A Comparison of Mathematical and Statistical Modeling with Longitudinal Data: An Application to Ecological Momentary Assessment of Behavior Change in Problem Drinkers

Kidist Bekele-Maxiwell

North Carolina State University

Judith Canner

California State University Monterey Bay

R. A. Everett

Haverford College

Sijing Shao

Northwell Health

Alexis Kuerbis

Hunter College at City University of New York

Destiny Stephenson

North Carolina State University

Jennifer Menda

North Carolina State University

Jon Morgenstern

Northwell Health

H.T. Banks

North Carolina State University

Abstract

Ecological momentary assessment (EMA) has been broadly used to collect real-time longitudinal data in behavioral research. Several analytic methods have been applied to EMA data to understand the changes of motivation, behavior and emotions on a daily or within-in daily basis. One challenge when utilizing those methods on intensive datasets in the behavioral field is to understand *when* and *why* the methods are appropriate to investigate particular research questions. In this manuscript, we compared two widely used methods (Generalized Estimating Equations and Generalized Linear Mixed Models) in behavioral research with three other less known methods (Markov Models, Generalized Linear Mixed-Effects Markov Models, and Differential Equations) in behavioral research but widely used in other fields. The aim is to provide the social science researchers a starting point to explore *what* new information can be obtained by mathematical models, which are usually lost in statistical models.

Keywords: ecological momentary assessment data, generalized estimating equations, generalized linear mixed models, markov models, generalized linear mixed-effects markov models, differential equations

Introduction

In behavioral research, intensive longitudinal data are ideal for understanding the dynamic processes of behavior change over time. Such data are especially useful for understanding constructs that change on a daily or within-day basis, such as motivation and self-efficacy for reducing drinking. Until relatively recently, such intensive data collection methods, referred to broadly as ecological momentary assessment (EMA), were too onerous to implement widely (Shiffman, Stone, & Hufford, 2008); however, advances in technology over the past two decades have allowed for the creation of inexpensive, efficient, and accurate real-time data collection methods (Morgenstern, Kuerbis, & Muench, 2014). These technological innovations have caused the proliferation of EMA across studies and behaviors (e.g., Carney, Armeli,

Tennen, Affleck, & O'Neil, 2000; Jahnel, Ferguson, Shiffman, Thrul, & Schuz, 2017; Kelly & Stephen, 2016; Tennen, Affleck, Armeli, & Carney, 2000; Todd et al., 2005), advancing the state of the science of behavioral research in cutting-edge ways, such as leading to the design and implementation of just in time interventions (e.g., Ford et al., 2015; Gustafson et al., 2014; Quanbeck et al., 2014).

Given the overwhelming amount and complexity of EMA data, several analytic procedures can be applied to investigate particular research questions. In the field of behavioral health, some procedures are more widely known and utilized than others. While advances have been made in applying new or rarely used methods to intensive longitudinal behavioral research, some methods appear to remain siloed among the few researchers who are aware of them.

The purpose of this manuscript is to illustrate the application of five distinct analytic methods to one dataset of intensive longitudinal data, highlighting the utility of each method—thus encouraging researchers to expand their ideas about what can be gleaned from using a single dataset. Specifically, we consider generalized estimating equations (GEE), generalized linear mixed-effects models (GLMM), Markov models (MM), generalized linear mixed-effects Markov models (GLMMM), and differential equations (DE) in addressing distinct yet related research questions. While we recognize that choosing analytic methods is reliant on both a specific conceptual model and the structure of the data (Hallgren, Atkins, & Witkiewitz, 2016), we tested closely related research questions based on a general conceptual model of factors that explain how problem drinkers maintain and/or might resist heavy drinking--highlighting the similarities, differences, strengths and limitations across methods.

Importantly, this is not an exhaustive list of possible analytic methods for intensive longitudinal data. We chose these five procedures as a means to compare two widely used

methods in behavioral research (GEE and GLMM) with three other methods (MM, GLMMM and DE) that are less well known or utilized within behavioral health but are widely used in areas such as biology, medicine, and physics. While others performed similar work comparing subsets of these four methods (e.g., Albert & Follmann, 2007; Shirley, Small, Lynch, Maisto, & Oslin, 2010), this would be the first manuscript known to these authors to include a comparison of DE with more commonly used statistical methods using EMA data.

Our work grew out of an interdisciplinary collaboration across mathematicians (RE, KB, HTB, SS), a statistician (JC), and social scientists (AK, JM). This exploratory exercise was instrumental--providing a foundational, mutual language for our work together. We aim to provide a starting point for others, particularly as mathematics becomes a more prominent part of analyzing data in social science. This manuscript is a summary of extensive work performed exploring each method, with an emphasis for a social science audience. Thus, the developmental history and mathematical details of each method, as well as the technical direction of how to implement each, will not be provided, but further details may be found in the provided references and code for each of these methods is available on request to the corresponding author.

The Data and Conceptual Model

Data utilized for this methodological comparison was collected during a study called Project SMART (Morgenstern et al., 2012). In a double-blind randomized controlled trial, Project SMART investigated the combined effectiveness of modified behavioral self-control therapy (MBSCT) and the medication naltrexone (NTX) for problem drinkers who wished to moderate rather than quit drinking. Problem drinkers (N=200), characterized by mild to moderate drinking problems with a high level of psychosocial functioning, were randomized to one of four conditions: placebo only (PBO), NTX only, MBSCT only, or combined NTX and MBSCT (NTX

+ MBSCT). The treatment period lasted 12 weeks, and participants completed an in-person battery of self-report measures at both baseline and week 13. During the 12-week treatment period, EMA data, specifically the subtype of EMA called daily diary data, was collected using Interactive Voice Recording (TELESAGE, 2005), in which participants completed a daily telephone survey each evening, with a total possible 84 days of data for each participant. Each survey took between 2-5 minutes to complete. Further details on the treatment interventions, IVR methodology, and study design are described in Morgenstern et al. (2012) and Mereish, Kuerbis, and Morgenstern (2018).

Participants

Participants were problem drinking men who have sex with men interested in moderated drinking rather than abstinence. The main criterion for study enrollment was that men had to consume > 24 standard alcoholic drinks per week, without demonstrating current or a history of withdrawal symptoms (Morgenstern et al., 2012). Participants were majority White (74%), attended at least some college or more (94%), employed (76%) and had a mean age of 40.3 years old (SD=11.1). Participants reported consuming a weekly mean of 43.1 standard drinks (SD=25.4) at baseline, and 93% met criteria for DSM-IV alcohol dependence.

Conceptual Model

As a part of our ongoing efforts in intervention research with problem drinkers (Bekele-Maxwell et al., 2017; Kuerbis, Armeli, Muench, & Morgenstern, 2014; Morgenstern et al., 2016), we adopted a “dual process” theoretical framework of substance abuse (Morgenstern, Naqvi, DeBellis, & Breiter, 2013) to select variables that we theorized related to alcohol consumption over time. This neurocognitive framework for addiction posits a “top-down, bottom-up” cognitive process in which executive functioning (“top down”, such as commitment

not to drink) exerts effort to inhibit both explicit and implicit responses to stimuli ("bottom up", such as the smell of alcohol). The variables selected to represent this framework were: alcohol consumption, personal norm violation, confidence to resist drinking heavily (self-efficacy), and commitment to resist drinking heavily (motivation). Craving, an important component of the dual process model, was excluded from this model formulation to be consistent with previous analyses. We then added treatment as an additional component impacting alcohol consumption.

Variables of Interest

Alcohol consumption. Alcohol use was assessed via multiple items within the IVR survey that asked participants to report the number of drinks they consumed in the last 24 hours. Each item asked about a different type of alcohol (e.g., beer, wine). Total number of drinks consumed in the last 24 hours was calculated by summing these items together to yield a daily count of drinks per day. We then averaged the daily data into weekly averages to yield a mean drinks per day (DPD) response variable.

In places where there may have been missing days of IVR data, Timeline Followback (TLFB, Sobell, Sobell, Leo, & Cancilla, 1988) data collected during post treatment assessments were used to calculate average DPD for the same week where there was a missing day. The calculated TLFB average replaced the missing daily value.

DPD was used in DE as a continuous response variable; however, for the purposes of an easier comparison of GEE, GLMM, and GLM/MM, as well as to avoid issues related to zero inflated regression analyses, we created a binary response variable: social drinking (DPD less than or equal to 4, including 0) versus heavy drinking (DPD greater than 4), using guidelines from the National Institute on Alcohol Abuse and Alcoholism guidelines (National Institute on Alcohol Abuse and Alcoholism, 2013).

Personal norm violation. Norm violation was assessed by asking, "*Do you consider the total amount you have had to drink since this time yesterday to be excessive? That is, was it more than you think you should have had?"* The response set ranged from 0 (*Definitely Not*) to 3 (*Definitely*). The thresholds (i.e., norms) individuals used to evaluate whether or not their drinking was excessive were not predefined, thus considered a "personal norm", and were not explicitly measured.

Confidence to resist drinking heavily. Confidence was measured by asking, "*How confident are you that you can resist drinking heavily (that is, resist drinking more than 5 standard drinks) over the next 24 hours?"* The response set ranged from 0 (*Not at all*) to 4 (*Extremely*).

Commitment to resist drinking heavily. Commitment was measured by asking, "*How committed are you not to drink heavily (that is, not to drink more than 5 standard drinks) over the next 24 hours?"* The response set ranged from 0 (*Not at all*) to 4 (*Extremely*).

MBSCT. A dichotomous variable indicated whether a participant was assigned to the behavioral therapy condition (MBSCT=1), in which they would receive 12 sessions of modified behavioral self-control therapy or no therapy (MBSCT=0).

NTX. A dichotomous variable was used to indicate whether a participant received active medication NTX (NTX=1) or placebo (NTX=0).

Time. To account for time, a count of the weeks during treatment was used, ranging from 1 to 12.

Model Descriptions and Results

For each of the modeling approaches described below, we introduce the method, propose the appropriate research question related to our conceptual model, and then describe the analytic

model design within the context of Project SMART. Next, we present model results and their interpretation. Initial comparison of the basic qualities of each of the methods, as well as their respective research questions, are described in Table 1. A comparison of each model relative to the others will be reviewed in the Discussion section.

Table 1: Comparison of all of all five models as they may be broadly applied to longitudinal studies of behavior and in this study

Model	**GEE**	**GLMM**	**MM**	**GLMMM**	**DE**
Research Question	*How do treatment (therapy (MBSCT) and/or medication (NTX)), norm violation, confidence to resist drinking, and commitment to resist drinking affect the likelihood of heavy drinking versus social drinking over time?*	*How do treatment (therapy (MBSCT) and/or medication (NTX)), norm violation, confidence to resist drinking, and commitment to resist drinking affect the likelihood of heavy drinking versus social drinking over time?*	*How does treatment (therapy and/or medication) affect the probability of transitioning from one level of alcohol consumption to another level over time?*	*How does treatment (therapy and/or medication), norm violation, confidence to resist drinking, and commitment to resist drinking affect the probability of transitioning from one level of alcohol consumption to another level over time?*	*How do the factors alcohol consumption, norm violation, confidence, and commitment relate to each other to cause the mechanisms for behavior change in individual participants?*
Applications and Assumptions of the Model	Used when interested in the average over intensive observations. Robust to violation of independence of observations.	Used when there are multiple levels and one is interested in exploring relationships at multiple levels. Robust to violation of independence of observations	Assumes first-order Markov process.	Assumes first-order Markov process. Similar to GLMM	Used when interested in investigating the complexity of the relationships that variables have with one another.
Assumptions about Missing Data	Missing completely at random (MCAR)	Missing at random (MAR)	MCAR	MCAR	Not Applicable
Structure of the Data	Long format (one row for each observation)	Long format Nested data	Long format Nested data	Long format Nested data	Long format
Level of Statistical Inference	**Population**	**Population** **Individual**	**Population** **Individual**	**Population** **Individual**	Population **Individual**
Level of Model Interpretation	**Population**	**Population** Individual	**Population** **Individual**	**Population** Individual	Population **Individual**
Response Variable	**Categorical** Continuous	**Categorical** Continuous	**Categorical** Continuous	**Categorical** Continuous	**Continuous**

Predictor Variables	**Categorical Continuous**	**Categorical Continuous**	**Categorical**	**Categorical Continuous**	**Continuous**
Type of Model	Marginal	Conditional	Marginal	Marginal	Not Applicable
Estimation Methods	**Quasilikelihood Estimation**	**Maximum Likelihood Estimation** Bayesian Estimation	**Maximum Likelihood Estimation** Bayesian Estimation	**Maximum Likelihood Estimation** Bayesian Estimation	Ordinary Least Squares **Iterative Reweighted Weighted Least Squares** Bayesian Estimation
Other studies that have used this method	Morgenstern et al. 2012 Morgenstern, et al. 2009	Mereish et al., 2018; Kuerbis et al., 2018 Livingston et al., 2017	Yeh et al., 2012	Shirley et al., 2010; Maruotti & Rocci, 2012	Banks et al., 2015
Software and related procedure or commands	**SAS** proc GEE or **proc genmod with repeated measures statement**	**SAS** proc mixed or **proc glimmix**	**R Package 'markovchain'** **Function markovchainFit()**	**R package 'LMest'** **Function est_mc_cov_latent()**	**MATLAB** **DE solver: dde23** **Function: fmincon for parameter estimation**

Note: Bolded text indicates the specific model choices in the application of the models to the Project SMART related research questions. Specific code is available on request. Supporting documentation for SAS programming for GEE can be found at http://support.sas.com/documentation/cdl/en/statug/68162/HTML/default/viewer.htm#statug_gee_overview.htm and https://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_genmod_sect008.htm. Supporting documentation for SAS and GLMM can be found at https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#mixed_toc.htm and https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#glimmix_toc.htm. Supporting documentation for R package 'markovchain' can be found at https://cran.r-project.org/package=markovchain Supporting documentation for R package 'LMest' can be found at https://www.jstatsoft.org/article/view/v081i04 and

https://cran.r-project.org/package=LMest
Supporting documentation for MATLAB package ‘dde23’ and ‘fmincon” respectively can be found at
https://www.mathworks.com/help/matlab/ref/dde23.html
https://www.mathworks.com/help/optim/ug/fmincon.html

Generalized Estimating Equations (GEE)

GEE (Zeger, Liang, & Albert, 1988) are a population-level statistical method commonly used to study longitudinal, repeated measures data. GEE extend generalized linear models (GLMs) by correcting for correlated observations (Stokes et al., 2000), and thus they are robust to the violation of the assumption of independence of observations. GEE are therefore appropriate for longitudinal panel designs or intensive longitudinal data-- data that is often collected in alcohol treatment clinical trials. As with GLMs broadly, GEEs allow response variables to be distributed non-normally. Parameters are estimated in GEE by solving a function similar to the quasilikelihood function (Liang & Zeger, 1986). In the formulation of the GEE, one must choose a correlation structure to model the relationship between time-dependent observations. There are several options, (e.g. exchangeable, autoregressive) depending on the known or hypothesized correction structure of the data, but the GEE method is generally robust to misspecification of the correlation structure (Liang and Zeger, 1986; Ballinger, 2004). Thus, parameter estimates are reasonably accurate even if the assumed correlation structure is "wrong".

Application to Project SMART. Within the context of Project SMART, one potential research question is: *How do treatment (therapy (MBSCT) and/or medication (NTX)), norm violation, confidence to resist drinking heavily, and commitment to resist drinking heavily affect the likelihood of heavy (versus social) drinking over time?* We used GEE to test for both the main effects and interactions between predictors and time, as well as for the combined effect of MBSCT and NTX. For this analysis, a binomial distribution with logit link function was specified due to the binary response of heavy or social drinking behavior. In addition, an autoregressive working correlation matrix was specified (Stokes et al., 2000), which assumes that the correlation is dependent upon the amount of time between observations, that is, average

alcohol consumption in a week is correlated more with one week ago than several weeks ago. Analyses were conducted using PROC GENMOD in the SAS statistical software program (SAS Institute Inc., 2002-2012).

We reduced the model predictors in a stepwise fashion, successively eliminating each term that was not statistically significant. The final model yielded four significant predictors of drinking across the treatment period: MBSCT, confidence to resist drinking heavily, norm violation, and time (Table 2). None of the interaction terms included in the original model were statistically significant. When keeping the other predictors constant in the model, participants in MBSCT had 49% lower odds of heavy drinking compared to those not in MBSCT. For every unit increase in confidence, participants had 41% lower odds of drinking heavily. Time also had a negative relationship with heavy drinking. For every week that passed during the treatment period, participants had 7% lower odds of drinking heavily. Only personal norm violation demonstrated a positive relationship with heavy drinking, such that for every unit increase in rating drinking as excessive (personal norm violation), participants were 3.7 times greater odds of drinking heavily that same week.

Table 2: Parameter estimates from GEE and GLMM predicting the likelihood of weekly heavy drinking compared to social drinking and GLMMM predicting transition probabilities to and from social and heavy drinking. Odds ratios (OR) and the lower limit (LLCI) and upper limit (ULCI) for the 95% confidence interval for the OR of each statistically significant predictor.

GEE					
Predictors	***β***	**SE**	**LLCI**	**OR**	**ULCI**
MBSCT	-0.67	0.22	0.41	0.51	0.64
Confidence to resist drinking	-0.53	0.08	0.54	0.59	0.64
Personal Norm Violation	1.30	0.12	3.30	3.70	4.10
Time (Week)	-0.07	0.02	0.91	0.93	0.95
GLMM					
Predictors	***β***	**SE**	**LLCI**	**OR**	**ULCI**
MBSCT	-0.97	0.37	0.26	0.38	0.55
Confidence to resist drinking	-0.72	0.12	0.43	0.47	0.55
Personal Norm Violation	2.16	0.18	7.20	8.60	10.40

Time (Week)	-0.15	0.03	0.84	0.86	0.89
GLMMM – Social to Heavy					
Predictors	***β***	**SE**	**LLCI**	**OR**	**ULCI**
Naltrexone	-0.97	0.26	0.23	0.38	0.63
MBSCT	-0.95	0.25	0.24	0.39	0.63
MBSCT*Naltrexone	1.12	0.35	1.53	3.07	6.14
Personal Norm Violation	0.91	0.11	2.00	2.50	3.11
Confidence to resist drinking	-0.27	0.08	0.65	0.76	0.90
Time (Week)	-0.07	0.03	0.88	0.93	0.98
GLMMM – Heavy to Social					
Predictors	***β***	**SE**	**LLCI**	**OR**	**ULCI**
Naltrexone	-0.53	0.24	0.37	0.59	0.93
MBSCT*Naltrexone	1.34	0.28	2.19	3.82	6.69
Personal Norm Violation	-1.17	0.14	0.24	0.31	0.41
Confidence to resist drinking	0.78	0.10	1.79	2.19	2.67

Generalized Linear Mixed Models (GLMMs)

Generalized linear mixed effects models (GLMMs) are another common extension of GLMs used for analyzing longitudinal data. The model specification is similar to the GEE, again using interaction between time and other predictors to test change in effect over time. GLMMs use both fixed-effects (population level) and random-effects (individual level) parameters. A random intercept in the model accounts for the correlation among observations on the same individual over time through the quantification of the variability between individuals, in contrast to GEEs which use a 'working' correlation structure. Therefore GLMMs also are appropriate to test for relationships at the population level, while controlling for individual variability in both the intercept and in change over time. If, however, the random effects are not statistically significant, indicating that there is insufficient individual variability to account for them, then the random effects would be removed from the model (Singer & Willett, 2003). Thus, what remains is a GLM more similar to the GEE above. To estimate parameter values, we used maximum likelihood estimation methods as we had sufficient power and data to use this method (Capanu, Gönen, & Begg, 2013).

Application to Project SMART. The research question for GLMM would be quite similar to that for GEE: *How do treatment (therapy (MBSCT) and/or medication (NTX)), norm violation, confidence to resist drinking heavily, and commitment to resist drinking heavily affect the likelihood of heavy (versus social) drinking over time?* We again tested for main effects, interaction effects between predictors and time, as well as for the combined effect of MBSCT and NTX. For this analysis, a binomial distribution was specified with a logit link function. Analyses were conducted using the GLIMMIX procedure in the SAS statistical software program (SAS Institute Inc., 2002-2012). All models included both random intercept and slope for the count of EMA days in the study, permitting individual variability in drinking over time.

Again we reduced the model in a stepwise fashion and successively eliminated each term that was not statistically significant. Not surprisingly, the final GLMM yielded the same four significant predictors of drinking across the treatment period as the GEE: MBSCT, confidence to resist drinking heavily, norm violation, and time (Table 2). None of the interaction terms included in the original model were statistically significant. When keeping the other predictors constant in the model, participants in MBSCT had 62% lower odds of heavy drinking compared to those not in MBSCT. For every unit increase in confidence, participants had 53% lower odds of drinking heavily. For every week that passed during the treatment period, participants had 14% lower odds of drinking heavily. For norm violation, for every unit increase in rating drinking as excessive (personal norm violation), participants had 8.6 times greater odds of drinking heavily.

Markov Models

While GEE and GLMM work well to look at the impact of predictors on the observable response, in the study of behavior, it is also informative to consider the influence of predictors on

the likelihood of *transition* from one behavioral state to another. In such situations, Markov models (MMs) are an appropriate tool to explore the transitions between behavior states over time. Markov models have been used in psychology applications for at least 60 years (e.g., Miller, 1952). A more recent example includes Yeh et al. (2012), where the authors use MMs to analyze smoking behavior, specifically to identify factors associated with the transitions between smoking and abstinence. MMs have been used previously in the application of alcohol-related studies (e.g., Shirley et al., 2010; Maruotti and Rocci, 2012), as well as extensions to Hidden Markov Models, which are beyond the scope of this paper. Regardless, MMs remain underutilized perhaps due to their relative inaccessibility in common software packages (de Haan-Rietdijk et al., 2017).

To define a MM, we must first define the Markov Process. We assume that any given time point, t, an individual, i, will be in a specific observable state, $s_{i,t}$. While one could assume any number of possible states, we will assume just two states, e.g., social (0) or heavy (1) drinking states. The first-order Markov Process assumes that the probability of an individual, i, being in a particular state at a specific time point, t, is dependent only on what state the individual was at in the previous time point, $t-1$. For example, the probability of being in state 0 at time t is dependent on whether an individual was in state 0 or 1 at time $t-1$, or

$$P\left(s_{i,t}=0 \middle| s_{i,t-1}=0\right)=\pi_{i,00}$$

and

$$P\left(s_{i,t}=0 \middle| s_{i,t-1}=1\right)=\pi_{i,10}.$$

Note that $\pi_{i,00}+\pi_{i,01}=1$ and $\pi_{i,11}+\pi_{i,10}=1$, as the sum of the probabilities of observing one of the two possible events at time t, given the same origin state at time $t-1$, total one. We

can define a "transition matrix" to model the transition probability from every state to every other state in a single time-step. For a two-state Markov Process that matrix would be

$$\Pi_i = \begin{bmatrix} \pi_{i,00} & \pi_{i,01} \\ \pi_{i,10} & \pi_{i,11} \end{bmatrix}.$$

From the transition matrix, we can also estimate the *steady-state solution*, i.e., the probability that at any time point a particular individual is in a particular state. The steady-state solution would be similar to the probabilities associated with the response variable for the GLMM and GEE, i.e. the probability of being in a specific state of drinking at a specific time-point.

The Markov Process can be assumed to be time-homogenous, i.e. the transition probabilities do not vary over time, as above, or time-inhomogeneous, such that the transition probabilities vary over time. Transition probabilities may be estimated empirically via maximum likelihood estimation (Maruotti and Rocci, 2012) or Bayesian estimation methods (Shirley et al. 2010; de Haan-Rietdijk et al., 2017).

Application to Project SMART: We assumed a simple first-order MM for all participants and we estimated the transition matrix, using maximum likelihood estimation, from social drinking (0) to heavy drinking (1) states and vice versa. Analyses were conducted using the R statistical software program (R Core Team, 2018) with the 'markovchain' package (Spedicato, 2017). In general, we observed that participants were more likely to remain in their current drinking behavior state (probability of social to social state transition is 0.838; probability of heavy to heavy state transition is 0.746) than transition to different state the next week across all twelve weeks and 181 participants (Table 3.1).

Table 3.1: Estimated transition probabilities and steady-states for all participants.

Transition	To	Social	Heavy
From	Social	0.838	0.162
	Heavy	0.254	0.746
Steady-States		0.61	0.39

The research question of interest for the simple MM is: *How does treatment (MBSCT and/or NTX) affect the probability of transitioning from one level of alcohol consumption to another level over time?* We examined this question by looking at individual Markov models and treatment group Markov models. We estimated individual participant transition matrices for four patients, one from each treatment group (Table 3.2). For participants 1761, 1460, and 1474, once they transitioned to social drinking, they remained in that state through the end of the 12 weeks, as indicated by the probabilities of 1in the row associated with transitions from social drinking and remaining in a social drinking state. While informative at a participant level, the models do not provide much information on differences in treatments, prohibiting generalization to the treatment groups. To make more generalized inferences about the effects of treatment on behavior changes, we estimated the transition probabilities for all participants within each treatment group (Table 3.3a; 3.3b).

Table 3.2: Estimated transition matrices for one participant from each treatment group.

Transition		PID 1859 Placebo		PID 1761 MBSCT		PID 1460 NTX		PID 1474 NTX+MBSCT	
	To	Social	Heavy	Social	Heavy	Social	Heavy	Social	Heavy
From	Social	0.60	0.40	1	0	1	0	1	0
	Heavy	0.17	0.83	0.25	0.75	0.09	0.91	0.11	0.89

Table 3.3a: Estimated transition matrices for time-homogeneous Markov model for all participants within each treatment group, with standard errors in parentheses. Steady-state solutions provided for each treatment group.

Transition		Placebo		MBSCT		NTX		NTX+MBSCT	
	To	Social	Heavy	Social	Heavy	Social	Heavy	Social	Heavy
From	Social	0.687 (0.059)	0.313 (0.040)	0.894 (0.052)	0.106 (0.018)	0.874 (0.059)	0.126 (0.022)	0.842 (0.048)	0.158 (0.021)
	Heavy	0.206 (0.026)	0.794 (0.051)	0.272 (0.039)	0.727 (0.064)	0.165 (0.027)	0.835 (0.060)	0.508 (0.063)	0.492 (0.062)
Steady-States		0.397	0.603	0.720	0.280	0.565	0.435	0.763	0.237

Table 3.3b: Estimated 95% confidence interval (lower, upper) bounds for the transition matrices for time-homogeneous Markov model for all participants within each treatment group.

Transition		Placebo		MBSCT		NTX		NTX+MBSCT	
	To	Social	Heavy	Social	Heavy	Social	Heavy	Social	Heavy
From	Social	(0.590, 0.785)	(0.250, 0.379)	(0.808, 0.980)	(0.077, 0.136)	(0.777, 0.970)	(0.090, 0.163)	(0.763, 0.921)	(0.124, 0.192)
	Heavy	(0.163, 0.248)	(0.711, 0.877)	(0.208, 0.337)	(0.622, 0.833)	(0.121, 0.208)	(0.737, 0.934)	(0.404, 0.611)	(0.390, 0.594)

While there were no statistically significant differences between the transition rates among the treatments (excluding the Placebo group) from social to heavy drinking, there was a significant difference in transition rates from heavy to social drinking (Table 3.3b). Specifically, the NTX+MBSCT group has the greatest transition probability to social drinking after a heavy drinking state, a contrast to the GEE and GLMM models that found no significant effect of the combined treatments on the probability of heavy drinking. The steady-state solutions show that treatment was better than placebo at supporting social drinking behavior states and specifically that MBSCT led to a higher probability of social drinking with or without NTX (Table 3.3a) and that the addition of NTX to MBSCT does not improve the probability of being in a social state, similar to the results of the GEE and GLMM.

We also examined the time-inhomogeneous transition probabilities by treatment group from social to heavy (Figure 1.1) or heavy to social (Figure 1.2) to compare the effect on treatments on transition probabilities that vary over time. Again, while all of the treatments had varying, but similar, effects on the probability of transition from social to heavy drinking states (Figure 1.1), it was the combination of MBSCT and NTX that had the greatest impact on the probability of transitioning out of a heavy drinking state into a social drinking state (Figure 1.2).

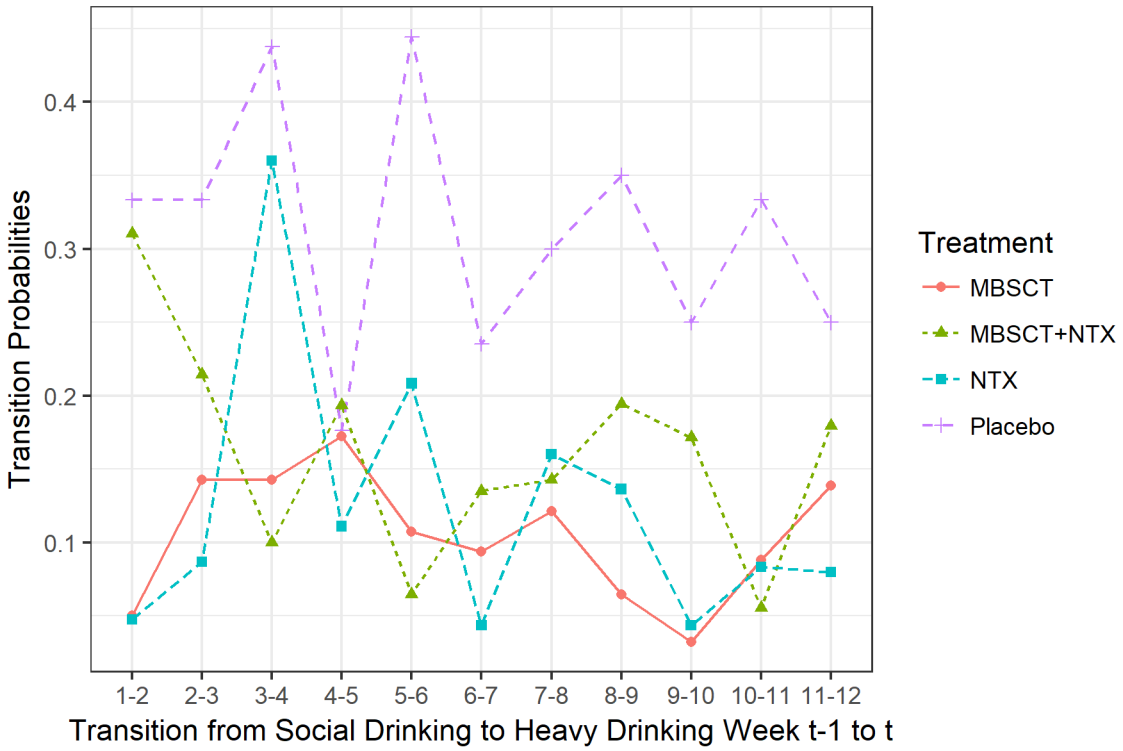

Figure 1.1: Estimated transition probabilities from week to week over the twelve weeks from social to heavy drinking.

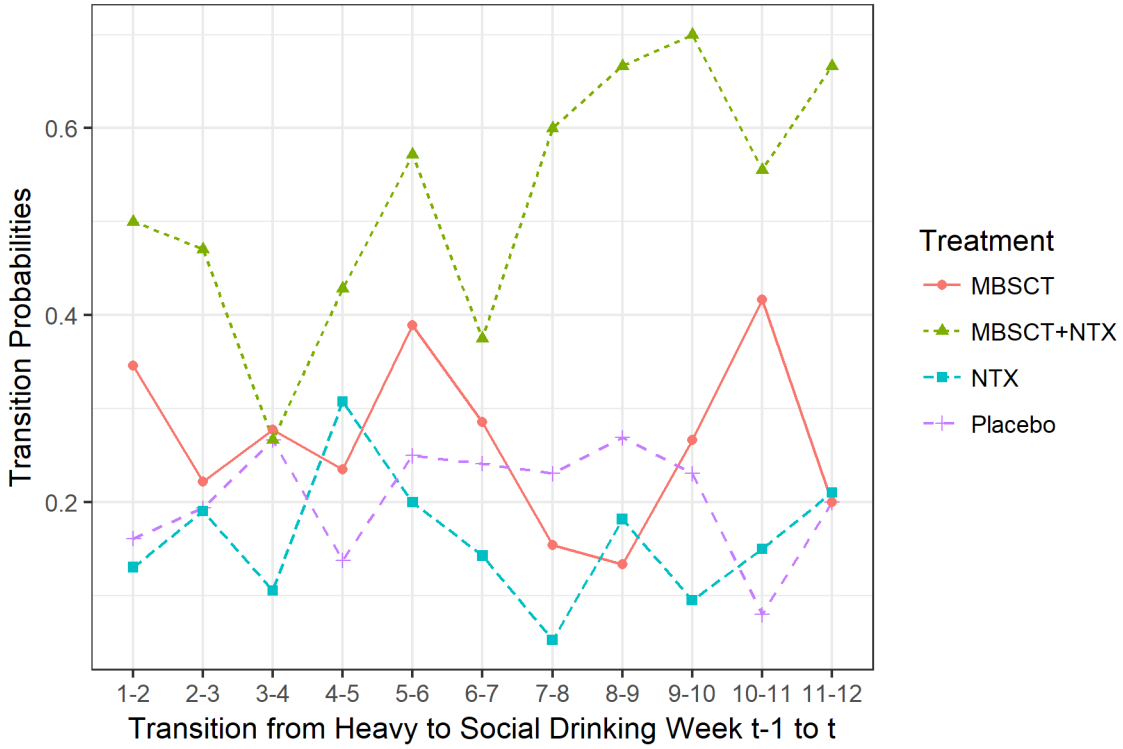

Figure 1.2: Estimated transition probabilities from week to week over the twelve weeks from heavy to social drinking

Generalized Linear Mixed-Effects Markov Model (GLMMM)

While useful at the group/population level, the previous MM approach did not allow for individual level effects and/or effects of time-varying predictors on transition probabilities, such as norm violation, confidence, and commitment. The erratic trend of the transition probabilities over time (Figure 1.1; Figure 1.2) may also indicate that other factors aside from treatment impact the transition probabilities. For GEE and GLMM, we developed linear models to examine the probability of being in a specific state at a specific time point. Here, we can extend MMs to build coupled GLMMs, one per each row of the transition matrix, to model the probability of transitioning from social to heavy drinking at a specific time point (Equation 1), and the probability of transitioning from heavy to social drinking at a specific time point (Equation 2) such that:

$$\log\left(\frac{\pi_{i,01}}{\pi_{i,00}}\right) = \text{logit}(\pi_{i,01}) = \mathbf{x}_i'\boldsymbol{\beta}_{01} + \boldsymbol{\gamma}_{i,01} \qquad (1)$$

$$\log\left(\frac{\pi_{i,10}}{\pi_{i,11}}\right) = \text{logit}(\pi_{i,10}) = \mathbf{x}_i'\boldsymbol{\beta}_{10} + \boldsymbol{\gamma}_{i,10}\ , \qquad (2)$$

where $\mathbf{x}_i{}'$ are the predictor values and $\boldsymbol{\beta}_{10}$ and $\boldsymbol{\beta}_{01}$are the coefficient estimates, and $\boldsymbol{\gamma}_{i,10}$ and $\boldsymbol{\gamma}_{i,01}$ are the random intercepts for each model respectively. Note that there is a time-inhomogeneous GLMMM where participant transition probabilities and coefficients of time-varying predictors are estimated over time. We recognize that a time-inhomogeneous GLMMM may be more practical for modeling behavior modification, but for the purposes of this paper, we implemented the simpler model, where transition probabilities were homogeneous over time.

Application to Project SMART: The research question for GLMMM would be: *How does treatment (MBSCT and/or NTX), norm violation, confidence to resist drinking, and commitment to resist drinking affect the probability of transitioning from one level of alcohol*

consumption to another level over time? Similar to GLMM and GEE, we tested for both the main effects and the interaction effects between predictors and time as well as for the combined effect of MBSCT and NTX. Analyses were conducted using R statistical software program (R core team, 2018) with the package 'LMest' (Bartolucci, Pandolfi, and Pennoni, 2017). The function for estimation of the models in 'LMest' required the imputation of missing data for weekly observations of norm violation, confidence, and commitment. We used predictive mean matching to impute the missing values for norm violation, confidence, and commitment via the R package 'mice' (van Buuren & Groothuis-Oudshoorn, 2011).

Again we reduced the two models (one to model the probability of transitioning from social to heavy drinking; and the second to model the probability of transitioning from heavy to social drinking) in a stepwise fashion and successively eliminated each term that was not statistically significant (Table 2). In contrast to GLMM and GEE results, time was not significant for the heavy to social transition model and had only a weak effect on social to heavy transitions (OR ~1) so there was only a slight decrease in likelihood of transitioning from social to heavy drinking over time (~7% decline). In addition, the interaction between MBSCT and NTX was significant in both models. We see that participants in a non-placebo group (NTX: OR=0.38; MBSCT: OR=0.39; MBSCT+NTX: OR=0.45) had 55-62% lower odds of transitioning from a social drinking state to a heavy drinking state, compared to the placebo group. The combination of MBSCT and NTX had less of a combined impact on the odds of transitioning from social to heavy drinking compared to MBSCT and NTX alone. In addition, as confidence to resist heaving drinking increased, participants were 24% less likely to transition from social to heavy drinking. Personal norm violation was significant as well, but merely verified that patients who transitioned to heavy drinking recognized that they were heavy drinking. In the model for heavy

to social transition probabilities, we see that a participant in the combined MBSCT and NTX group were 2.25 times more likely to transition from heavy to social drinking, compared to the other treatment or placebo groups. Additionally, as confidence increased the likelihood of transitioning to social drinking from heavy drinking also increased.

Differential Equations (DE)

Differential equations (DE) are mathematical representations of how variables relate to themselves and each other over time. DE model the rate of change of an observed response over time as it relates to other predictors and their change over time. In a simple example, we consider that population size changes over time and may grow or decline depending on the number of births or deaths within the population. The number of births and deaths are proportional to how many individuals are within a population at a given time. Therefore, the rate of change in the population size is directly dependent on the current population size at any given time point as well as other variables and associated parameters. In mathematics, the rate of change is often called the *derivative with respect to time* of a function, often notated as $\frac{df(x,t)}{dt}$. DE may be developed from theory regarding cognitive processes related to addiction and then used to test those theories by determining if the model fits the data for variables measured related to those processes. While statistical methods, such as GEE, GLMM, or GLMMM, model the direct effect of predictors on a response variable, DE can model nonlinear effects of the predictors on the response; the effect of the predictors on each other; and the effect of the response on the predictors over time. Therefore, while there are similarities to the types of inference we might make with DE and statistical methods, DE allow for a more nuanced exploration of the relationships between all the variables of interest. However, in the current work, we consider continuous DE that assume the predictor and response variables are continuous. Similar to

statistical models, DE parameters are often estimated using least squares regression-like methods of minimizing the residual error between the model prediction and the data observation, often using ordinary least squares or iterative-reweighted weighted least squares regression techniques (Banks, et al., 2014). That is, the DE model solution represents the mean of the distribution of the response variable, and distributional assumptions are often incorporated in a statistical error model.

Application to Project SMART: We implemented a differential equation model to investigate the research question: *How do the factors alcohol consumption, norm violation, confidence to resist drinking heavily, and commitment to resist drinking heavily relate to each other to cause the mechanisms for behavior change in (four) individual participants?* In the context of Project SMART, we consider the change in *amount* of alcohol consumed over time instead of the two categories of social and heavy drinking as we did in our analysis using GLMM, GEE, and GLM/MM. DE will allow us to quantify the complex and nonlinear relationships between alcohol consumption, norm violation, confidence, and commitment over time to cause the mechanisms for behavior change in individual participants.

To create a mathematical representation of the conceptual model previously discussed, we let A represent the daily alcohol consumption, or the number of alcoholic drinks a person has consumed in the past 24 hours from time t (i.e., from time $t-1$ to time t); V represents norm violation in the past 24 hours; C_f represents the confidence level a person feels at time t to resist heavy drinking in the next 24 hours; and C_h represents commitment a person makes to not drink heavily in the next 24 hours at time t (i.e. from time t to time $t+1$). In addition, we defined a *latent* variable, personal norm, or A^*, which means it was not directly measured in the SMART study. We define this variable to be the threshold (i.e., norm) an individual used to evaluate

whether or not his drinking was excessive. The mathematical model is given by the following system of differential equations and the psychological processes considered in this mathematical model are very briefly explained directly below each term:

$$\frac{dA}{dt} = \underbrace{a_1}_{\text{desire}} - \underbrace{a_2 V(t-1)}_{\substack{\text{norm violation}\\ \text{affects drinking}}} - \underbrace{a_3 C_h(t-1)}_{\substack{\text{commitment}\\ \text{affects drinking}}} - \underbrace{a_4 C_f(t-1) C_h(t-1)}_{\substack{\text{interaction of confidence and}\\ \text{commitment affect drinking}}}$$

$$\frac{dV}{dt} = \underbrace{\chi_{(A>A^*)} v_1 \frac{d(A-A^*)}{dt}}_{\substack{\text{if drink more than personal norm:}\\ \text{rate that drinking approaches personal}\\ \text{norm affects norm violation}}} - \underbrace{\chi_{(A\le A^*)} v_2 V}_{\substack{\text{if drink less than personal norm:}\\ \text{norm violation soln decreases to 0}}}$$

$$\frac{dC_f}{dt} = \underbrace{-\chi_{(A>4)} d_1 \frac{dA}{dt} C_h(t-1)}_{\substack{\text{if drink heavily:}\\ \text{interaction of rate of drinking and}\\ \text{commitment affects norm violation}}} + \underbrace{\chi_{(A\le 4)}\, d_2 \left(C_f - \alpha\right)\left(1 - \frac{C_f}{n}\right)}_{\substack{\text{if drink socially:}\\ \text{confidence soln increases logistically}\\ \text{to max level over time}}}$$

$$\frac{dC_h}{dt} = \underbrace{m C_h \left(1 - \frac{C_h}{K}\right)}_{\substack{\text{commitment soln increases}\\ \text{logistically to max level over time}}},$$

where

$$A^* = \underbrace{b e^{-rt} + l}_{\substack{\text{personal norm decreases}\\ \text{to min level over time}}},$$

and χ is the indicator function defined as follows

$$\chi_{(A>x)} = \begin{cases} 1 \; if \; A > x \\ 0 \; else \end{cases}, \; \chi_{(A\le x)} = \begin{cases} 1 \; if \; A \le x \\ 0 \; else \end{cases},$$

where $x = A^*$ in the second differential equation and $x = 4$ in the third differential equation. We refer the reader to (Bekele-Maxwell et al., 2017) for further details of the mathematical model formulation, though they used a different threshold for heavy drinking. We estimated the DE parameters for each individual using an iterative reweighted weighted least squares (IRWLS) method (Banks, Hu, & Thompson, 2014) that aims to minimize a weighted distance between the weekly averaged data and weekly averaged model solution. Analyses were conducted using the

built-in function 'fmincon' in MATLAB (The MathWorks, 2015). The estimated parameter values for each individual are given in Table 4. Each resulting parameter set represents a patient and, combined with the model, produces individual solutions.

Table 4: DE model parameter estimates for each individual case

	PID			
Parameters	**1460**	**1474**	**1761**	**1859**
a_1	0.111	0.362	0.573	1.318
a_2	0.084	0.171	0.310	0.389
a_3	0.033	0.041	0.430	0.291
a_4	0.017	0.071	0.040	0.002
v_1	0.250	0.187	0.126	0.323
v_2	0.349	0.053	0.061	0.152
d_1	0.077	0.141	0.026	0.358
b	9.319	1.788	6.313	1.613
r	0.020	0.007	0.310	0.001
l	6.127	2.351	2.074	1.001
m	0.029	0.022	0.027	0.050
K	3.755	4.044	3.984	3.832
A_0	11.172	7.568	12.592	3.085
V_0	0.515	1.893	2.142	0.324
C_{f0}	1.955	0.868	1.476	3.396
C_{h0}	1.713	0.518	1.872	2.433
d_2	0.248	0.290	0.093	0.000
α	0.751	1.055	1.275	1.871
n	3.945	3.657	3.693	3.829

To best compare the model solutions to the previous methods, we present results for individual participants from each treatment group (Figure 2). The model solutions are able to reasonably capture the overall trend in the data for each participant, suggesting the accuracy in the hypothesized relationships among the factors. These results indicate that alcohol consumption, norm violation, confidence to resist drinking heavily, and commitment to resist drinking heavily influence each other (as described by the mathematical model) to produce the behavior change. In addition, the results indicate the importance of the "personal norm" threshold, which might be of interest for future studies.

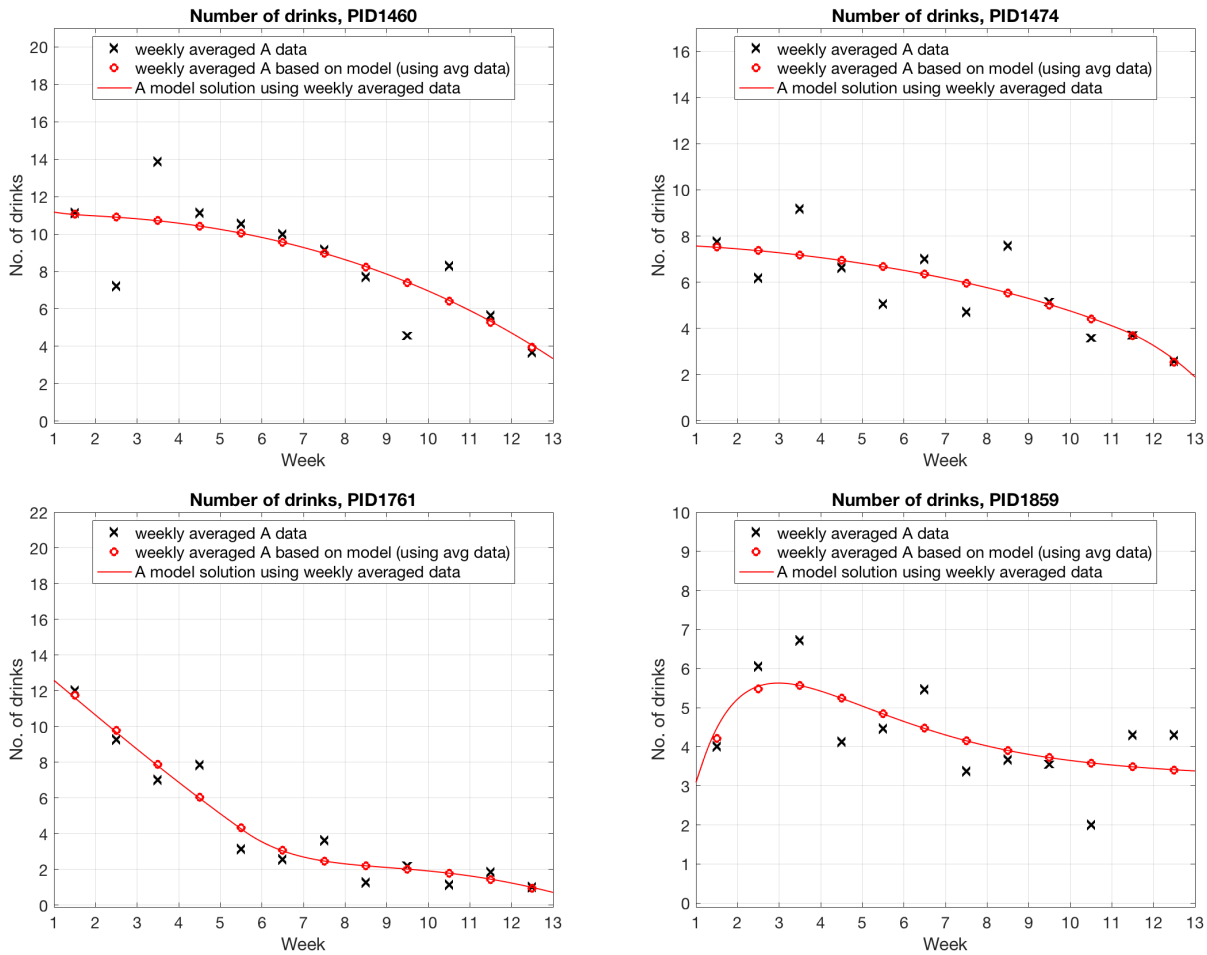

Figure 2: DE alcohol consumption model solutions for PID 1474 (MBSCT+NTX), PID 1761 (MBSCT), PID 1460 (NTX), and PID 1859 (PBO). The red line represents the model solution (solved on the daily scale), the red o's represent the weekly averaged alcohol consumption based on the model, and the black x's represent the weekly averaged alcohol consumption data.

Discussion

When designing a study, it is important to first identify the research question(s) of interest and the appropriate model for analysis so that the data collection design and process is effective and efficient. In this paper, we demonstrated how each modeling approach evaluated the conceptual model for changing or maintaining problem drinking, using observations from Project SMART. The two types of generalized linear models, GEE and GLMM, produced the same support for the proposed conceptual framework of the dual process neurocognitive model

of addiction. Both MBSCT and confidence to resist heavy drinking (self-efficacy) were negatively related to heavy drinking compared to social drinking, potentially indicating the increased use of learning, skills, and agency (top down executive functions) to moderate drinking.

In contrast, the GLM/MM results showed a combined effect for medication and psychotherapy on the probability of transitioning out of a heavy drinking state--a finding consistent with other research exploring combined treatments (Donovan et al., 2008). Though this appears counter to our GEE and GLMM results, as well as previous analyses (Morgenstern et al., 2012), it is simply due to a shift in perspective–the addition of naltrexone to therapy supports behavior change, but may only slightly increase time spent in a social drinking state overall. Given that this combined effect is slightly greater than the individual effects alone, the GLM/MM potentially supports the possibility of combined treatments addressing the top down-bottom up nature of addiction, as outlined in the conceptual model: naltrexone is thought to inhibit the strength of "bottom up" stimuli (craving), while MBSCT enhances "top down" executive process and skills to reduce drinking. These mechanisms of action would need to be tested explicitly to confirm these relationships. The fact that commitment was also a significant contributor to the transition between heavy to social drinking is consistent with our previous studies and the conceptual model, as commitment may be the single biggest top down contributor to changes in drinking behavior.

DE also validates the conceptual model—but provides very different information. While DE validates the selection of variables initially chosen, similar to statistical models, DE also demonstrates the potential complexity of the relationships that variables may have with one another—a complexity that is lost in statistical models due to the restrictions of linear models.

While DE can be applied at the population level, individual level analysis provides the opportunity to construct a much richer model than might emerge utilizing a group from the beginning. One may think of this as akin to the contributions of qualitative research—the purpose of qualitative research is not to provide generalizability to a population but instead to provide more in-depth information about a few individuals ultimately offering potential road maps for future investigation at the population level. By exploring data with one individual at a time with DE, in depth, non-linear relationships between variables can be described, providing the researcher with new information about how they may proceed constructing a mathematical model to be applied to a group, as well as inform the construction of potential statistical models, where feasible.

Choosing the "Best" Modeling Approach

We examined five different modeling methods in this paper, all of which could be further modified to include other formulations of our variables or perspectives on our research questions. Regardless, we see that each model has its purposes when examining behavior changes over time (Table 1). Too often, a method is chosen simply because it was employed previously for a similar dataset. While this choice might in fact be appropriate, it is necessary to understand *why* it is suitable. In order to do this, one must have an understanding of the method and underlying assumptions.

GLMMs and GEEs are both extensions of GLMs, although they handle the correlation within an individual over time differently. GEEs consider a 'working' correlation matrix and evaluates population-level effects only whereas GLMMs use random effects to handle time dependency and allow for inference for both fixed (population-level) and random (individual-level) effects. While GEEs or GLMMs can classify a participant into a specific category of

behavior, such as heavy drinking, smoking, or manic state, they do not directly provide a measure of the rate of behavior changes between states, such as from abstinence to relapse rates, which are conditional on prior behavior states. GLM/MMs allow us to model behavior changes and rates of behavior change over time, such as periods of remission or recidivism (e.g., Yeh et al. 2012). In addition, focusing on the behavior changes between states instead of only the observed behavior (outcome) may illuminate why a certain treatment is effective or ineffective. In focusing on the transition, we can observe if it failed to prevent a participant from transitioning to a "bad" behavior rather than only on helping the participant return to a healthy behavior. By focusing only on the outcome, the treatment may appear ineffective, when it may in fact be a powerful prevention measure. More complex GLM/MMs, like Hidden Markov models, can detect overarching behavior patterns, such as remission and relapse, and not just observed behavior states (e.g., Shirley et al., 2010; Maruotti and Rocci, 2012).

GLMMs, GEEs, and GLM/MMs determine the direct relationship of the predictors on some form of the observed response. In contrast, DEs quantify the often complex and nonlinear relationship between the predictor variables and the rate of change (i.e., derivative) of the mean of the response variable. Both DE and MM models may be fit to population level data or to data for a specific individual. This allows for personalization of the model, to examine an individual's behavior or to make predictions about an individual's behavior under different scenarios or perturbations of the parameter estimates.

Other Modeling Considerations

In the process of model building, it is important to recognize, and be transparent about, the choices that are statistical in nature, such as significance testing and model comparison measures, and the choices that are scientific in nature, such as the initial choice to include

specific predictors or model structure (nonlinear vs. linear formulation). One choice in the model building process that must be considered is the time-scale of the data. For example, time-lag for predictors may be appropriate in a mathematical or statistical model given the conceptual model being tested. The DE treats norm violation prospectively—if one violates their norm one day it influences drinking levels the next day. Another time-scale decision is the choice to represent data measured daily as a weekly average or some other time-scaled average. The choice to average data over time may depend on the research question, the type of behavior studied, or characteristics of the data itself, such as cases in which there may be missing data. Time-scaled averages do not necessarily dilute the impact of EMA, as a study of alcohol treatment clinical trials found that models testing for main effects of treatments using aggregate data performed equivalently and in some cases better than the more complicated data structures (Hallgren et al., 2016). Ultimately, the research question should drive the form of the data used in the modeling process.

Conclusion

While GEE and GLMM remain common approaches to analysis in longitudinal studies due to their utility and their accessibility in common statistical software, we showed how other modeling approaches, such as MMs, GLMMMs, and DEs might provide alternative perspectives and insight into the validity of conceptual models of behavior. Hopefully, as more researchers adopt these other modeling methods, statistical software will provide more accessible ways to build and analyze these types of models for future longitudinal studies of behavior. Until that time, we encourage the continued collaboration between the mathematical, statistical, and social scientists to support the important work of modeling behavior for treatment and intervention.

Acknowledgment This research was supported in part by the U.S. Air Force Office of Scientific Research under grant AFOSR FA9550-18-1-0457.

References

Albert, P. S., & Follmann, D. A. (2007). Random effects and latent processes approaches for analyzing binary longitudinal data with missingness: a comparison of approaches using opiate clinical trial data. *Stat Methods Med Res, 16*(5), 417-439. doi:10.1177/0962280206075308

Ballinger, G. A. (2004). Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods*, 7, 127-150.

Banks, H. T., Hu, S., & Thompson, W. C. (2014). *Modeling and inverse problem in the presence of uncertainty*. Boca Raton, FL: Taylor and Francis Group.

Bartolucci F, Pandolfi S, Pennoni F (2017). LMest: An R Package for Latent Markov Models for Longitudinal Categorical Data. Journal of Statistical Software, 81(4), 1-38. doi: 10.18637/jss.v081.i04 (URL: http://doi.org/10.18637/jss.v081.i04).

Bekele-Maxwell, K., Everett, R. A., Shao, S., Kuerbis, A., Stephenson, L., Banks, H. T., & Morgenstern, J. (2017). Dynamical systems modeling to identify a cohort of problem drinkers with similar mechanisms of behavior change. *Journal for Person-Oriented Research, 3*(2), 100-118.

Capanu, M., Gönen, M., & Begg, C. B. (2013). An assessment of estimation methods for generalized linear mixed models with binary outcomes. *Statistics in Medicine*, *32*(26), 10.1002/sim.5866. http://doi.org/10.1002/sim.5866

Carney, M. A., Armeli, S., Tennen, H., Affleck, G., & O'Neil, T. P. (2000). Positive and negative daily events, perceived stress, and alcohol use: A diary study. *Journal of Consulting and Clinical Psychology, 68*(5), 788-798.

de Haan-Rietdijk, S., Kuppens, P., Bergeman, C.S., Sheeber, L.B., Allen, N.B., Hamaker, E.L. (2017). On the Use of Mixed Markov Models for Intensive Longitudinal Data. *Multivariate Behavioral Research, 52*(6), 747–767.

Donovan, D. M., Anton, R. F., Miller, W. R., Longabaugh, R., Hosking, J. D., & Youngblood, M. (2008). Combined pharmacotherapies and behavioral interventions for alcohol dependence (the COMBINE study): Examination of posttreatment drinking outcomes. *Journal of Studies on Alcohol and Drugs, 69*, 5-13.

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2012). Applied longitudinal analysis. (Vol. 998): John Wiley & Sons.

Ford, J. H., 2nd, Alagoz, E., Dinauer, S., Johnson, K. A., Pe-Romashko, K., & Gustafson, D. H. (2015). Successful Organizational Strategies to Sustain Use of A-CHESS: A Mobile Intervention for Individuals With Alcohol Use Disorders. *J Med Internet Res, 17*(8), e201. doi:10.2196/jmir.3965

Gustafson, D. H., McTavish, F. M., Chih, M. Y., Atwood, A. K., Johnson, R. A., Boyle, M. G., . . . Shah, D. (2014). A smartphone application to support recovery from alcoholism: a randomized clinical trial. *JAMA Psychiatry, 71*(5), 566-572. doi:10.1001/jamapsychiatry.2013.4642

Hallgren, K. A., Atkins, D. C., & Witkiewitz, K. (2016). Aggregating and analyzing daily drinking data in clinical trials: A comparison of type I errors, power, and bias. *Journal of Studies on Alcohol and Drugs, 77*(6), 986-991.

Jahnel, T., Ferguson, S. G., Shiffman, S., Thrul, J., & Schuz, B. (2017). Momentary smoking context as a mediator of the relationship between SES and smoking. *Addictive Behaviors*. doi:10.1016/j.addbeh.2017.12.014

Kelly, A. C., & Stephen, E. (2016). A daily diary study of self-compassion, body image, and eating behavior in female college students. *Body Image, 17*, 152-160. doi:10.1016/j.bodyim.2016.03.006

Kuerbis, A., Armeli, S., Muench, F., & Morgenstern, J. (2014). Profiles of confidence and commitment to change as predictors of moderated drinking: A person-centered approach. *Psychology of Addictive Behaviors, 28*(4), 1065-1076. doi:10.1037/a0036812

Liang, K., & Zeger, S. L. (1986). Longitudinal analysis using generalized linear models. *Biometrika, 73*(1), 13-22.

The MathWorks, Inc. (2015). MATLAB 2015a. Natick, MA.

Maruotti, A., & Rocci, R. (2012). A mixed non-homogenous hidden Markov model for categorical data, with application to alcohol consumption. *Statistics in Medicine 31*(9), 871-886.

Mereish, E. H., Kuerbis, A., & Morgenstern, J. (2018). A daily diary study of stressful and positive events, alcohol use, and addiction severity among heavy drinking sexual minority men. *Drug and Alcohol Dependence, 187*, 149-154. doi:https://doi.org/10.1016/j.drugalcdep.2018.03.003

Morgenstern, J., Kuerbis, A., Chen, A., Kahler, C. W., Bux, D. A., & Kranzler, H. (2012). A randomized clinical trial of naltrexone and behavioral therapy for problem drinking men-who-have-sex-with-men. *Journal of Consulting and Clinical Psychology, 80*(5), 863-875.

Morgenstern, J., Kuerbis, A., Houser, J., Muench, F. J., Shao, S., & Treloar, H. (2016). Within-person associations between daily motivation and self-efficacy and drinking among problem drinkers in treatment. *Psychology of Addictive Behaviors, 30*(6), 630-638. doi:10.1037/adb0000204

Morgenstern, J., Kuerbis, A., & Muench, F. (2014). Ecological momentary assessment and alcohol use disorder treatment. *Alcohol Research: Current Reviews, 36*(1), 101-109.

Morgenstern, J., Naqvi, N., DeBellis, R., & Breiter, H. (2013). The contributions of cognitive neuroscience and neuroimaging to understanding mechanisms of behavior change in addiction. *Psychology of Addictive Behaviors, 27*(2), 336-350.

National Institute on Alcohol Abuse and Alcoholism. (2013). *Rethinking drinking*. Retrieved from Bethesda, MD:

Quanbeck, A. R., Gustafson, D. H., Marsch, L. A., McTavish, F., Brown, R. T., Mares, M. L., . . . McDowell, H. (2014). Integrating addiction treatment into primary care using mobile health technology: protocol for an implementation research study. *Implement Sci, 9*, 65. doi:10.1186/1748-5908-9-65

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

SAS Institute Inc. (2002-2012). SAS software, Version 13.1 for Windows. Cary, NC: Author.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4*, 1-32. doi:10.1146/annurev.clinpsy.3.022806.091415

Shirley, K. E., Small, D. S., Lynch, K. G., Maisto, S. A., & Oslin, D. W. (2010). Hidden Markov models for alcoholism treatment trial data. *Annals of Applied Statistics, 4*(1), 366-395. doi:10.1214/09-AOAS282

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Sobell, L. C., Sobell, M. B., Leo, G. I., & Cancilla, A. (1988). Reliability of a timeline method: Assessing normal drinkers' reports of recent drinking and a comparative evaluation across several populations. *British Journal of Addiction, 83*(4), 393-402.

Spedicato, G.A. (2017) Discrete Time Markov Chains with R. The R Journal. R package version 0.6.9.7. (URL: https://journal.r-project.org/archive/2017/RJ-2017-036/index.html)

TELESAGE, I. (2005). *SmartQ 5.2 automated telephone survey software*. Chapel Hill, NC: Author.

Tennen, H., Affleck, G., Armeli, S., & Carney, M. A. (2000). A daily process approach to coping: Linking theory, research, and practice. *American Psychologist, 55*, 626-636.

Todd, M., Armeli, S., Tennen, H., Carney, M. A., Ball, S. A., Kranzler, H., & Affleck, G. (2005). Drinking to cope: A comparison of questionnaire and electronic dairy reports. *Journal of Studies on Alcohol, 66*, 121-129.

van Buuren, S. & Groothuis-Oudshoorn K. (2011). mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software, 45(3), 1-67. (URL: http://www.jstatsoft.org/v45/i03/)

Zeger, S. L., Liang, K. Y., & Albert, P. S. (1988). Models for longitudinal data: a generalized estimating equation approach. *Biometrics, 44*(4), 1049-1060.