;17(1):61–77.

69. Shiyko MP, Lanza ST, Tan X, Li R, Shiffman S. Using the time-varying effect model (TVEM) to examine dynamic associations between negative affect and self confidence on smoking urges: differences between successful quitters and relapsers. *Prev Sci*. 2012;13(3):288–299.

70. Hedeker D, Mermelstein RJ, Demirtas H. Analysis of binary outcomes with missing data: missing = smoking, last observation carried forward, and a little multiple imputation. *Addiction*. 2007;102(10):1564–1573.

71. Jebb AT, Tay L. Introduction to time series analysis for organizational research: methods for longitudinal analyses. *Organ Res Methods*. 2017;20(1):61–94.