Psychology of Addictive Behaviors

Pooled and Person-Specific Machine Learning Models for Predicting Future Alcohol Consumption, Craving, and Wanting to Drink: A Demonstration of Parallel Utility

Peter D. Soyster, Leighann Ashlock, and Aaron J. Fisher

Online First Publication, April 22, 2021. http://dx.doi.org/10.1037/adb0000666

CITATION

Soyster, P. D., Ashlock, L., & Fisher, A. J. (2021, April 22). Pooled and Person-Specific Machine Learning Models for Predicting Future Alcohol Consumption, Craving, and Wanting to Drink: A Demonstration of Parallel Utility. *Psychology of Addictive Behaviors*. Advance online publication. http://dx.doi.org/10.1037/adb0000666

AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychology of Addictive Behaviors

ISSN: 0893-164X

https://doi.org/10.1037/adb0000666

Pooled and Person-Specific Machine Learning Models for Predicting Future Alcohol Consumption, Craving, and Wanting to Drink: A Demonstration of Parallel Utility

Peter D. Soyster, Leighann Ashlock, and Aaron J. Fisher
Department of Psychology, University of California, Berkeley

OPEN DATA

Background and Aims: The specific factors driving alcohol consumption, craving, and wanting to drink, are likely different for different people. The present study sought to apply statistical classification methods to idiographic time series data in order to identify person-specific predictors of future drinking-relevant behavior, affect, and cognitions in a college student sample. ***Design:*** Participants were sent 8 mobile phone surveys per day for 15 days. Each survey assessed the number of drinks consumed since the previous survey, as well as positive affect, negative affect, alcohol craving, drinking expectancies, perceived alcohol consumption norms, impulsivity, and social and situational context. Each individual's data were split into training and testing sets, so that trained models could be validated using person-specific out-of-sample data. Elastic net regularization was used to select a subset of a set of 40 variables to be used to predict either alcohol consumption, craving, or wanting to drink, forward in time. ***Setting:*** A west-coast university. ***Participants:*** Thirty-three university students who had consumed alcohol in their lifetime. ***Measurements:*** Mobile phone surveys. ***Findings:*** Averaging across participants, accurate out-of-sample predictions of future drinking were made 76% of the time. For craving, the mean out-of-sample R^2 value was .27. For wanting to drink, the mean out-of-sample R^2 value was .27. ***Conclusion:*** Using a person-specific constellation of psychosocial and temporal variables, it may be possible to accurately predict drinking behavior, affect, and cognitions before they occur.

Public Health Significance Statement

This study utilized both person-specific and between-subjects approaches to predict future drinking, craving, and wanting to drink. Results indicated that both approaches performed well, but may have different applied utility.

Keywords: machine learning, prediction, alcohol, idiographic, time series

Supplemental materials: https://doi.org/10.1037/adb0000666.supp

Psychological and medical scientists have studied the causes, correlates, and consequences of alcohol use for more than 150 years. Given the broad demographics of those who use alcohol, and the wide-ranging consequence of alcohol use, there have been specific research subdomains related to alcohol consumption in men, women, and children for nearly as long (Madden, 1884). These literatures reveal invaluable contributions detailing topics as broad as the demographics of alcohol use, the biological mechanisms of alcohol metabolism, and the acute and chronic psychological effects of alcohol intoxication. Despite this attention, many fundamental questions in alcohol research remain unanswered. Perhaps chief among these is, *what causes someone to drink in a way that hurts their well-being, and how can they be helped to modify their drinking behavior*? Many potential mechanisms driving alcohol use have been identified, such as intentions to drink (Conner et al., 1999), personality characteristics like impulsivity (Stewart & Devine, 2000), and contextual factors like social norms (Kuntsche et al., 2006). Yet uncertainty remains regarding how to determine the relative influence of these mechanisms on actual consumption of alcohol for any given person. Methods for determining which of these mechanisms are relevant to a given population or individual are needed but are not yet optimized.

Peter D. Soyster https://orcid.org/0000-0003-4571-4106
Leighann Ashlock https://orcid.org/0000-0002-0506-0032
Aaron J. Fisher https://orcid.org/0000-0001-9754-4618

The data analyzed in this manuscript has been made publicly available through the Open Science Framework website. To our knowledge, these data have been included in one publication (Soyster et al., 2019; DOI: https://doi.org/10.17505/jpor.2019.06), which characterized survey compliance in a series of ecological momentary assessment studies.

The data are available at https://osf.io/q7upd/.

Correspondence concerning this article should be addressed to Peter D. Soyster, Department of Psychology, University of California, 2121 Berkeley Way, Berkeley, CA 94720, United States. Email: petersoyster@berkeley.edu

College Drinking

Studying the causes of alcohol use in college students is particularly important (Jennison, 2004). While higher education is a protective factor for many negative health behaviors (Collins et al., 2009), several findings suggest that college is a risk factor for hazardous alcohol use (Slutske, 2005; Turrisi et al., 2006). Researchers consider college students to be a distinct population, due to their higher risk for problematic drinking (Slutske, 2005). College students' unique drinking characteristics distinguish them from the general population. Moreover, they represent an early developmental stage of drinking behavior over the life course. Thus, they may provide an opportunity to examine specific mechanisms purported to lead to drinking in the general population—although it remains unclear if factors leading to alcohol use in college students occur similarly in other groups. For example, alcohol craving is widely regarded as a mechanism that leads to alcohol consumption (Iwanicka & Olajossy, 2015; Lowman et al., 2000; McHugh et al., 2016), yet some research has reported that due to the episodic nature of college drinking, situational factors like the day of the week likely play a more predominant role in consumption behavior (Kuntsche & Labhart, 2012). However, given a growing body of work in idiographic science, evidence points to the likelihood that, even within this population, the relationship between mechanisms such as craving and alcohol consumption will differ from person to person.

Idiographic and Nomothetic Approaches

Idiographic science is an area of human subjects research that places the individual at the center of hypotheses and analyses, rather than extrapolating person-level inferences from aggregated data (Fisher et al., 2018). Molenaar and Campbell (2009) have demonstrated that the statistical properties of mechanistic effects derived from between-subjects data cannot be presumed to represent the intra-individual mechanisms of any specific individual. Due to the heterogeneity of intra-individual processes, the influences of person-level mechanisms can often differ from person to person, creating different data generating processes across persons, ultimately resulting in distinct and differentiable statistical properties within each individual. However, this does not preclude the possibility that there are also genuine nomothetic data generating processes in the population. That is, whereas the particular influence of variables such as craving or delayed gratification may be idiosyncratic in their presentation and, thus, require idiographic methodologies to properly study, other phenomena may be more general and consistent from person to person. For instance, alcohol consumption could be reasonably expected to increase, generally, during evenings and weekends, or relatively uniformly across specific subpopulations during holidays or sporting events. Thus, there may be utility in pursuing parallel idiographic and nomothetic methods for predicting drinking behavior—in order to identify possible shared versus unique sources of influence on individuals.

For this reason, the present study pursued complementary analyses of idiographic and nomothetic data structures for predicting alcohol consumption, craving, and wanting to drink. Idiographic approaches allow the possibility that the data generating processes that drive individual drinking behavior will differ from person to person, and nomothetic approaches assume that pooling across individuals will produce generalizable predications that are applicable across participants. Importantly, to make such a comparison, study conceptualization and design must be predicated on the requirements of idiographic science. That is, because idiographic approaches require the collection of intensive repeated measures from individual participants in order to carry out unique, individualized analyses for each person, an idiographic approach obligates researchers to consider data density and temporality in ways that nomothetic analyses do not. To wit, individual data structures can be concatenated and aggregated to produce group-level data structures, whereas group-level data structure cannot be disaggregated to a level of granularity that does not exist in the primary data collection. From the idiographic perspective, participants ostensibly become a collection of N-of-1 studies, which can then be subsequently aggregated or compared. In the context of the present analysis, a nomothetic approach asks *what, generally, predicts when college students will drink alcohol*? Conversely, the idiographic approach asks *what predicts when this specific college student will drink alcohol*? As social science and medical fields work towards precision and personalized intervention models, we argue that the idiographic approach, if feasible, may provide information that is both superior and more granular for clinical applications. Nevertheless, this remains an empirical question, one that should be interrogated through direct comparison.

The Present Study

The present study used an ecological momentary assessment (EMA) paradigm to identify potential predictors of alcohol consumption, alcohol craving, and wanting to drink, in a sample of college students. We hypothesized that the precise mechanisms driving alcohol use, craving, and wanting to drink, were likely to differ markedly across participants, yielding unique prediction models with variable out-of-sample accuracy. However, we likewise acknowledged the possibility that common influences on drinking behavior may provide complimentary predictive utility, prompting the comparison of the two approaches across individuals. For fully idiographic predictions, we generated separate prediction models for each participant. For the nomothetic comparison, we pooled all participant data into a single prediction model. In both cases, prediction accuracy was tested on *person-specific* holdout data, assessing the degree to which idiographic or nomothetic models accurately predict future outcomes at the person level.

In producing both idiographic and nomothetic prediction models, the present study employed a set of procedures for defining a large set of potential predictors, utilizing variable selection methods for dimension reduction, and generating predictions via penalized regression models. Generating accurate and consistent predictions requires the target behavior to be subject to a systematic, relatively consistent data generating process (i.e., a set of contextual, interpersonal, and intrapersonal processes dictating the behavior). Thus, in order to generate models that can make accurate out-of-sample predictions, the data capture must measure the individual over time, and include the specific contextual variables that correspond to the behavior—including thoughts, emotions, physical sensations, and contextual features. Additionally, models should be able to represent the shape of the behavior's temporal variation (e.g., cyclic, based on time-of-day or day-of-the-week; Fisher & Bosley, 2020). For these reasons, we employed a bottom-up, data-driven approach to model construction, starting with the maximum number of potential variables (i.e., the feature space), and using variable selection procedures to reduce the total space to a subset of person-relevant features. In the present study we utilized the least absolute shrinkage and selection operator (LASSO; see, Section "Method"), however, it should be noted that other feature selection approaches exist in the literature. The approach employed in the present study facilitates parsimony in shaping model construction, while also reducing overfitting through regularization.

The present study presupposed that accurate prospective predictions would be derived from idiosyncratic combinations of temporal and psychosocial variables within individual participants (see Method). We hypothesized that consumption, craving and wanting to drink are systematic and predictable processes, which should be reflected in indices such as area under the curve (AUC), sensitivity, and specificity. Specifically, we hypothesized that we would observe a medium to large effect size in model prediction, based on converting model AUCs to Cohen's *D* (i.e. medium = .50, large = .80; Cohen, 2013). We predicted that a substantial proportion of variance in craving and wanting to drink (as indexed by R^2) would be returned in a majority of participant models. More precisely, we hypothesized a medium to large R^2 value for the craving and wanting to drink models (i.e. medium = .13, large = .26; Cohen, 2013). Finally, given the possibility that drinking behavior may be determined by both idiosyncratic and general data generating processes, we compared the performance of full idiographic predictions to a single, pooled prediction model, tested on a person-by-person basis on individual holdout data.

Method

Participants

Thirty-three participants completed the present study, with 19 providing sufficient data for prediction modeling of drinking events. Twenty-eight and 27 participants provided sufficient data for predictive modeling of craving and wanting to drink, respectively. In total, $N = 30$ unique participants were retained for modeling. Retained participants were adults ($M_{age} = 19$; $SD_{age} = 2.75$; range = 18–30) living in the San Francisco Bay Area who self-identified as having used alcohol in their lifetime. Twenty-seven participants (90%) identified as female and two identified as male, with one person who preferred not to disclose. Participants were diverse with respect to race/ethnicity (37% white, 33% Asian/Pacific Islander, 7% Hispanic or Latino, and 20% mixed or "other," with 3% who preferred not to disclose) and sexual orientation (80% heterosexual, 3% homosexual, 13% bisexual/queer, with 3% who preferred not to disclose). At baseline, participants reported consuming alcohol on an average of 1.67 days in the past week ($SD = 1.37$, range = 0–4) with an average of 4.23 ($SD = 2.42$, range = 1–12) drinks consumed on the average drinking day. Demographic information for retained and excluded participants is provided in Table 1.

Table 1
Participant Demographics

Variable	Retained participants (N = 30) N (% of sub sample)	Excluded participants (N = 3) N (% of sub sample)
Gender		
Male	2 (7%)	0 (0%)
Female	27 (90%)	3 (100%)
Did not disclose	1 (3%)	0 (0%)
Age		
Mean (*SD*)	19 (2.75)	21.33 (1.53)
Sexual orientation		
Heterosexual	24 (80%)	3 (100%)
Homosexual	1 (3%)	0 (0%)
Bisexual/queer	4 (13%)	0 (0%)
Did not disclose	1 (3%)	0 (0%)
Race/ethnicity		
White	11 (37%)	1 (33%)
Asian	10 (33%)	2 (66%)
Hispanic/Latino	2 (7%)	0 (0%)
Mixed race	6 (20%)	0 (0%)
Did not disclose	1 (3%)	0 (0%)
Mean drinking days per week		
Mean (*SD*)	1.67 (1.37)	1.33 (2.31)
Mean drinks per drinking occasion		
Mean (*SD*)	4.23 (2.42)	1.67 (1.53)
Age of first drink		
Mean (*SD*)	15.60 (2.04)	14.33 (4.16)

Procedure

All study procedures were approved by the University of California, Berkeley Committee for the Protection of Human Subjects and all participants provided informed consent prior to participation.

Participants were recruited through an undergraduate research participation pool. Interested individuals were directed to an online survey to screen for study eligibility. Inclusion required at least 18-years-of-age, English-language proficiency, regular access to a web-enabled mobile phone, having consumed alcohol (at least one sip) in their lifetime, and having had thoughts of consuming alcohol at least once in the previous month. Neither desire nor motivation to change drinking behavior was required to participate.

Eligible participants were invited to our lab at UC Berkeley for enrollment and baseline procedures. Participants completed a computer-based baseline assessment. Participants were then introduced to the EMA survey system and sampling protocol, and were trained to accurately estimate standard drink equivalents. Participants then completed EMA surveys eight-times-daily for 15 days. Surveys were sent on a semi-random schedule, randomly pushing a survey within 2 hr windows throughout the day. Notifications were sent to participants' mobile phones as hyperlinked text messages. EMA surveys asked participants to rate their current experience of each variable on a 0 (*not at all*) to 100 (*as much as possible*) visual analog slider. Participants who completed at least 80% of the mobile phone surveys were reimbursed with partial course credit.

Baseline Measures

Demographics

Demographic measures included participants' age, biological sex, gender identity, sexual orientation, racial/ethnic identity, alcohol use history, and psychological functioning.

EMA Survey Measures

Survey Procedure

Each survey assessed the number of standard drinks consumed since the previous survey, as well as positive affect, negative affect, alcohol craving, drinking expectancies, perceived alcohol consumption norms, impulsivity, and social and situational context. Table 2 contains the complete list of survey items and their wordings. The mean number of completed surveys across all participants ($N = 33$) was 107.6 ($SD = 25.6$), with a minimum of 11 and a maximum of 129. The mean number of completed surveys across retained participants ($N = 30$) was 114 ($SD = 12.22$), with a minimum of 79 and a maximum of 129.

Data Preparation and Analysis

Each participant's EMA data was prepared for analysis separately. First, survey time stamps were used to calculate continuous time since the first EMA survey. Next, we ensured that the data frame had a uniform structure that reflected eight surveys per day, starting with the first morning survey. Next, we created variables for survey number (i.e., 1st, 2nd, etc. survey of the day), day of the week, and cycles of 12 hr, 24 hr, and weekly frequency. The latter variables were constructed from sine and cosine terms, consistent with methods provided by Flury and Levri (1999).

Table 2
Possibility Space for All Prediction Models

Items assessed through EMA
Since your last survey; how many alcoholic drinks have you had?
I feel comfortable in my current location
I feel stressed
I feel down/depressed
I feel calm/relaxed
I currently feel pressure to drink
I feel enthusiastic
I feel happy
I am having conflict/fighting with others
I am craving alcohol
I am feeling impulsive
A drink would make me feel better right now
What % of [university] students do you think are drinking alcohol right now?
I would like to drink
I feel able to delay gratification
I feel angry
Variables derived from time
Linear time
Quadratic time
Cubic time
12 hr cosine
12 hr sine
24 hr cosine
24 hr sine
Weekly cosine
Weekly sine
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Survey 1 of the day
Survey 2 of the day
Survey 3 of the day
Survey 4 of the day
Survey 5 of the day
Survey 6 of the day
Survey 7 of the day
Survey 8 of the day

For the models of alcohol consumption, participants' data were assessed for the frequency and distribution of drinking events. Participants were omitted from analyses for insufficient frequency of survey responses ($N = 3$) or for reporting so few drinking occasions during the survey period that k-fold cross-validation could not be run ($N = 11$). Drinking data were then transformed such that each time point represented whether or not any drinking had occurred since the last time point. For the craving and wanting to drink models, participants' data were assessed for variability in reported craving and wanting to drink. Participants were omitted from these analyses for insufficient frequency of survey responses ($N = 3$) or for having insufficient variation in reported craving or wanting to drink (e.g., reporting zero craving at nearly every time point; $N = 2$ in craving model, $N = 2$ in wanting to drink model).

Finally, all temporal variables and the dichotomous drinking variable were set forward in time by one observation. The resulting data frame provided row-wise relationships wherein all temporal variables were aligned in time with the occurrence of drinking events and all remaining predictor variables were lagged predictors. This data structure allowed for k-fold cross-validation of a time series without violating the contiguity of the data. All prediction models in the present analysis were lagged models—predicting drinking events, craving, and wanting to drink roughly two hours in the future. Table 2 presents the complete feature space for all prediction models (number of possible variables = 40).

Constructing Prediction Models

Each participant's time series was evenly split into two data sets to be used for model training and testing. The specific observations included in the training and testing sets were chosen randomly and were not required to be temporally contiguous. Following the division of the time series, elastic net regularization was used to select independent variables for idiographic prediction models.

Elastic Net Regularization

Elastic net regularization (Zou & Hastie, 2005) was used for variable selection in all training models. It should be noted that a variety of other techniques exist for classification (e.g., random forests, support vector machines, gradient boosting machines) and variable selection (e.g., random forest variable importance, penalized SVM), and competing methods could, ostensibly, provide stronger predicted performance than the elastic net algorithm. However, elastic net has several characteristics which make it an attractive approach for the present aims. Elastic net is a regularized regression technique that produces sparse models through coefficient penalization. Beyond protecting against model overfitting, such penalization is useful for variable selection—coefficients shrunk to zero are removed from the model. Given the ubiquity and accessibility of regression-based methods in social sciences, the elastic net thus represents a sensible bridge between classical and machine learning approaches. To complete these analyses, we used the *cv.glmnet* function in the *glmnet* package (Friedman et al., 2010) in *R*. Elastic net integrates L_1 and L_2 penalization (i.e. LASSO and ridge regression) via the alpha parameter. Alpha varies from 0 (*exclusively ridge regression*) and 1 (*exclusively LASSO* *regression*), with .50 representing an even balance between the two. All models were initially run at alpha =.50. In the case that no variables were retained, alpha was decreased incrementally by values of .05 until at least one nonzero coefficient was retained. A *k*-fold cross-validation with 10 folds was used to select the optimal model. The lambda threshold with the minimum mean cross-validated error was used to select the final model for each participant.

Prediction Models

Following the elastic net procedure, retained variables were used as independent variables in the prediction models. We constructed a binary logistic regression model for each participant, using the training data set. The resulting coefficients were then used to predict drinking events in the testing data set. Following this, we used the *pROC* function in the *pROC* package (Robin et al., 2011) in *R* to determine AUC, sensitivity, specificity, and a brier score for predictions of each participant's future drinking events in the testing data set. Brier score represents the mean squared difference between an observed binary outcome (e.g., drank/did not drink) and the model estimated probability of that outcome occurring, with lower scores indicating stronger predictive performance (Wilks, 2010). To calculate effect size, we then used established methods for converting AUC to Cohen's *D* (Salgado, 2018). We employed a similar procedure for the craving and wanting to drink models, using linear regression models to make predictions of future ratings of the dependent variables, and calculating R^2 values to estimate model fit in the testing data set.

To increase the utility of study results to generate hypotheses for future research, we conducted exploratory post-hoc analyses to assess whether participant characteristics could be used to predict model performance. In a series of bivariate linear regressions, we regressed AUC, sensitivity, specificity, R^2_{craving}, and R^2_{wanting} onto participant age, gender, sexual orientation, race/ethnicity, mean drinking days per week, mean drinks per drinking occasion, number of EMA surveys completed, number of drinking occasions during EMA period, and whether in the previous year they met abuse criteria for cannabis, hallucinogens, stimulants, or polysubstance abuse.

Pooled Prediction Models

In order to test the relative accuracy of nomothetic predictions, we concatenated the time series for all 33 participants in order to generate a single aggregate data frame. Across human subjects research in medicine and the social sciences, multilevel modeling (MLM) is the most common method for handling repeated measures data. However, support for MLM data structures is not commonly found in machine learning methods, and no MLM alternatives currently exist for LASSO or elastic net procedures. Thus, we took a two-step approach to pursuing pooled analyses. It has been well-demonstrated that pooling across repeated measures can bias data analyses due to the nesting of repeated measures within individuals. The variance attributable to repeated measures can often be accounted for via the random intercept in a MLM. Thus, in the first step of our pooled approach, we person-mean-centered each variable for each participant to account for nesting within persons across multiple measurements. Next, we created a lag-1 structure within each individual data set prior to concatenation, in order to ensure that no between-subject lags existed in the aggregated data. We then concatenated the 33 Lagged, person-mean-centered time series into a long-format nomothetic (i.e. aggregated) data frame. The elastic net procedures outlined above were then applied to the pooled data. The aggregated model was produced with alpha = .50. The fixed-effect coefficients from this model were then applied to the out-of-sample holdout data for each individual to predict drinking, craving, and wanting to drink, person by person.

Finally, due to the relative novelty of our pooled approach—and in order to compare our results to the more ubiquitous MLM approach—we ran a second pooled-data model within an MLM framework. Using the 24 variables selected by the elastic net procedure, we ran a MLM in the *lme4* (Bates et al., 2015) package in *R*, using the *glmer* function for consumption and the *lmer* function for craving and wanting to drink. Because it is possible that some bias and inflation of model performance may have still been possible under the person-mean-centered conditions, we wanted to compare these results to an industry-standard approach with a random intercept.

Results

Complete data, analysis code, and results for all participants are available on the Open Science Framework at https://osf.io/q7upd/.

Elastic Net Regularization

Dichotomous (Drank/Did Not Drink) Predictions

The results of all 19 person-specific dichotomous prediction models are presented in Table 3. The mean out-of-sample AUC was .76 (*SD* = .16, range = .53–.98, Cohen's *D* = .99). Out-of-sample AUC was not normally distributed. Median out-of-sample AUC was .78. The mean sensitivity was .81 (*SD* = .21; range = .33–1.00), and the mean specificity was .78 (*SD* = .24; range = .30–1.00). The mean Brier score was .097 (*SD* = .038; range = .045–.186). Together, these findings indicate that models in the present study were 76% accurate on average, with an average of 81% accuracy for predicting presence and 78% accuracy for predicting absence of a drinking episode.

Continuous Predictions for "Craving" and "Wanting to Drink"

The results for all person-specific models are presented in Table 4. For craving, the mean out-of-sample R^2 value was .27 (*SD* = .20; range = .00–.69). For wanting to drink, the mean out-of-sample R^2 value was also .27 (*SD* = .21; range = .00–.67).

Frequencies of variable retention by person-specific model type are presented in Table 5. The median number of retained variables for the idiographic drinking models was 11. The median number of retained variables for the idiographic craving and wanting to drink models was 5 and 8.5, respectively.

Pooled Prediction Models

The results of all pooled prediction models are presented in Table 6. These results reflect the accuracy of the fixed-effect coefficients of the pooled models in predicting future outcomes in the out-of-sample idiographic time series.

The pooled elastic net regression model retained 24 variables to predict drinking events. The mean out-of-sample AUC for

Table 3
Person-Specific Dichotomous (Drank/Did Not Drink) Prediction Model Results

ID	AUC	Sensitivity	Specificity	Brier	No. of retained variables	Retained variables	No. of retained self-report variables
1	.57	1.00	.30	.15	24	1,3,4,5,8,9,11,12,13,14,15,18,20,21,23,29,31,32,33,34,35,38,39,40	11
2	.78	.83	.81	.09	8	5,9,11,16,18,21,32,34	3
3	.98	1.00	.95	.07	8	5,9,11,12,16,17,18,36	4
4	.56	.75	.54	.08	8	4,5,11,12,13,31,36,39	5
5	.79	.75	.98	.06	12	4,5,8,9,13,16,23,28,32,37,39,40	5
6	.96	1.00	.90	.05	12	2,4,7,9,14,15,16,17,29,35,36,27	6
7	.94	1.00	.86	.06	16	2,3,6,7,12,14,16,17,20,22,23,29,30,33,35,37	6
9	.90	.75	.97	.12	17	3,7,9,10,11,13,14,15,16,17, 21,23,28,30,32,36,38	8
10	.66	1.00	.33	.11	14	4,8,9,11,12,13,15,28,29,30,32,33,37,40	7
12	.94	1.00	.88	.09	10	1,3,7,9,11,12,13,25,32,34	7
14	.72	.71	.90	.09	14	4,5,11,13,14,16,17,29,30,32,36,37,38,39	5
15	.53	.80	.42	.16	4	7,14,32,33	2
17	.68	.40	.97	.19	1	19	0
18	.83	.75	1.00	.05	10	5,10,12,13,15,23,31,33,35,38	5
21	.92	1.00	.76	.05	17	5,8,10,11,12,15,16,20,21,26,27,28,29,30,36,38,40	6
23	.74	.56	.91	.11	9	5,9,13,19,29,30,32,34,37	3
24	.58	.33	.96	.11	11	2,4,7,16,19,20,29,33,36,38,40	3
29	.87	.71	.90	.08	13	1,3,9,12,21,25,26,29,32,33,34,38,39	4
32	.53	1.00	.41	.15	11	5,6,8,15,20,26,28,32,34,35,38	4

Note. 1 = Comfortable, 2 = stressed 3 = down/depressed, 4 = calm, 5 = pressure to drink, 6 = enthusiastic, 7 = happy, 8 = experiencing conflict, 9 = craving, 10 = impulsive, 11 = positive alcohol expectancy, 12 = perception of drinking norms, 13 = want to drink,14 = able to delay gratification, 15 = angry, 16 = drank/did not drink, 17 = Monday, 18 = Tuesday, 19 = Wednesday, 20 = Thursday, 21 = Friday, 22 = Saturday, 23 = Sunday, 24 = survey 1, 25 = survey 2, 26 = survey 3, 27 = survey 4, 28 = survey 5, 29 = survey 6, 30 = survey 7, 31 = survey 8, 32 = linear time, 33 = quadratic time, 34 = cubic time, 35 = 12 hr cosine, 36 = 12 hr sine, 37 = 24 hr cosine, 38 = 24 hr sine, 39 = weekly cosine, 40 = weekly sine.

person-level predictions from the pooled elastic net prediction coefficients was .78 (SD = .14, range = .55–.99, Cohen's D = 1.05). Out-of-sample AUC was not normally distributed. Median out-of-sample AUC was .81. The mean sensitivity was .85 (SD = .19; range = .30–1.00), and the mean specificity was .81 (SD = .18; range = .35–1.00). The mean Brier score was .111 (SD = .169; range = .024–.933). Together, these findings indicate that the person-level tests of the pooled elastic net predictors were 78% accurate on average, with an average of 85% accuracy for predicting presence and 81% accuracy for predicting absence of a drinking episode. The pooled elastic net models retained 15 and 23 variables to predict craving and wanting to drink, respectively. For craving, the mean out-of-sample R^2 value was .31 (SD = .22; range = .00–.73). For wanting to drink, the mean out-of-sample R^2 value was also .33 (SD = .22; range = .00–.87).

As noted above, MLM models were run using the predictors selected by the elastic net models, in order to test the potential effect of utilizing a random intercept on model performance. Results for all MLM models are presented in Table 7. The mean out-of-sample AUC for person-level predictions from MLM prediction coefficients was.77 (SD = .14, range = .49–1.00, Cohen's D = 1.05). Out-of-sample AUC was not normally distributed. Median out-of-sample AUC was .81. The mean sensitivity was .82 (SD = .17; range = .50–1.00), and the mean specificity was .80 (SD = .17; range = .30–1.00). The mean Brier score was .075 (SD = .053; range = .003–.266). Taken together, MLM results were comparable—in fact, nearly identical—to the results of pooled elastic net models.

A series of repeated measures t-tests indicated that out-of-sample tests exhibited significantly better R^2_{craving} ($M_{\text{difference}} = -.44$, $t[27] = -4.30$, $p < .001$), and R^2_{wanting} ($M_{\text{difference}} = -.28$, $t[26] = -2.68$, $p = .01$), based on the predictions generated from the single, pooled model. There were no significant differences between the idiographic and pooled predictions for AUC, sensitivity, specificity, or Brier score. We found a similar pattern of results for the out-of-sample MLM models. Out-of-sample R^2_{craving} ($M_{\text{difference}} = -.05$, $t[27] = -2.16$, $p = .04$), and R^2_{wanting} ($M_{\text{difference}} = -.05$, $t[26] = -2.40$, $p = .02$) were significantly better based on MLM predictions.

Exploratory Analyses

Baseline variables did not significantly predict person-specific model AUC, sensitivity, or specificity. However, person-specific Brier scores were significantly lower for participants who identified as male when compared to those who identified as female ($b = -.059$, $t[16] = -2.25$, $p = .04$) and R^2_{craving} was significantly larger for participants who met criteria for hallucinogen ($b = .24$, $t[26] = 2.16$, $p = .04$), stimulant ($b = .33$, $t[26] = 3.74$, $p < .001$), and polysubstance abuse ($b = .20$, $t[26] = 2.33$, $p = .03$). Person-specific R^2_{wanting} was significantly larger for participants who had greater mean drinking days per week ($b = .06$, $t[25] = 2.31$, $p = .03$), and those who met criteria for hallucinogen ($b = .26$, $t[25] = 2.22$, $p = .04$) and stimulant abuse ($b = .26$, $t[25] = 2.57$, $p = .02$).

Baseline variables did not significantly predict pooled model AUC. Pooled specificity was significantly larger ($b = .01$, $t[25] = 3.00$, $p = .01$) and pooled sensitivity was significantly smaller ($b = -.01$, $t[25] = -2.55$, $p = .02$) for participants who

Table 4
Person-Specific R² *Values for 'Craving' and 'Want to Drink' Models*

ID	R^2 craving	No. of retained variables	Retained variables	No. of retained self-report variables	R^2 for Want to Drink	No. of retained variables	Retained variables	No. of retained self-report variables
1	.10	4	11,12,14,29	3	.12	6	3,5,11,12,14,33	5
2	.29	12	1,3,5,11,12,13,16,17,20,21,27,40	6	.38	7	1,9,11,12,13,22,32	5
3	.19	8	1,8,9,11,13,19,39,40	5	.15	9	9,11,13,14,18,21,23,28,40	4
4	.01	1	13	1	.05	10	8,12,13,18,21,29,30,31,36,39	3
5	.47	12	3,5,6,11,13,17,21,30,33,35,37,39	5	.60	7	3,6,11,13,21,33,39	4
6	.28	7	6,9,28,33,34,36,37	2	.23	9	2,3,13,19,21,25,28,36,38	3
7	.40	6	6,9,11,12,13,36	5	.34	1	13	1
8	.51	4	8,9,12,13	4	.53	10	2,9,10,12,13,22,23,26,30,38	5
9	.51	5	8,9,13,31,36	3	.42	3	9,11,13	3
10	.20	1	13	1	.47	4	9,11,13,36	3
12	.44	6	13,27,30,34,36,37	1	.52	17	1,2,6,7,10,11,13,17,18,21,25,27, 29,30,36,37,39	7
13	NA	NA	NA	NA	.00	39	1,2,3,4,5,6,7,8,9,10, 11,12,13,14,15,16, 17,18,19,20,21,22,23,25,26,27,28,29,30,31, 32,33,34,35,36,37,38,39,40	15
14	.69	8	5,11,13,14,27,30,24,40	4	.50	5	11,13,29,30,31	2
15	.00	5	9,11,13,30,40	3	.03	5	9,11,13,32,33	3
16	.02	7	1,2,9,11,28,32,39	4	.19	1	11	1
17	.17	3	5,15,29	2	NA	NA	NA	NA
18	.00	2	21,32	0	.09	9	1,5,16,19,20,29,32,33,34	2
19	.09	17	6,7,8,10,11,13,17,19,20,21, 22,26,29,31,32,37,38	6	.13	1	24	0
20	.51	7	3,5,8,9,13,15,32	6	.67	8	4,6,9,12,13,30,38,40	5
21	.57	12	4,5,6,12,13,15,16,23,28,31,39,40	6	.54	19	1,2,3,5,6,9,11,12,13,14,15,17, 21,23,25,28,33,35,39	11
22	.20	8	2,9,11,12,13,15,22,39	6	.17	9	6,9,10,11,13,15,20,21,34	6
23	.43	3	11,29,32	1	.42	2	9,13	2
24	.19	3	9,29,32	1	.02	21	3,4,5,7,9,10,12,14,15,18,19,20,22,25,26, 28,30,32,36,38,40	9
25	.00	3	9,12,31	2	.16	2	12,31	1
26	.25	5	6,9,13,29,34	3	.09	8	4,5,6,7,9,13,22,34	6
27	.46	1	13	1	.08	18	1,2,4,10,13,14,15,17,19,20,21,22, 23,28,29,33,38,40	7
28	NA	NA	NA	NA	NA	NA	NA	NA
29	.31	5	1,7,11,20,28	3	.42	8	1,5,7,11,19,21,28,39	4
30	.17	5	11,15,18,32,40	2	.15	14	1,2,6,8,9,12,13,18,22,28,29,30,31,40	7
32	.12	2	12,30	1	.01	13	1,6,9,12,13,18,19,21,22,26,28,29,34	5

Note. 1 = comfortable, 2 = stressed, 3 = down/depressed, 4 = calm, 5 = pressure to drink, 6 = enthusiastic, 7 = happy, 8 = experiencing conflict, 9 = craving, 10 = impulsive, 11 = positive alcohol expectancy, 12 = perception of drinking norms, 13 = want to drink, 14 = able to delay gratification, 15 = angry, 16 = drank/did not drink, 17 = Monday, 18 = Tuesday, 19 = Wednesday, 20 = Thursday, 21 = Friday, 22 = Saturday, 23 = Sunday, 24 = survey 1, 25 = survey 2, 26 = survey 3, 27 = survey 4, 28 = survey 5, 29 = survey 6, 30 = survey 7, 31 = survey 8, 32 = linear time, 33 = quadratic time, 34 = cubic time, 35 = 12 hr cosine, 36 = 12 hr sine, 37 = 24 hr cosine, 38 = 24 hr sine, 39 = weekly cosine, 40 = weekly sine.

completed a greater number of EMA surveys. Pooled Brier scores were significantly lower for participants who completed a greater number of EMA surveys ($b = -.005$, $t[26] = -2.15$, $p = .04$) and significantly higher for those who had a larger number of average drinks per drinking occasion ($b = .04$, $t[26] = 3.32$, $p = .002$), and those who met criteria for hallucinogen ($b = .207$, $t[26] = 2.48$, $p = .02$), stimulant ($b = .194$, $t[26] = 2.57$, $p = .02$), or polysubstance abuse ($b = .151$, $t[26] = 2.20$, $p = .04$). Pooled $R^2_{craving}$ was significantly larger for those who met criteria for cannabis ($b = .17$, $t[28] = 2.25$, $p = .03$) or hallucinogen abuse ($b = .23$, $t[28] = 2.11$, $p = .04$). $R^2_{wanting}$ was significantly larger for those who had more drinking days per week, on average ($b = .06$, $t[28] = 2.24$, $p = .03$).

Baseline variables did not significantly predict MLM model AUC, specificity, sensitivity, $R^2_{craving}$, or $R^2_{wanting}$. MLM Brier score was significantly larger of participants who reported more drinking occasions during the EMA period ($b = .004$, $t[28] = 4.68$, $p < .001$).

The same set of variables were used to investigate if there were significant patterns in the groups of individuals who were better predicted by the single pooled prediction or MLM models versus the multiple person-specific models. When comparing the person-specific models to the pooled estimate and MLM models, none of the variables were significantly associated with having

Table 5
Frequencies of Variable Retention by Person-Specific Model Type

Variable	Drinking	Craving	Want to drink
Comfortable	3	4	10
Stressed	3	3	8
Down/depressed	5	3	6
Calm	7	1	5
Pressure to drink	10	7	8
Enthusiastic	2	6	9
Happy	6	2	5
Conflict	5	5	3
Craving	10	13	15
Impulsive	3	2	7
Positive expectancy	9	13	15
Drinking norms	9	8	12
Want to drink	9	17	21
Delay gratification	6	2	6
Angry	7	6	6
Drink (yes/no)	9	2	2
Monday	5	3	4
Tuesday	3	1	7
Wednesday	3	2	7
Thursday	5	3	3
Friday	5	4	11
Saturday	1	3	10
Sunday	5	1	6
Survey 1	0	0	2
Survey 2	2	0	5
Survey 3	3	1	4
Survey 4	1	3	3
Survey 5	5	4	9
Survey 6	9	6	8
Survey 7	6	5	8
Survey 8	3	5	6
Linear time	11	7	5
Quadratic time	7	2	7
Cubic time	6	5	5
12 hr cosine	5	1	2
12 hr sine	7	5	6
24 hr cosine	6	5	4
24-sine	8	2	6
Weekly cosine	5	6	7
Weekly sine	5	6	6

better pooled or MLM predictions for AUC, sensitivity, specificity, Brier score, R^2_{craving}, or R^2_{wanting}.

Discussion

The present study sought to identify relative sets of person-specific and pooled predictors of alcohol consumption, alcohol craving, and wanting to drink in a sample of college students. Additionally, we aimed to assess the ability of prediction models comprised of psychosocial and temporal variables to accurately predict future consumption, craving, and wanting to drink. At the idiographic level of analysis, our model-building procedures returned a unique number and combination of retained variables and coefficients for each model for each participant. Despite this variability—or perhaps because of it (Fisher et al., 2018; Molenaar, 2004; Soyster & Fisher, 2019)—we observed strong sample-wide prediction accuracy for future drinking occasions. Slightly exceeding our expectations, a mean AUC = .76 corresponded to a Cohen's $D = .99$, indicating a large effect size (Cohen, 2013). Based on each participant's unique subset of variables selected from our feature space (see Table 2), accurate predictions of future drinking behavior could be made 76% of the time (on average) in the present models. Further, our exploratory post-hoc analyses indicated that the modeling approaches we present for predicting alcohol-relevant affect and cognitions may be particularly well-suited to individuals who meet criteria for other substance abuse. It may be the case that in the context of more intensive substance use (contrasted against less intense substance use), such methods provide increased ability to accurately predict craving and wanting to drink.

However, the present study also included two sets of analyses based on aggregated data—one based on machine learning analysis of person-mean-centered data and a multilevel model (MLM) utilizing the predictors selected by the elastic net procedure. These models allowed us to contrast our findings against a more traditional,

Table 6
Pooled Elastic Net Model Results

ID	AUC	Sensitivity	Specificity	Brier	R^2 craving	R^2 want to drink
1	.62	.50	.86	.125	.18	.10
2	.68	.83	.60	.086	.31	.44
3	.96	1.00	.93	.070	.27	.32
4	.55	.75	.54	.090	.02	.17
5	.82	.75	.98	.046	.53	.50
6	.99	1.00	.98	.046	.34	.32
7	.63	1.00	.35	.076	.60	.59
8	.94	1.00	.89	.037	.60	.57
9	.83	.80	.81	.208	.53	.52
10	.79	.83	.83	.102	.28	.48
12	.90	.88	.88	.076	.45	.56
13	.94	1.00	.94	.029	.11	.28
14	.80	.71	.95	.078	.73	.55
15	.56	.30	.93	.154	.00	.00
16	NA	NA	NA	NA	.11	.29
17	.55	NA	NA	.225	.40	.49
18	.81	.75	1.00	.038	.08	.08
19	.93	1.00	.90	.041	.17	.12
20	.78	1.00	.56	.049	.72	.87
21	.92	1.00	.76	.058	.45	.46
22	.63	.67	.88	.051	.24	.32
23	.69	.56	.95	.112	.42	.32
24	.78	1.00	.58	.101	.34	.28
25	.85	.67	1.00	.055	.00	.26
26	NA	NA	NA	NA	.19	.03
27	.85	1.00	.85	.024	.14	.07
28	.79	1.00	.64	.933	.17	.07
29	.82	.86	.69	.095	.63	.61
30	.95	1.00	.95	.031	.34	.15
32	.61	1.00	.58	.072	.04	.09

Note. For drinking prediction model, retained variables = Stressed, down/depressed, pressure to drink, happy, conflict, craving, impulsive, positive alcohol expectancy, drinking norms, want to drink, angry, drank/did not drink, Tuesday, Wednesday, Friday, Saturday, survey 2, survey 6, survey 7, survey 8, 24 hr cosine, 24 hr sine, weekly cosine, weekly sine. For craving model, retained variables = pressure to drink, enthusiastic, conflict, craving, positive alcohol expectancy, drinking norms, want to drink, drank/did not drink, Friday, Saturday, survey 3, survey 6, survey 7, 12 hr sine, 24 hr sine. For wanting to drink model, retained variables = stressed, pressure to drink, enthusiastic, conflict, craving, positive alcohol expectancy, drinking norms, want to drink, delay gratification, angry, drank/did not drink, Monday, Tuesday, Friday, Saturday, Sunday, survey 2, survey 6, survey 7, survey 8, cubic time, 24 hr sine, weekly sine.

Table 7
Multilevel Model Results

ID	AUC	Sensitivity	Specificity	Brier	R^2 craving	R^2 want to drink
1	.63	.50	.84	.127	.17	.09
2	.69	.83	.56	.086	.32	.44
3	.97	1.00	.95	.057	.30	.32
4	.54	.75	.61	.089	.04	.19
5	.79	.75	.96	.052	.53	.51
6	1.00	1.00	1.00	.036	.31	.32
7	.53	.67	.68	.081	.61	.60
8	.94	1.00	.89	.037	.62	.58
9	.84	.80	.81	.156	.55	.54
10	.78	.83	.83	.102	.30	.43
12	.92	.88	.88	.080	.45	.55
13	.89	1.00	.89	.025	.11	.30
14	.84	.71	.98	.073	.66	.51
15	.57	.60	.65	.162	.00	.00
16	NA	.63	.85	.003	.10	.29
17	.49	.90	.30	.266	.41	.50
18	.79	.75	1.00	.038	.09	.10
19	.82	1.00	.68	.043	.13	.11
20	.82	1.00	.64	.049	.73	.83
21	.90	1.00	.71	.061	.46	.46
22	.63	.67	.88	.048	.24	.33
23	.73	.56	.98	.118	.35	.27
24	.78	1.00	.58	.108	.35	.26
25	.82	.67	.97	.058	.00	.20
26	NA	.63	.85	.010	.21	.02
27	.87	1.00	.87	.019	.12	.04
28	.82	1.00	.75	.071	.00	.00
29	.80	.57	.92	.096	.62	.60
30	.95	1.00	.95	.031	.28	.11
32	.54	1.00	.47	.080	.04	.11

Note. For drinking prediction model, retained variables = stressed, down/depressed, pressure to drink, happy, conflict, craving, impulsive, positive alcohol expectancy, drinking norms, want to drink, angry, drank/did not drink, Tuesday, Wednesday, Friday, Saturday, survey 2, survey 6, survey 7, survey 8, 24 hr cosine, 24 hr sine, weekly cosine, weekly sine. For craving model, retained variables = Pressure to drink, enthusiastic, conflict, craving, positive alcohol expectancy, drinking norms, want to drink, drank/did not drink, Friday, Saturday, survey 3, survey 6, survey 7, 12 hr sine, 24 hr sine. For wanting to drink model, retained variables = stressed, pressure to drink, enthusiastic, conflict, craving, positive alcohol expectancy, drinking norms, want to drink, delay gratification, angry, drank/did not drink, Monday, Tuesday, Friday, Saturday, Sunday, survey 2, survey 6, survey 7, survey 8, cubic time, 24 hr sine, weekly sine.

pooled approach to data analysis. Within this nomothetic approach, participant time series were pooled into a single data set, producing a single set of coefficients for all participants. Thus, we pooled across individuals in generating singular predictions for drinking events, craving, and wanting to drink, from the elastic net and MLM, respectively. For each dependent variable, we used the single set of coefficients from the pooled model to make predictions in the out-of-sample time series for each individual. Although we hypothesized that the person-specific models would provide strong AUC, sensitivity, specificity, and R^2, due to their ability to better-learn the idiosyncratic features of each individual's data, it should be emphasized that pooling across participants has the advantage of borrowing information from the aggregated sample in places where the person-level data may be missing or incomplete. Consistent with this idea, we were able to generate prediction models for individuals who were excluded from idiographic modeling due to insufficient data. Moreover, whereas idiographic approaches emphasize idiosyncratic, person-level drivers of drinking behavior, nomothetic models capture general (and generalizable) data generating processes that may exist across individuals in a population.

Importantly, the pooled approach yielded equivalent results for predicting future drinking events, with mean out-of-sample predictions greater than, but not statistically significant from the person-specific models. However, the pooled model and MLM both did a significantly better job in predicting future ratings for craving and wanting to drink. This pattern of results indicated that for craving and wanting to drink, the shared information between participants was concordant to the extent that stronger predictions were achieved. This may be due to the relative homogeneity of our sample; all of our participants were undergraduate students from the same university and were largely female and heterosexual. Our small between-subjects sample size prevents us from speculating about the extent to which the retained variables and coefficients from the pooled model and MLM may generalize to the broader population of college students. Results were consistent for both person-mean-centered elastic net and MLM, indicating that the elastic net models did not necessarily benefit from a lack of random effects. It is important to emphasize that the predictions in question were still generated from idiographic time series—data held out from each of the 33 participants' EMA series and tested on a person-by-person basis. Nevertheless, the pooled data were, in a single set of coefficients, able to predict unseen outcomes as well as or better than the more finely-tuned, idiosyncratic coefficients (which differed in number and strength from person to person). Future research should examine conditions under which pooled or independent data sources produce better predictions. Likely due to our small between-subjects sample size, we did not observe and significant relationships between being better-predicted by the person-specific or between-subjects models. In possible future clinical applications, person-level data remains the most plausible and practical application of the present methodology, allowing a single clinician to work with a single patient without the need for pooling or aggregating other patients. However, in research paradigms in which both idiographic and nomothetic methods are equally practical, it may be useful to alternate approaches based on specific person-level or contextual variables.

Regarding possible clinical applications of the current study, it is important to recall that these models utilized a set of time-lagged independent variables, validated using out-of-sample holdouts from each person's time series data. Thus, the reported predictive accuracy and variance accounted for in each dependent variable pertains to unseen data, approximately two hours in the future. Specifically, person-specific models were able to accurately classify an average of 76% of time points as either drinking or non-drinking events approximately two hours before they occurred, and we returned equally promising results in the variance accounted for in the continuous ratings of *craving* and *wanting to drink*. Ostensibly, these models could be applied to ongoing monitoring and early warning systems, allowing clinicians to identify optimal inflection points for the delivery of intervention materials.

Thus, the present methodology holds potential promise for personalizing substance use treatment either by supporting just-in-time mobile interventions, or by augmenting existing therapist-delivered psychosocial interventions. Both the idiographic and aggregated methods in the present paper could provide clinicians with insights into the timing of craving and consumption behavior, identifying times in which the individual is most at risk for drinking. Additionally, the idiographic approach in the current study could facilitate the construction of personalized interventions by identifying potentially idiosyncratic mechanisms of craving, wanting, and consumption. In concert, these data could be used to identify *when* intervention is needed and help guide *which* intervention components may be most effective.

Despite the surprising success of the pooled models in the present study, it is crucial to emphasize the importance of sufficient person-level data in looking ahead to possible clinical and real-time applications of the current methodology. That is, these were intensively measured data, captured during the course of each individual's day-to-day life in situ, producing multivariate time series of sufficient length to allow division into training and testing sets.

Limitations and Future Directions

Generalizability

One possible limitation of the present study is the external validity of the analyses and the generalizability to college-aged drinkers writ large. Specifically, one might argue that idiographic models generalize only to the individual under analysis, and not to other individuals. This is accurate. It is not our contention that idiographic *models* generalize. Such a position would be a violation of the ergodic theorem (see, Fisher et al., 2018; Molenaar, 2004). Instead, we contend that idiographic *methods* generalize. That is, the methods for data collection, preparation, analysis, and application are generalizable from person to person. The formal position of idiographic science is that human subjects research should operate at the level of the individual. Methods and analyses should be applied to each individual, separately, yielding correspondingly unique results and conclusions. However, the theory and general methodology underlying a given study should be consistent from person to person. This is the model we employ here and it is this *process* that is generalizable.

Nevertheless, the present study also produced accurate out-of-sample predictions from pooled data models. These predictions may generalize outside of the present study and should be tested in additional samples, in order to evaluate their replicability and external validity.

Clarifying the Feature Space

While, overall, our results are encouraging, they may belie important variability in model results across the sample. For example, in some participants (e.g., P003) we observed nearly perfect predictive accuracy for future drinking events, while in others (e.g., P032) the model performed at the level of chance. We observed similar variability in the craving and wanting to drink models. These results highlight that even when utilizing personalized methods, there may exist certain individuals (or groups of individuals) for whom such methods are insufficient to accurately predict future drinking, craving, and wanting to drink. Alternatively, it may be that the drinking-relevant behavior, affect, and cognitions for these individuals *are* theoretically predictable, but that the current feature space failed to assess the subset of variables that would have facilitated more accurate modeling. Future studies should seek to continue to develop methods for identifying an optimal subset of potential mechanisms to assess through EMA, as this likely varies person to person (Soyster & Fisher, 2019).

Assessing Model Accuracy

Despite the promising results in the current study, we should emphasize that, in the case of the drinking prediction models employed here, AUC may be a somewhat misleading metric (Lobo et al., 2008). For example, AUC does not weight type 1 (*false positive*) and type 2 (*false negative*) errors differentially. For our models, predicting a drinking event occurred when it actually did not is likely less problematic than predicting a drinking event did not occur when it actually did. Depending on the base-rates of the phenomenon under investigation, AUC can be problematically biased. In the case of rare events, a model could incorrectly predict every occurrence of the behavior and still have high overall accuracy as indexed by AUC. This is relevant to the present analysis because there was marked variability in the overall frequency of drinking events across the EMA period. Future studies should examine and compare other metrics for the assessment of model accuracy; perhaps the optimal method varies as a function of the applied goal of the models (e.g., is it more important to predict the occurrence, or absence, or some other outcome?)

Sampling Frequency, Variable Selection, and Model Construction

The ability of these models to accurately (and sometimes inaccurately) predict future use, craving, and wanting to drink, relied on several design factors of this study. For example, in the present study we chose to measure the constructs of interest every 2 hr. As our variables were time-lagged, this merits careful consideration: Conceivably, some of the independent variables in our feature space could have exerted effects on dependent variables every four hours, or every day, rather than every two hours. If any time-lagged effects unfold at faster or slower time scales, these might be obscured by the selected sampling frequency. Additionally, various independent variables could exert effects on selected dependent variables at different lag lengths.

Finally, the present study utilized elastic net regularization for variable selection, but other approaches exist such as random forest variable selection (Catal & Diri, 2009; Ho, 1995). Future studies should examine how different modeling paradigms affect model, including sampling frequency, lag length, variable selection procedures, and classification methods. Different combinations among these choices may help to optimize the current methodology.

Conclusion

The present study aimed to predict drinking-relevant behavior, affect, and cognition in college students by constructing idiographic prediction models from a set of 40 possible predictor variables. These models achieved higher predictive accuracy than expected. As these models demonstrated such accuracy in the prediction of behavior hours later in out-of-sample data, they may hold promise for just-in-time interventions and other potential clinical applications.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Catal, C., & Diri, B. (2009). Investigating the effect of dataset size, metrics sets, and feature selection techniques on software fault prediction problem. *Information Sciences*, *179*(8), 1040–1058. https://doi.org/10.1016/j.ins.2008.12.001

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge. https://doi.org/10.4324/9780203771587

Collins, M. A., Neafsey, E. J., Mukamal, K. J., Gray, M. O., Parks, D. A., Das, D. K., & Korthuis, R. J. (2009). Alcohol in moderation, cardioprotection, and neuroprotection: Epidemiological considerations and mechanistic studies. *Alcoholism, Clinical and Experimental Research*, *33*(2), 206–219. https://doi.org/10.1111/j.1530-0277.2008.00828.x

Conner, M., Warren, R., Close, S., & Sparks, P. (1999). Alcohol consumption and the theory of planned behavior: An examination of the cognitive mediation of past behaviorid. *Journal of Applied Social Psychology*, *29*(8), 1676–1704. https://doi.org/10.1111/j.1559-1816.1999.tb02046.x

Fisher, A. J., & Bosley, H. G. (2020). Identifying the presence and timing of discrete mood states prior to therapy. *Behaviour Research and Therapy*, *128*, 103596. https://doi.org/10.1016/j.brat.2020.103596

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(27), E6106–E6115. https://doi.org/10.1073/pnas.1711978115

Flury, B. D., & Levri, E. P. (1999). Periodic logistic regression. *Ecology*, *80*(7), 2254–2260. https://doi.org/10.1890/0012-658(1999)080[2254:PLR]2.0.CO;2

Friedman, J. H., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1), 1–22. https://doi.org/10.18637/jss.v033.i01

Ho, T. K. (1995). Random decision forests. *Proceedings of 3rd International Conference on Document Analysis and Recognition*, *1*, 278–282. https://doi.org/10.1109/ICDAR.1995.598994

Iwanicka, K., & Olajossy, M. (2015). The concept of alcohol craving. *Psychiatria Polska*, *49*(2), 295–304. https://doi.org/10.12740/PP/27538

Jennison, K. M. (2004). The short-term effects and unintended long-term consequences of binge drinking in college: A 10-year follow-up study. *The American Journal of Drug and Alcohol Abuse*, *30*(3), 659–684. https://doi.org/10.1081/ADA-200032331

Kuntsche, E., Knibbe, R., Gmel, G., & Engels, R. (2006). Who drinks and why? A review of socio-demographic, personality, and contextual issues behind the drinking motives in young people. *Addictive Behaviors*, *31*(10), 1844–1857. https://doi.org/10.1016/j.addbeh.2005.12.028

Kuntsche, E., & Labhart, F. (2012). Investigating the drinking patterns of young people over the course of the evening at weekends. *Drug and Alcohol Dependence*, *124*(3), 319–324. https://doi.org/10.1016/j.drugalcdep.2012.02.001

Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2008). AUC: A misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, *17*(2), 145–151. https://doi.org/10.1111/j.1466-8238.2007.00358.x

Lowman, C., Hunt, W. A., Litten, R. Z., & Drummond, D. C. (2000). Research perspectives on alcohol craving: An overview. *Addiction*, *95*(Suppl. 2), S45–S54. https://doi.org/10.1046/j.1360-0443.95.8s2.14.x

Madden, T. M. (1884). Alcoholism in childhood and youth. *British Medical Journal*, *2*(1234), 358–359. https://doi.org/10.1136/bmj.2.1234.358

McHugh, R. K., Fitzmaurice, G. M., Griffin, M. L., Anton, R. F., & Weiss, R. D. (2016). Association between a brief alcohol craving measure and drinking in the following week. *Addiction*, *111*(6), 1004–1010. https://doi.org/10.1111/add.13311

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research and Perspectives*, *2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1

Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology: *Current Directions in Psychological Science*, *18*(2), 112–117. https://doi.org/10.1111/j.1467-8721.2009.01619.x

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*(1), 77. https://doi.org/10.1186/1471-2105-12-77

Salgado, J. F. (2018). Transforming the Area under the Normal Curve (AUC) into Cohen's D, Pearson's rpb, Odds-ratio, and natural log odds-ratio: Two conversion tables. *The European Journal of Psychology Applied to Legal Context*, *10*(1), 35–47. https://doi.org/10.5093/ejpalc2018a5

Slutske, W. S. (2005). Alcohol use disorders among U.S. college students and their non–college-attending peers. *Archives of General Psychiatry*, *62*(3), 321–327. https://doi.org/10.1001/archpsyc.62.3.321

Soyster, P. D., & Fisher, A. J. (2019). Involving stakeholders in the design of ecological momentary assessment research: An example from smoking cessation. *PLOS ONE*, *14*(5), Article e0217150. https://doi.org/10.1371/journal.pone.0217150

Stewart, S. H., & Devine, H. (2000). Relations between personality and drinking motives in young adults. *Personality and Individual Differences*, *29*(3), 495–511. https://doi.org/10.1016/S0191-8869(99)00210-X

Turrisi, R., Mallett, K. A., Mastroleo, N. R., & Larimer, M. E. (2006). Heavy drinking in college students: Who is at risk and what is being done about it? *The Journal of General Psychology*, *133*(4), 401–420. https://doi.org/10.3200/GENP.133.4.401-420

Wilks, D. S. (2010). Sampling distributions of the brier score and brier skill score under serial dependence. *Quarterly Journal of the Royal Meteorological Society*, *136*(653), 2109–2118. https://doi.org/10.1002/qj.709

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, *67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

Received May 19, 2020
Revision received September 23, 2020
Accepted October 19, 2020 ■