



# Using Smartphone Survey Data and Machine Learning to Identify Situational and Contextual Risk Factors for HIV Risk Behavior Among Men Who Have Sex with Men Who Are Not on PrEP

Tyler B. Wray<sup>1</sup> · Xi Luo<sup>2,3</sup> · Jun Ke<sup>2</sup> · Ashley E. Pérez<sup>4</sup> · Daniel J. Carr<sup>1</sup> · Peter M. Monti<sup>1</sup>

Published online: 9 May 2019  
© Society for Prevention Research 2019

## Abstract

“Just-in-time” interventions (JITs) delivered via smartphones have considerable potential for reducing HIV risk behavior by providing pivotal support at key times prior to sex. However, these programs depend on a thorough understanding of when risk behavior is likely to occur to inform the timing of JITs. It is also critical to understand the most important momentary risk factors that may precede HIV risk behavior, so that interventions can be designed to address them. Applying machine learning (ML) methods to ecological momentary assessment data on HIV risk behaviors can help answer both questions. Eighty HIV-negative men who have sex with men (MSM) who were not on PrEP completed a daily diary survey each morning and an experience sampling survey up to six times per day via a smartphone application for 30 days. Random forest models achieved the highest area under the curve (AUC) values for classifying high-risk condomless anal sex (CAS). These models achieved 80% specificity at a sensitivity value of 74%. Unsurprisingly, the most important contextual risk factors that aided in classification were participants’ plans and intentions for sex, sexual arousal, and positive affective states. Findings suggest that survey data collected throughout the day can be used to correctly classify about three of every four high-risk CAS events, while incorrectly classifying one of every five non-CAS days as involving high-risk CAS. A unique set of risk factors also often emerge prior to high-risk CAS events that may be useful targets for JITs.

**Keywords** HIV · Sexual behavior · Men who have sex with men · Machine learning · Ecological momentary assessment

## Introduction

Overall, HIV incidence in the USA has declined in recent years, reflecting the success of advances in biomedical,

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11121-019-01019-z>) contains supplementary material, which is available to authorized users.

---

✉ Tyler B. Wray  
tyler\_wray@brown.edu

- <sup>1</sup> Department of Behavioral and Social Sciences, Center for Alcohol and Addictions Studies, Brown University School of Public Health, 121 South Main Street, Providence, RI 02912, USA
- <sup>2</sup> Department of Biostatistics, Brown University School of Public Health, Providence, RI 02906, USA
- <sup>3</sup> Department of Biostatistics and Data Science, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA
- <sup>4</sup> Department of Social and Behavioral Sciences, University of California, San Francisco, San Francisco, CA 94118, USA

behavioral, and social prevention (Centers for Disease Control and Prevention (CDC) 2013). However, new infections continue to rise in certain subgroups of MSM (CDC 2016b), especially young MSM (aged 25–34; CDC 2017). New biomedical prevention methods, like pre-exposure prophylaxis (PrEP), show promise for reducing incidence (Centers for Disease Control and Prevention 2016c; Punyacharoensin et al. 2016). However, with PrEP uptake currently at less than 5% of all eligible MSM (Siegler et al. 2018), condoms continue to be the most widely used and widely available method of HIV/STI prevention. Even for MSM on PrEP, adherence is often imperfect (Hosek et al. 2016; Liu et al. 2014), and rates of other sexually transmitted infections (STIs) are high (Mayer et al. 2016), highlighting the need for continued condom use (US Public Health Service 2018). Together, this research emphasizes the need to continue developing innovative ways of encouraging condom use among MSM.

One particularly promising strategy for encouraging condom use involves providing “just-in-time” interventions (JITs; Nahum-Shani et al. 2017). JITs provide support at

specific times when they may be most needed or when the recipient is most receptive, based on ongoing assessments of the recipient's unique risk factors, internal state, and context (Nahum-Shani et al. 2017). This approach draws from the principle that certain antecedents, states of vulnerability (e.g., affect), or situations/contexts (e.g., alcohol intoxication) increase the likelihood of engaging in risk behavior (Mustanski 2007; Wray et al. 2016, 2019) and that providing the right support or skills in these moments can ensure that they are salient and relevant, making them easier to use (Smyth and Heron 2016). Most JITs rely on smartphones given their ubiquity and integration into everyday life (Goldstein et al. 2017). These devices are typically within reach for 90% of owners' daily lives (Dey et al. 2011), enabling nearly continuous in-the-moment monitoring and support. They may also be valuable intervention tools for MSM, as the vast majority have used smartphone apps to meet partners (Rosser et al. 2011).

The first step toward building robust JITs, however, is developing a thorough understanding of the most important states of vulnerability that precede risk behavior, so that researchers can determine which states JITs might address. Intensive longitudinal methods, like ecological momentary assessment (EMA), are well-suited for this purpose. EMA involves assessing participants' behaviors and several possible situational risk factors (e.g., affect, attitudes, contexts) in near real-time as participants go about their lives (Shiffman 2009). It relies on self-report data (often called "active" data, since it requires participant engagement) and uses a blend of experience sampling, diary methods, and event-contingent surveys to assess relevant constructs as close as possible to when they occur (Shiffman et al. 2008).

EMA research has been ongoing in the fields of addiction (Wray et al. 2015), mental health (Wenze and Miller 2010), and other health behaviors (Stone and Shiffman 1994) for over two decades and has produced critical discoveries about the dynamics of these conditions/behaviors in natural settings. It has also directly informed the development of JITs (Heron and Smyth 2010). Yet, although several pilot and feasibility studies have been reported (Paolillo et al. 2018; Swendeman et al. 2015; Wray et al. 2016; Yang et al. 2015), we are aware of no large-scale EMA studies of HIV risk behavior in MSM. Daily diary and event-level studies, however, may serve as a useful starting point for identifying candidate situational/contextual risk factors. For example, binge drinking and stimulant drug use have been shown to consistently co-occur with condomless anal sex (CAS) the same day (Vosburgh et al. 2012). Given that heavy drinking is a key risk factor for HIV acquisition in MSM and nearly half of MSM report drinking at this level in the past month (Sander et al. 2013), alcohol use may be an especially critical risk factor. Various affective states have also been shown to precede CAS events, including sexual arousal (Groves et al. 2010), sad/anxious affect

(Mustanski 2007), and positive affect (Mustanski 2007), although with some inconsistent results across studies. Other, more unique factors have also been identified, such as partner characteristics (Groves et al. 2016), motives for sex (e.g., to affirm one's attractiveness; Puterman 2009), and plans/intentions for sex (Parsons et al. 2015).

Although these findings are important, most of these studies have adopted a piecewise approach to identify vulnerable states, wherein traditional statistical models are used to test whether a specific state is associated with health risk behavior, one-by-one, while assuming a specific type of relationship (usually a linear function; Spruijt-Metz and Nilsen 2014). Given this, they are not well-suited for exploring the field of possible risk factors and arriving at a specific subset that reliably precede risk and thus might serve as potent targets of intervention (Bi et al. 2013; Spruijt-Metz et al. 2015). Machine learning (ML) approaches may be more appropriate for this goal. ML can utilize both traditional statistical approaches (e.g., logistic regression) and several less-familiar models (e.g., decision trees) to help researchers understand patterns in the data, use them to classify outcomes with high accuracy (i.e., make predictions), and identify a set of key variables that contribute to prediction (Etchings 2017). These goals are often distinct from many traditional statistical approaches in which the aim is to test whether a specific model is supported by the data (Breiman 2001).

In this study, we used EMA methods to study high-risk CAS events (i.e., those with non-exclusive or unknown HIV status partners) among HIV-negative MSM who are not on PrEP. For 30 days, participants completed a daily diary survey each morning and an experience sampling survey up to six times each day. Guided by the existing literature, surveys assessed a variety of constructs, including alcohol/drug use, affect, and motives for sex, as well as plans, attitudes, and confidence about using a condom (among others, see Supplemental Digital Content 1 for a full list). Since participants also reported the time sex began with each partner, we then applied ML methods using all available data collected prior to these events each day (or the time since the last partner) to classify high-risk CAS events (versus "safe sex" or no sex), and to (1) determine the overall accuracy with which these events could be predicted using data on these factors, and (2) identify which of these risk factors was most important in predicting risk.

## Methods

### Participants

Eighty participants were recruited from gay-oriented smartphone dating applications (e.g., Grindr, Scruff, Hornet), general social media sites (e.g., Facebook, Instagram), and in-person outreach (e.g., flyers). Eligible participants were (1) 18+ years old, (2)

assigned male sex at birth, (3) HIV-negative or unknown status, (4) not currently prescribed or taking PrEP, and (5) reported having had CAS with a non-exclusive partner at least once in the past 30 days. Those eligible also (6) consumed five or more drinks on a single occasion in the past month.

## Procedures

Participants were instructed to complete two types of assessments on their personal smartphones over a 30-day period: (1) A self-initiated, daily diary assessment, to be completed upon waking each day, and up to (2) six signal-prompted experience sampling assessments per day, delivered at random times in three-hour windows between 9 am and midnight. Participants were first screened online before being contacted by a research staff to schedule either an in-person or a videoconference orientation session. At these sessions, the staff reviewed study procedures and obtained informed consent before helping participants download an EMA app onto their devices (MetricWire; <https://www.metricwire.com/>). The staff then trained participants on the software's features and walked them through a typical day in the study. Participants were coached to achieve response rates to random prompts of 80% or greater and to complete 100% of morning assessments. Each week, feedback on each participant's response rates was provided via email, and coaching was made available for those who failed to achieve target rates. Coaching consisted of brief discussions about ways that participants could improve their response rates, like keeping their phone in view, keeping their ringer on, and setting alarms for morning surveys. Payments were contingent on response rates: they earned \$2 for each morning survey, plus a "bonus" of \$10 for every 10 days when all these assessments were completed, as well as \$0.50 for each random survey, with a "bonus" of \$10 for every 10 days if they have completed > 80% (total of \$210 possible). All procedures were approved by Brown University's IRB.

## Measures

**Individual-Level Data** Although the focus of this study was primarily on situational risk factors, participants also completed several baseline measures relevant to key outcomes, including general sexual history and alcohol/drug use patterns and problems. Sexual behavior in the past 30 days was also assessed using an online Timeline Followback (TLFB) procedure (Sobell and Sobell 1992; Wray et al. 2018, 2019). In this task, participants were presented with a calendar of the past 30 days and asked to identify days on which they had oral, anal, or vaginal sex. After identifying all days, they indicated the number of sex partners they had on each day (up to 4), as well as each partner's gender, whether they were a new partner, were a sexually exclusive partner or not, whether they

asked about each partner's HIV status or the last time they tested, and if so, what their status was, and whether each partner was on PrEP. If they were unsure for any reason what each partner's HIV status was or whether they were on PrEP, they were instructed to select "do not know." They were then asked to report which sex acts they engaged in with each partner (oral, insertive anal, receptive anal, vaginal sex) and whether they used a condom for each act. The number of new anal sex partners and CAS events in the last 30 days, as well as a number of other demographic characteristics, was entered into models as person-level predictors of daily risk behavior.

Daily diary (DD) surveys assessed sexual behavior over the past 24 h in a manner very similar to the TLFB, but in greater detail. For each partner that participants reported having oral, anal, or vaginal sex with (0–4 partners per day), they also logged how long they had known them, where they met them, their key motivations for having sex with them, and the time sex began. These daily reports were used to construct the primary outcome of the classification models: whether a given day involved "high-risk CAS," meaning condomless anal sex with a non-sexually exclusive partner, or a partner of unknown HIV status (either because they did not ask about HIV status or were not sure for some other reason), or a partner they were not sure was on PrEP (if HIV-negative/unknown). This was compared with days involving either no sex or "safe sex," meaning participants either engaged in oral sex only, had anal sex with a condom, or had CAS with a sexually exclusive partner that they had explicitly asked about HIV status/PrEP and were confident they were either on PrEP, were HIV-negative, or were HIV-positive. Although we did not inquire about whether participants had discussed treatment status with their HIV-positive partners, we characterized CAS with sexually exclusive partners that they were confident were HIV-positive as "safe sex" events because it suggested these participants had explicitly had a conversation with their partner about HIV status and risk and both partners had committed to having sex with only each other. However, it is important to note that only one of three total CAS events reported with HIV-positive partners occurred with a sexually exclusive partner, so these events are unlikely to have influenced results in either direction. Daily diary assessments also collected detailed data about past-day alcohol and drug use (e.g., number of drinks consumed, level of intoxication/high), as well as day-level data about participants' plans and intentions for sex, use of protection, and alcohol/drug use for the coming day.

Experience sampling (ES) surveys were prompted via push notification randomly within three-hour intervals between 9 a.m. and midnight each day. These surveys were used to collect data on dynamic constructs, including the type of location participants were currently in, the type of other people they were around, and several dimensions of affect at the time of assessment. These surveys also inquired about participants' desire to engage in anal sex with a man "in the next few hours," as well as their likelihood of doing so, and their

intentions for using a condom if they did so, as well as their attitudes about condoms at the time of assessment, their ability to use one if they wanted to (perceived behavioral control; Ajzen 1991), and the social norms of condom use among those they were currently around. Each question was posed in every prompt, regardless of whether or not participants indicated that sex was likely.

## Analysis

**Pre-Processing** DD data for each day was first lagged, and then, each day was classified by whether it involved no sex/“safe” sex or “high-risk” CAS. ES data for each day was then further aligned with the specific time of day (e.g., 11 p.m.) in which participants reported that sex events began, so that each day of data included only ES surveys collected prior to sex events that same day. When two or more sex events occurred on the same day that had different classifications (i.e., one involved safe sex and another involved high-risk CAS), these days were split and ES surveys collected prior to each sex event were nested within that event. However, this was rare, applying to only 0.9% ( $n = 21$ ) of all study days. Given that several ES surveys were often collected each day leading up to a sex event, we then used these data to calculate several day-level functions, including the mean, peak, slope, and variability of affective constructs (e.g., positive affect, anxiety, boredom, rejection), as well as the last rating collected prior to sex. Plans and intentions for alcohol use, drug use, and sexual behavior were summarized at the day level. Each model also included possible two-way interactions between features. See Supplemental Digital Content 1 for a table of the features included in each model, including the construct they assessed and when they were collected.

**Modeling Approach** We used receiver operating characteristic (ROC) curves (sensitivity vs. specificity) and area under the curve (AUC) to explore the peak sensitivity and specificity of several ML models, including random forests and logistic regression, in classifying high-risk CAS events. Our models also included one-day historical data collected from the day before, an approach that was consistent with that described in Bae et al. (2017). Models were trained on a random subset of 70% of available person-days and validated on the remaining 30%. After selecting the model with the highest AUC, we used the amount each feature decreased the standardized Gini index to assess the importance of that feature. We then evaluated the prediction performance of the top 20% of predictors using the same procedure. DeLong tests were used to compare differences in AUC values. Mixed logistic models were then used to calculate the univariate effect sizes of the top 20% of predictors. All predictors were entered as fixed effect variables, after adjusting for age, education, and racial/ethnic minority status, with participants entered as random effects to adjust the

dependence within participants when assessing statistical significance. Statistical significance was assessed using two-sided  $p$  values ( $< .05$ ). All analyses were conducted using R.3.3.1.

## Results

See Table 1 for participant characteristics. Across the 80 enrolled participants, 2287 person-days of daily diary data were collected, for an overall response rate of 95.3% for these surveys. A total of 11,515 experience sampling surveys were collected from these participants over 30 days, for an overall response rate of 78.9%.

Participants reported a total of 519 sex events (involving oral sex, insertive or receptive anal sex, vaginal sex) on 19.4% of all person-days. Of these sex events, 62.6% percent ( $n = 325$ ) involved anal sex, 88.3% of which occurred with high-risk partners ( $n = 287$ ) and 69.9% of which did not involve condom use ( $n = 227$ ). A total of 205 high-risk CAS events were reported across 68.8% of participants ( $n = 55$ ) and on a total of 8.2% of all collected person-days.

## Classification Accuracy

Random forest models performed better than more traditional statistical models (e.g., logistic regression; AUC = 0.88 vs. 0.75). These models achieved 80% specificity at a sensitivity value of 74%. That is, there was a 74% chance that a randomly picked CAS day would be accurately classified as involving high-risk CAS when such an event occurred (versus a “safe” sex event or no sex day). At this level of sensitivity, these models achieved a specificity value of 80%.

## Risk Factor Importance

Table 2 shows the top 20% of risk factors for high-risk CAS, listed by the extent to which they decreased the Gini index. A random forest model with only these top 20% of risk factors performed nearly as well as models that included all risk factors (AUC = 0.86,  $p = .230$ , see Fig. 1). Unsurprisingly, high ratings of the likelihood of sex and the desire for sex (in various combinations) were among the most important risk factors for predicting whether high-risk CAS events would occur. Specifically, high-risk CAS events tended to occur most often after participants provided consistently high ratings of these two items over the course of the day (prior to sex), with particularly high ratings in the hours before sex. High-risk CAS also tended to occur after participants provided consistently high ratings of sexual arousal over the course of the same day, as well as the previous day. These results suggest that the effects of sexual arousal on a given day could “spill over” to affect decisions about sex on subsequent days.

**Table 1** Demographic and behavioral characteristics of the study sample ( $N = 80$ )

Characteristics	Mean (SD) or $N$ (%)
Age (Range: 18–53, $M \pm SD$ )	27.1 (7.8)
Race	
White	59 (78.7)
Black/African American	4 (5.3)
American Indian/Alaska Native	1 (1.3)
Asian	5 (6.7)
Multiracial	6 (8.0)
Ethnicity (Hispanic or Latino)	13 (16.3)
Currently in exclusive relationship <sup>a</sup>	5 (6.3)
College degree	41 (51.3)
Low income <sup>b</sup>	22 (27.5)
Unemployed	9 (11.3)
Sexual identity	15 (100.0)
Gay	63 (78.8)
Bisexual	11 (13.8)
Other	4 (5.0)
Not sure	2 (2.5)
Total # of new anal sex partners, past 30 days <sup>c</sup>	2.3 (2.7)
Total # condomless anal sex (CAS) events, past 30 days <sup>c</sup>	2.8 (3.9)
Total # of high-risk CAS events, past 30 days <sup>c</sup>	1.4 (2.1)

<sup>a</sup> Represents participants who reported currently being in a sexually exclusive, monogamous relationship with one partner. <sup>b</sup> Represents those with a household annual income < \$30,000/year. <sup>c</sup> As assessed in the baseline TLFB

Beyond these more intuitive factors, however, general affective states also contributed to predicting later high-risk CAS events. In particular, days on which high-risk CAS occurred tended to be marked by consistently high levels of positive affect throughout the day and in the hours leading up to sex, as well as throughout the day before. Although Gini index values from the multivariate random forest model suggested that high-risk CAS events were also preceded by low levels of peak negative affect both throughout the same day and day before, the univariate odds ratios for these features were not significant. The number of experience sampling surveys submitted on a given day was also among the top features that were predictive of high-risk CAS events. Finally, two individual-level characteristics also appeared to meaningfully contribute to prediction: the number of new sex partners and number of CAS events participants reported in the month prior to enrollment. Overall, these results show that a specific pattern of changes in key factors, like sexual desire, arousal, and motivation, as well as positive dimensions of affect, emerge in the hours before sex and serve as hallmarks of future high-risk CAS events.

## Discussion

In one of the first large EMA studies of HIV risk behavior in MSM, we explored whether smartphone surveys that captured data on situational risk factors in the day-to-day lives of MSM

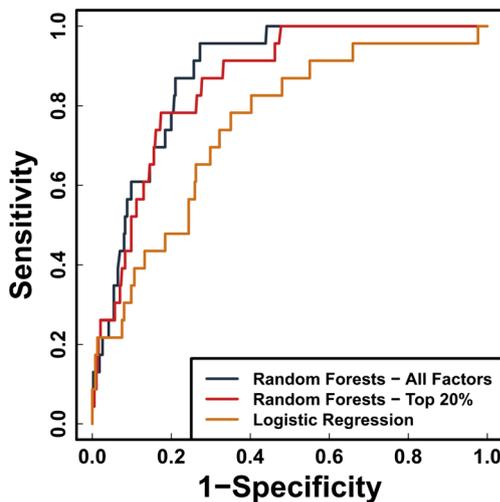
(who are not on PrEP) could predict high-risk CAS events. A key goal of this step of the analysis was to explore the potential these intensive longitudinal data have for guiding the delivery of JIT support that encourages safer choices in advance of high-risk sex events. To be appropriate for this, however, these data (taken together) must be capable of triggering support before most of the target events occur, and only when it is relevant (Nahum-Shani et al. 2017). Overall, our results suggest that “active” survey data can indeed be used to help accurately classify a majority of high-risk sex events before they occur, but may trigger irrelevant support too frequently. Given the sensitivity and specificity of our best performing model, an intervention relying on similar data would (on average) miss one out of every four high-risk sex events and trigger irrelevant support on one out of every five non-CAS days. Although these rates might be acceptable for some users, they are far from ideal. In particular, delivering support every five days when it is not needed may lead some users to begin ignoring prompts and lose trust.

However, there may be several ways of improving the performance of these models. First, although past studies suggest that self-report data captured in as close to real-time as possible is largely accurate (Hjorthøj et al. 2012; Simons et al. 2015), it is undoubtedly imperfect. Collecting self-report data via smartphone also requires near-continuous engagement from users throughout the day, a burdensome task in the context of an intervention program (Nahum-Shani et al. 2017), especially among those who are not highly motivated to use prevention

**Table 2** Gini index values and univariate odds ratios (ORs) of top 20% of risk factors contributing to prediction of high-risk condomless anal sex (CAS) events

Variable	Level	Function	Gini	OR	<i>p</i>
Likelihood of sex	Experience sampling	Mean	4.493	3.272	<0.001
Desire for sex × likelihood	Experience sampling	Mean	3.823	1.920	<0.001
Desire for sex × likelihood	Experience sampling	Last	3.210	1.777	<0.001
Sexual arousal	Experience sampling	Mean	3.100	2.081	<0.001
Desire for sex	Experience sampling	Mean	3.004	3.471	<0.001
Sexual arousal (previous day)	Experience sampling	Mean	2.874	1.885	<0.001
Condom use intentions	Experience sampling	Mean	2.680	0.440	<0.001
Likelihood of sex (previous day)	Experience sampling	Mean	2.528	2.080	<0.001
Negative affect (previous day)	Experience sampling	Peak	2.511	1.004	0.980
Number of CAS events, past month	Individual-level		2.487	1.141	<0.001
Number of new sex partners, past month	Individual-level		2.425	1.196	<0.001
Positive affect	Experience sampling	Mean	2.411	2.022	0.001
Positive affect (previous day)	Experience sampling	Mean	2.377	1.687	0.015
Total number of surveys submitted	Experience sampling	Total	2.336	0.836	0.052
Desire for sex × likelihood (previous day)	Experience sampling	Mean	2.331	1.319	0.102
Desire for sex (previous day)	Experience sampling	Mean	2.308	2.051	<0.001
Negative affect	Experience sampling	Peak	2.252	0.810	0.213
Condom use intentions (previous day)	Experience sampling	Mean	2.225	0.547	0.004
Desire for sex	Experience sampling	Last	2.178	1.797	<0.001
Positive affect	Experience sampling	Last	2.176	1.890	<0.001

methods. High user demand is also a common reason for abandoning health-related apps (Choe et al. 2014). In addition to their ability to collect survey data, smartphones also produce a variety of other “passive” data (i.e., sensor and phone use metadata) that reflect both an individual’s use of the device and users’ underlying individual and social behaviors. For example, phone call and text message metadata reflect aspects of users’ social engagement (e.g., whether they met someone new recently)



**Fig. 1.** Receiver operating characteristic (ROC) curves for models classifying high-risk condomless anal sex (CAS) events versus “safe” sex and no sex. The AUC values for logistic regression, random forests with all risk factors, and random forests with the top 20% risk factors are 0.75, 0.88 and 0.86 respectively.

and global positioning system (GPS) data reflect users’ travel, physical context, and current activity (e.g., driving, walking, running; Google Developers 2018). Software that can observe these data “passively” (i.e., without requiring ongoing interaction from the user) and without identifying individual users has recently been developed and is being used to construct “digital phenotypes” of various behaviors, emotional states, and health conditions (Onnela and Rauch 2016). The idea is that, given its richness, the overall pattern of this passive smartphone data could produce a reliable “digital fingerprint” of these behaviors and states that may help identify when key events are likely to occur and allow developers to deliver interventions without asking users for repeated interaction. This concept is clearly appealing to those interested in JIT interventions, since it could allow researchers to deliver timely, relevant support based on nothing more than users’ normal use of their devices. So far, this approach has been used to detect stress (Muaremi et al. 2013), mood (LiKamWa et al. 2013), smoking (Shoaib et al. 2015), and alcohol use (Bae et al. 2017). Given this, an important next step involves exploring whether these passive data improve prediction, either independently or jointly with “active” data.

**Risk Factor Identification**

Another key goal of this study was to identify a concise set of risk factors that most reliably predicted CAS events among MSM. Perhaps, the most important insight from these findings was that, although our models included a number of

individual-level variables, situational and contextual risk factors contributed information that was most relevant to predicting the occurrence of high-risk CAS events on a given day. This pattern of findings suggests that, although participants' past behavior in similar situations (such as their general tendency to use condoms during anal sex and their use of condoms with new/casual partners) influenced their decisions prior to a given sex event, moment-to-moment changes in psychological and emotional states also exerted relatively strong influences on these choices.

Our results also showed that a specific pattern of changes in psychological and emotional states also often emerged prior to high-risk CAS. Not surprisingly, CAS events frequently occurred after participants consistently reported a high likelihood of sex throughout the day. They also frequently occurred after participants reported high motivation for sex and sexual arousal throughout a given day, extending the findings of past laboratory studies (George et al. 2009) to show that MSM often make decisions that put them at higher risk for HIV/STI transmission after experiencing high levels of sexual arousal in their daily lives. Finally, consistent with past studies (Parsons et al. 2015), low condom use intentions also contributed meaningfully to prediction, but interestingly, seemingly less so than sexual motivation and arousal. One possible explanation for this pattern is that questions reflecting motivation for sex (i.e., desire, likelihood, arousal) capture some overall risk construct, such as participants' overall willingness to have sex even if it is with a high-risk partner and even if it involves forgoing condom use, more so than asking about condom use specifically.

Beyond these factors, shifts in several more general emotional states, most of which involved various dimensions of positive affect, also often preceded high-risk CAS events when compared with "safer" sex or no sex. In particular, high-risk CAS events were characterized by consistently high levels of positive affect (e.g., happiness, joy, excitement) over the course of the day before, day of, and in the hours leading up to them. High-risk CAS events also often occurred on days in which participants also reported low negative affect (e.g., sadness, anxiety, hostility), but this was not significant in univariate models. This pattern of results suggests that low positive affect alone may not be a reliable predictor of risk and may only be relevant in the context of other risk factors. Similar past studies exploring day-level associations between positive affect and sexual behavior in MSM have yielded conflicting results (Groves et al. 2010; Mustanski 2007). However, our results extend these studies by showing that increases in positive affect in the days and hours prior to sex were associated with decisions to engage in high-risk CAS. Finally, high-risk CAS also tended to occur on days in which participants submitted a fewer number of experience sampling surveys. Although this could suggest some degree of reactivity as has been reported in some past studies (Newcomb and Mustanski

2013), it more likely reflects that more surveys tended to be available on days that did not involve sex because we only used surveys submitted prior to the beginning of sex on days when it occurred.

These findings are also notable because of the variables that were not identified among the top 20% that contributed to prediction. For example, although alcohol and certain types of drugs (e.g., stimulants) have been shown to increase the probability of engaging in high-risk CAS in MSM at the event-level (Vosburgh et al. 2012), they were not among the top factors that characterized high-risk CAS events in this sample. Likewise, other emotional and behavioral contexts that have been hypothesized to increase risk in MSM, such as loneliness, rejection, and discrimination, also did not appear among the most important risk factors. It is important to note, however, that classification models like these do not test mechanisms by which risk behavior may occur, only which factors are most useful in distinguishing high-risk CAS from other classes of events. So, other factors that were not identified in the current sample could still be important, but may be more distal to the outcome. For example, alcohol and drug use may result in increases in positive affect that, in turn, increase the risk for high-risk CAS. In this scenario, alcohol and drug use may not appear among the top variables contributing to prediction, but would still play an important role in the process leading to HIV risk behavior. Since similar studies are rare, many of these mechanisms have not yet been tested.

### Implications for Intervention and Monitoring

Like overall classification accuracy, risk factor identification also has important implications for designing systems intended to help monitor MSM's risk for engaging in HIV risk behavior or for intervening at critical times. First, rather than monitoring a wide variety of possible risk factors, our results suggest that these programs might predict high-risk CAS events with a high degree of precision by tracking just seven key risk factors over time. That is, our findings show that monitoring changes in participants' estimations of the likelihood of sex, their desire for sex, and sexual arousal throughout the day, as well as more general positive and negative affect, could successfully predict many high-risk CAS events. Our findings also provide important information about optimal timing of surveys. Although we included day-level variables for many key risk factors that reflected participants' ratings in the morning on a given day, similar variables collected over the course of the day often appeared to be more useful in predicting high-risk CAS events. That is, providing consistently high or low ratings across several surveys collected later in the day was almost always more useful for prediction than broader day-level ratings. Given this, future research and intervention programs might explore various ways of monitoring this more limited set of factors throughout the day that

avoid excessive burden and minimize interruption (e.g., “micro-EMA” approaches; Intille et al. 2016). Finally, our results also showed that the effects of several risk factors experienced the day before influenced behavior the next day. For example, both sexual arousal throughout the current day and the previous day contributed to predicting high-risk CAS events. For this reason, tracking and intervention programs should consider including at least one-day historical data in prediction models trained on similar survey data.

Identifying a set of key risk factors is also helpful for identifying important targets of JIT interventions aimed at interrupting users’ trajectory toward risk. Although some of the most important risk factors identified through these analyses may themselves be relatively unmalleable (e.g., likelihood/desire for sex), delivering JITs directly addressing key barriers to using prevention methods or encouraging simple risk reduction techniques specifically when these risk factors are high could still help reduce risk behavior. For example, showing users a map of nearby locations to obtain free condoms, reminding them of their long-term health-related goals, or recapping other harm reduction techniques specifically when their desire for sex is high could ensure these interventions are relevant and serve as an important “nudge” to action. Other situational risk factors (e.g., positive affect, sexual arousal), though, may be malleable, meaning that interventions may be able to directly reduce the influence these states have on future decisions. For example, JIT interventions could highlight the connection between these states and unpleasant consequences and suggest that avoiding risk may be one way to help users maintain their positive mood (Aspinwall 1998).

## Limitations

Although this study has many strengths, several limitations should also be noted. First, this study relied almost entirely on self-report data, both in terms of the risk factors monitored and the outcome itself. Such data are always subject to biases or inaccuracies (Shiffman 2009) and could explain some of our misclassification rate. Second, only a total of 205 high-risk CAS events were available for analysis. Although there are no universal rules for the minimum number of cases needed for ML models, having more risk events available for analysis would likely have allowed us to achieve better prediction performance. Future research should study larger samples over longer time periods. Third, our results were produced by training models on data collected from a sample of mostly White, high-risk MSM. As such, these results may not be applicable to other populations (e.g., heterosexual men and women) or more diverse samples of MSM. Given especially high HIV incidence particularly among African-American and Hispanic/Latino men gay and bisexual men (CDC 2016a),

similar research should aim to understand situational factors predicting HIV risk behavior among these men.

In summary, this study demonstrates one way that modern data analysis techniques can be used to help identify patterns in large data sets, rather than testing whether the data support a finite set of pre-specified models. It also illustrates how these methods can help inform approaches to monitor and prospectively predict health risk behaviors in digital health interventions, as well as identify a concise set of factors (among many) that commonly emerge prior to risk behaviors that may be useful intervention targets in themselves or help researchers determine when to intervene. Our results suggested that survey data collected via smartphone throughout the day could help predict specific HIV risk behaviors in MSM before they occur with a reasonable degree of accuracy and that a key set of situational risk factors may help signal future risk events. These findings contribute to the existing body of literature on momentary and situational risk factors for HIV risk behavior among MSM and provide a guide for tracking risk and possible intervention targets for those developing digital health interventions intended to improve prevention in this population.

**Acknowledgements** This manuscript was supported by P01AA019072 (to PM) and L30AA023336 (to TW) from the National Institute on Alcohol Abuse and Alcoholism.

## Compliance with Ethical Standards

**Conflict of Interest** The authors declare that they have no conflict of interest.

**Research Involving Human Participants and/or Animals** All procedures in this study were approved by the Brown University Institutional Review Board.

**Informed Consent** All participants in this study provided informed consent prior to enrollment.

## References

- Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, 50, 179–211.
- Aspinwall, L. G. (1998). Rethinking the role of positive affect in self-regulation. *Motivation and Emotion*, 22, 1–32.
- Bae, S., Ferreira, D., Suffoletto, B., Puyana, J. C., Kurtz, R., Chung, T., & Dey, A. K. (2017). Detecting drinking episodes in young adults using smartphone-based sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1, 5.
- Bi, J., Sun, J., Wu, Y., Tennen, H., & Armeli, S. (2013). A machine learning approach to college drinking prediction and risk factor identification. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 4, 72.
- Breiman, L. (2001). Statistical modeling: the two cultures (with comments and a rejoinder by the author). *Statistical Science*, 16, 199–231.

- Centers for Disease Control and Prevention. (2013). *HIV prevention: progress to date*. Atlanta: U.S. Department of Health and Human Services.
- Centers for Disease Control and Prevention. (2016a). Lifetime risk of HIV diagnosis. Retrieved from: <https://www.cdc.gov/nchhstp/newsroom/2016/croi-press-release-risk.html>.
- Centers for Disease Control and Prevention. (2016b). Trends in HIV diagnoses, 2005–2014. Retrieved from: <http://www.webcitation.org/6vAPQZG6c>.
- Centers for Disease Control and Prevention. (2016c). As many as 185,000 new HIV infections in the U.S. could be prevented by expanding testing, treatment, PrEP. Retrieved from: <https://www.cdc.gov/nchhstp/newsroom/2016/croi-press-release-prevention.html>.
- Centers for Disease Control and Prevention. (2017). *HIV among African American gay and bisexual men*. Atlanta: National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention.
- Choe, E. K., Lee, N. B., Lee, B., Pratt, W., & Kientz, J. A. (2014). Understanding quantified-selfers' practices in collecting and exploring personal data. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1143–1152). New York: ACM.
- Dey, A. K., Wac, K., Ferreira, D., Tassini, K., Hong, J. H., & Ramos, J. (2011). Getting closer: an empirical investigation of the proximity of user to their smart phones. *Proceedings of the 13th International Conference on Ubiquitous Computing* (pp. 163–172). New York: ACM.
- Etchings, J. A. (2017). *Strategies in biomedical data science: driving force for innovation*. Hoboken: John Wiley & Sons.
- George, W. H., Davis, K. C., Norris, J., Heiman, J. R., Stoner, S. A., Schacht, R. L., et al. (2009). Indirect effects of acute alcohol intoxication on sexual risk-taking: the roles of subjective and physiological sexual arousal. *Archives of Sexual Behavior*, *38*, 498–513.
- Goldstein, S. P., Evans, B. C., Flack, D., Juarascio, A., Manasse, S., Zhang, F., & Forman, E. M. (2017). Return of the JITAI: applying a just-in-time adaptive intervention framework to the development of m-health solutions for addictive behaviors. *International Journal of Behavioral Medicine*, *24*, 673–682.
- Google Developers. (2018). Google APIs for Android: detected activity. Retrieved from: <http://www.webcitation.org/72BzwQGD0>.
- Grov, C., Golub, S. A., Mustanski, B., & Parsons, J. T. (2010). Sexual compulsivity, state affect, and sexual risk behavior in a daily diary study of gay and bisexual men. *Psychology of Addictive Behaviors*, *24*, 487–497.
- Grov, C., Rendina, H. J., Ventuneac, A., & Parsons, J. T. (2016). Sexual behavior varies between same-race and different-race partnerships: a daily diary study of highly sexually active Black, Latino, and White gay and bisexual men. *Archives of Sexual Behavior*, *45*, 1453–1462.
- Heron, K. E., & Smyth, J. M. (2010). Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. *British Journal of Health Psychology*, *15*, 1–39.
- Hjorthøj, C. R., Hjorthøj, A. R., & Nordentoft, M. (2012). Validity of timeline follow-back for self-reported use of cannabis and other illicit substances—systematic review and meta-analysis. *Addictive Behaviors*, *37*, 225–233.
- Hosek, S., Landovitz, R., Rudy, B., Kapogiannis, B., Siberry, G., & Rutledge, B. (2016). An HIV pre-exposure prophylaxis (PrEP) demonstration project and safety study for adolescent MSM ages 15–17 in the United States (ATN 113). Paper presented at the International AIDS Conference.
- Intille, S., Haynes, C., Maniar, D., Ponnada, A., & Manjourides, J. (2016). *μEMA: microinteraction-based ecological momentary assessment (EMA) using a smartwatch*. Paper presented at the proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing.
- LiKamWa, R., Liu, Y., Lane, N. D., & Zhong, L. (2013). *Moodscope: building a mood sensor from smartphone usage patterns*. Paper presented at the proceeding of the 11th annual international conference on Mobile systems, applications, and services.
- Liu, A., Glidden, D. V., Anderson, P. L., Amico, K. R., McMahan, V., Mehrotra, M., et al. (2014). Patterns and correlates of PrEP drug detection among MSM and transgender women in the Global iPrEx Study. *Journal of Acquired Immune Deficiency Syndromes* (1999), *67*, 528–537.
- Mayer, K., Maloney, K., Levine, K., King, D., Grasso, C., Krakower, D., & Boswell, S. (2016). *HIV infection and PrEP use are independently associated with diagnoses of bacterial sexually transmitted infections in men accessing care at a Boston community health center*. Paper presented at the ID week, New Orleans, LA.
- Muaremi, A., Arnrich, B., & Tröster, G. (2013). Towards measuring stress with smartphones and wearable devices during workday and sleep. *BioNanoScience*, *3*, 172–183.
- Mustanski, B. (2007). The influence of state and trait affect on HIV risk behaviors: a daily diary study of MSM. *Health Psychology*, *26*, 618–626.
- Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2017). Just-in-time adaptive interventions (JITAs) in mobile health: key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine*, *52*, 446–462.
- Newcomb, M. E., & Mustanski, B. (2013). Diaries for observation or intervention of health behaviors: factors that predict reactivity in a sexual diary study of men who have sex with men. *Annals of Behavioral Medicine*, *47*, 325–334.
- Onnela, J. P., & Rauch, S. L. (2016). Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. *Neuropsychopharmacology*, *41*, 1691–1696.
- Paolillo, E. W., Obermeit, L. C., Tang, B., Depp, C. A., Vaida, F., Moore, D. J., & Moore, R. C. (2018). Smartphone-based ecological momentary assessment (EMA) of alcohol and cannabis use in older adults with and without HIV infection. *Addictive Behaviors*, *83*, 102–108.
- Parsons, J. T., Rendina, H. J., Grov, C., Ventuneac, A., & Mustanski, B. (2015). Accuracy of highly sexually active gay and bisexual men's predictions of their daily likelihood of anal sex and its relevance for intermittent event-driven HIV Pre-Exposure Prophylaxis. *Journal of Acquired Immune Deficiency Syndromes (JAIDS)*, *68*, 449–455.
- Punyacharoensin, N., Edmunds, W. J., De Angelis, D., Delpech, V., Hart, G., Elford, J., et al. (2016). Effect of pre-exposure prophylaxis and combination HIV prevention for men who have sex with men in the UK: a mathematical modelling study. *The Lancet HIV*, *3*, e94–e104.
- Puterman, E. (2009). Bringing risk prevention into the bedroom: sex motives and risky behaviors in men who have sex with men. Doctoral dissertation, University of British Columbia.
- Rosser, B. S., Wilkerson, J. M., Smolenski, D. J., Oakes, J. M., Konstan, J., Horvath, K. J., et al. (2011). The future of Internet-based HIV prevention: a report on key findings from the Men's INternet (MINTS-I, II) Sex Studies. *AIDS and Behavior*, *15*, 91–100.
- Sander, P. M., Cole, S. R., Stall, R. D., Jacobson, L. P., Eron, J. J., Napravnik, S., et al. (2013). Joint effects of alcohol consumption and high-risk sexual behavior on HIV seroconversion among men who have sex with men. *Aids*, *27*, 815–823.
- Shiffman, S. (2009). Ecological momentary assessment (EMA) in studies of substance use. *Psychological Assessment*, *21*, 486–497.
- Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, *4*, 1–32 Retrieved from: <http://www.ncbi.nlm.nih.gov/pubmed/18509902>.
- Shoab, M., Bosch, S., Scholten, H., Havinga, P. J., & Incel, O. D. (2015). Towards detection of bad habits by fusing smartphone and smartwatch sensors. *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)* (pp. 591–596). St. Louis: IEEE.

- Siegler, A., Mouhanna, F., Giler, R., McCallister, S., Yeung, H., Jones, J., ... Sullivan, P. S. (2018). *Distribution of active PrEP prescriptions and the PrEP-to-need ratio, US, Q2 2017*. Paper presented at the conference on retroviruses and opportunistic infections (CROI), Boston, MA.
- Simons, J. S., Wills, T. A., Emery, N. N., & Marks, R. M. (2015). Quantifying alcohol consumption: self-report, transdermal assessment, and prediction of dependence symptoms. *Addictive Behaviors, 50*, 205–212.
- Smyth, J. M., & Heron, K. E. (2016). Is providing mobile interventions "just-in-time" helpful? An experimental proof of concept study of just-in-time intervention for stress management. *Paper presented at IEEE Wireless Health 2016* (pp. 1–7). Retrieved from: [https://sites.psu.edu/shadelab/files/2016/11/Smyth\\_Heron\\_2016\\_JIT\\_IEEE-1vleki4.pdf](https://sites.psu.edu/shadelab/files/2016/11/Smyth_Heron_2016_JIT_IEEE-1vleki4.pdf).
- Sobell, L. C., & Sobell, M. B. (1992). Timeline follow-back *measuring alcohol consumption* (pp. 41–72): Springer.
- Spuijt-Metz, D., & Nilsen, W. (2014). Dynamic models of behavior for just-in-time adaptive interventions. *IEEE Pervasive Computing, 13*, 13–17.
- Spuijt-Metz, D., Hekler, E., Saranummi, N., Intille, S., Korhonen, I., Nilsen, W., et al. (2015). Building new computational models to support health behavior change and maintenance: new opportunities in behavioral research. *Translational Behavioral Medicine, 5*, 335–346.
- Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine, 16*, 199–202.
- Swendeman, D., Ramanathan, N., Baetscher, L., Medich, M., Scheffler, A., Comulada, W. S., & Estrin, D. (2015). Smartphone self-monitoring to support self-management among people living with HIV: perceived benefits and theory of change from a mixed-methods, randomized pilot study. *Journal of Acquired Immune Deficiency Syndromes (JAIDS), 69*, S80–S91.
- US Public Health Service. (2018). *Preexposure prophylaxis for the prevention of HIV infection in the United States - 2017 update*. New York: US Department of Health and Human Services.
- Vosburgh, H. W., Mansergh, G., Sullivan, P. S., & Purcell, D. W. (2012). A review of the literature on event-level substance use and sexual risk behavior among men who have sex with men. *AIDS and Behavior, 16*, 1394–1410.
- Wenze, S. J., & Miller, I. W. (2010). Use of ecological momentary assessment in mood disorders research. *Clinical Psychology Review, 30*, 794–804.
- Wray, T. B., Merrill, J., & Monti, P. M. (2015). Using ecological momentary assessment (EMA) to assess situation-level risk factors for heavy drinking and alcohol-related consequences. *Alcohol Research & Health, 36*, 19–27.
- Wray, T. B., Kahler, C. W., & Monti, P. M. (2016). Using ecological momentary assessment (EMA) to study sex events among very high-risk men who have sex with men (MSM). *AIDS and Behavior, 20*, 2231–2242. <https://doi.org/10.1007/s10461-015-1272-y>.
- Wray, T. B., Adia, A. C., Pérez, A. E., Simpanen, E. M., Woods, L.-A., Celio, M. A., & Monti, P. M. (2018). Timeline: A web application for assessing the timing and details of health behaviors. *The American Journal of Drug and Alcohol Abuse, 45*, 141–150. <https://doi.org/10.1080/00952990.2018.1469138>.
- Wray, T. B., Celio, M. A., Pérez, A. E., DiGuseppi, G. T., Carr, D. J., Woods, L. A., & Monti, P. M. (2019). Causal effects of alcohol intoxication on sexual risk intentions and condom negotiation skills among high-risk men who have sex with men (MSM). *AIDS and Behavior, 23*, 161–174.
- Yang, C., Linas, B., Kirk, G., Bollinger, R., Chang, L., Chander, G., ... & Latkin, C. (2015). Feasibility and acceptability of smartphone-based ecological momentary assessment of alcohol use among African American men who have sex with men in Baltimore. *JMIR mHealth and uHealth, 3*, e67.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.