



HHS Public Access

Author manuscript

Am J Drug Alcohol Abuse. Author manuscript; available in PMC 2020 January 01.

Published in final edited form as:

Am J Drug Alcohol Abuse. 2019 ; 45(2): 141–150. doi:10.1080/00952990.2018.1469138.

Timeline: A Web Application for Assessing the Timing and Details of Health Behaviors

Tyler B. Wray¹, Alexander C. Adia¹, Ashley E. Pérez^{1,2}, Erik M. Simpanen¹, Laura-Allison Woods¹, Mark A. Celio¹, and Peter M. Monti¹

¹ Center for Alcohol and Addictions Studies, Brown University School of Public Health, Providence, RI 02906

² Department of Social and Behavioral Sciences, University of California, San Francisco, San Francisco, CA 94118

Abstract

Background: Timeline Followback (TLFB) interview methods are used to assess a variety of health behaviors, including alcohol use, drug use, and sexual behavior. While several online TLFBs have been developed, most focus on single behaviors, and few studies have explored their validity in assessing multiple risk behaviors using a single online TLFB.

Objective: To examine the validity of a customizable web application (Timeline) for assessing alcohol use, drug use, and sexual behavior among high-risk men who have sex with men.

Methods: Participants ($N = 15$ men) completed standardized survey instruments before undergoing a 30-day daily diary procedure where they submitted daily reports of health risk behaviors via smartphone. They then completed a Timeline at the end of the 30-day period covering the same time interval.

Results: Comparing a baseline administration of Timeline with popular surveys of health risk behaviors supported Timeline's validity ($r = 0.41$ – 0.59 for alcohol use, $r = 0.83$ for drug use, and $r = 0.34$ – 0.52). While participants reported similar amounts of each behavior via daily diary as they did on a follow-up Timeline ($r = 0.55$ – 0.88 for alcohol use, $r = 0.69$ for drug use, and $r = 0.87$ – 0.92 for sexual behaviors), results provided evidence of underreporting on the Timeline. Timing of behaviors also frequently disagreed across these methods.

Conclusions: Timeline is valid for assessing overall engagement in alcohol use, drug use, and sexual behavior over a 30-day window. However, researchers interested in the specific timing of behaviors within assessment intervals may wish to use smaller follow-up intervals (e.g., 7-days, 14-days) or more intensive reporting methods (e.g., daily diary).

Correspondence regarding this article may be sent to: Tyler B. Wray, Ph.D., Center for Alcohol and Addictions Studies, Brown University School of Public Health, 121 South Main Street, Providence, RI 02912, [Phone](mailto:tyler_wray@brown.edu): 401-863-6659, [Fax](mailto:tyler_wray@brown.edu): 401-863-6697, tyler_wray@brown.edu.

Conflicts of Interest

The authors have no conflicts of interest to report.

Keywords

HIV-risk; alcohol use; drug use; assessment; longitudinal research

BACKGROUND

The Timeline Followback (TLFB; 1, 2) is an approach to assessing health behaviors that originally focused on alcohol use. The TLFB presents respondents with a calendar of a specific time period (e.g., the past 30 days) and asks them to identify important events that occurred (e.g., birthdays, anniversaries) during that time to aid in recall. Respondents are then asked to identify specific days on which they engaged in a behavior (e.g., drinking), before reporting details of their behavior on that day (e.g., number of drinks consumed), beginning from the most recent day to the most distal. Thus, the TLFB collects both cumulative and detailed, day-level data about behaviors over a given time interval. This approach may improve recall by avoiding the need for respondents to aggregate behaviors over broad recall intervals (3). The TLFB approach has been applied to a variety of behaviors, including alcohol use, other drug use, and sexual behavior (e.g., 1, 4, 5), and has been used in a variety of populations, including college students, those with substance use disorders, and others (e.g., 6, 7, 8).

While current research supports the reliability of TLFB methods across a variety of behaviors and delivery methods (9–11), studies exploring the convergent validity of TLFB suggest that it may vary based on the behaviors assessed and measures against which TLFB is compared. Early studies comparing TLFBs with aggregate surveys assessing the same behaviors have shown sizable associations between these two measures (4, 5). Findings from studies of alcohol and drug use that compared TLFBs with relevant biomarkers have largely supported the validity of TLFBs (12, 13). A key limitation is that these studies often cannot establish validity of *timing* of behaviors on the day level as reported on the TLFB. One exception is a recent study that found one-week TLFBs largely agreed with biomarker data collected in real-time using transdermal alcohol sensors (14).

Another approach to exploring the validity of the timing of behaviors reported on the TLFB at the day level is to compare these reports with intensive surveys that also assess the behavior daily over the same period of time, using daily phone calls or online “diaries.” Although some studies demonstrated considerable agreement between these methods of assessment (15, 16), others showed that alcohol use was significantly underreported on the TLFB compared to daily assessments (17, 18). Studies of sexual behavior echo these findings (19, 20), adding that participants’ reports on a given day are often discrepant across these methods (21). These findings suggest that while the TLFB may provide reasonable estimates of overall behaviors, respondents may underreport their behaviors on the TLFB and may have problems accurately reporting *timing* of behaviors. More studies that compare TLFBs with daily self-reports are needed, especially those including sexual behavior.

The TLFB has been adapted for computer and online delivery, which has many advantages, including standardizing delivery, reducing researcher/participant burden, enabling remote data collection, and increasing participant privacy/confidentiality, key concerns for studies of

drug use and sexual behavior (22, 23). While several studies have shown no differences between TLFB data collected via computer/online versus interviews (24–26), at least one found that respondents may report more behaviors via online-delivered TLFBs compared to interviews (25). These results suggest online and computerized TLFBs are at least equivalent to interview-delivered TLFBs for assessing alcohol and drug use.

However, online TLFB tools have important areas for improvement. First, most currently existing tools were developed for use in specific studies and assess only one or two behaviors. A more flexible solution could allow researchers to customize types of behaviors assessed and details of interest for each. Second, many existing online TLFB tools employ overly simplified user interfaces displaying text and plain fields presented in calendar-like layouts. A more graphically-oriented interface could increase user engagement and assist in recall, especially when assessing different types of behaviors in the same study (e.g., alcohol and drug use).

To address these limitations, we developed a tool (Timeline) that provides a flexible and interactive approach to collecting data on multiple health behaviors. In this manuscript, we describe Timeline and its functionality. We also describe a small validation study that we conducted during the initial phase of a larger study. This study preliminarily explores the convergent validity of this tool for collecting reports of alcohol use, drug use, and sexual behavior among high-risk men who have sex with men (MSM), by comparing this data with other common survey instruments and daily diary assessments collected via smartphone. We hypothesized that Timeline-collected assessments of alcohol/drug use and sexual behavior would be strongly positively associated with reported behaviors on daily diary surveys, but that Timeline-reported behaviors may be underreported compared to daily reports.

METHODS

Participants

Participants were recruited from Rhode Island and the greater Boston, Massachusetts area from popular gay-oriented smartphone apps (e.g., Grindr), social media, or in-person outreach to participate in a daily diary study on alcohol/drug use and HIV risk. Eligible participants were (a) males, (b) over age 18, who (c) self-reported being HIV-negative, (d) reported at least one condomless anal sex event with a casual male partner in the last 30 days, and (e) were “at-risk” drinkers, meaning they reported consuming either 14+ drinks per week or 5+ drinks on a single occasion at least once in the last month (NIAAA). Participants were ineligible if they were (a) currently taking pre-exposure prophylaxis (PrEP) for HIV prevention, (b) were not fluent in English, (c) were currently receiving alcohol/drug treatment, or (d) had been diagnosed with a serious mental illness (e.g., schizophrenia, bipolar disorder). Although we recently reported data from a *different* TLFB in a feasibility study of intensive longitudinal methods (blinded), the sample reported here represents a unique sample of fifteen participants who enrolled in the full daily diary study and used the most recent version of Timeline app described below. Thus, the sample reported here does not overlap with this previous sample. Descriptive characteristics of this sample are provided in Table 1.

297 participants were screened for eligibility for this and other studies via our recruitment pool between March to July 2017. 61 met eligibility criteria for the larger study, and 18 were enrolled. Three participants (17%) withdrew prior to completing the four-week study, most often in the first two weeks. The most common reason for withdrawal was that frequent completion of smartphone surveys was incompatible with work/school schedules or other demands. 55.6% were recruited via Grindr, 22.2% from FB, and 22.2% from Instagram.

Measures

Timeline—is an interactive, customizable web application that allows researchers to assess many behaviors online, using a TLFB-like approach. In an ‘admin’ dashboard, administrators create new studies, specify their start/end dates, and set a recall window: 7, 14, 30, or 60 days. Administrators can specify type and order of pre-specified behaviors they wish to assess (alcohol/drug use, sexual behavior, medication adherence), or add custom forms and icons to assess new behaviors. To add a new participant to each study, administrators add a unique four-digit ID number to the study’s panel and enter a ‘start date’ for each that corresponds with the beginning of that participant’s recall period.

After a participant’s ID number and start date have been added to the admin panel, participants complete Timeline by entering this administrator-assigned ID number and setting a password for their account. Participants then respond to any parent questions before being presented with a calendar that reflects the recall interval specified during study set-up. These intervals are ‘rolling’ based on the start date specified for each participant, so that they begin recalling for this first interval the first time they log in, regardless of when they access Timeline. If the study is longitudinal, recall will begin at the last date participants completed during their last login. After the calendar is presented, participants are asked to identify important days that occurred during the recall window, and then specify why each day was significant (e.g., birthday, anniversary). Then, participants are asked to identify the days on which each specified behavior occurred during that interval, and each day is marked by a behavior-specific icon (see Figure 1). After all days on which each behavior occurred are identified, a ‘detail’ view then allows participants to respond to a set of administrator-specified questions about the behaviors that occurred on each specific day (e.g., “how many standard drinks did you consume?”). A calendar with identified behaviors is provided on these screens as a visual aid (see Figure 2).

Since Timeline uses ID numbers instead of personal information, the data collected, stored, and exported is inherently de-identified to ensure confidentiality. Data collected is exported for each study in .csv format through the admin panel. To connect participants with Timeline, researchers can provide them with either a main study URL or an ID-specific URL, which may prevent ID entry errors. Using ID-specific URLs facilitates connecting Timeline with popular online survey tools (e.g., Qualtrics®), using embedded data and query strings.

In this study, Timeline was used to assess alcohol use, drug use, and sexual behavior over two 30-day periods: one assessing the 30 days prior to enrollment, and the other 30 days later, at study completion. For alcohol use, participants were asked how many standard drinks (12 oz. beer, 5 oz. wine, 1 oz. of liquor) they consumed each drinking day and were

provided with a visual key to assist in responding. For each day on which drug use was reported, participants were asked which drugs they used. Finally, for each day participants reported engaging in sexual behavior, participants were asked how many partners they had (up to 4 could be reported), the gender of each partner, whether or not they engaged in oral, insertive anal, receptive anal, or vaginal sex that day, and whether or not a condom was used for each act. Participants could identify that multiple behaviors occurred on a single day.

Daily diary assessments.—Participants submitted daily reports of their alcohol use, drug use, and sexual behavior the previous day using MetricWire (<http://metricwire.com/>), a smartphone app used for survey research. Daily surveys asked about alcohol use, drug use, and sexual activity in a way that mirrored Timeline’s assessment of specific days. Days with missing data using either method were deleted casewise

Baseline measures.—Several individual-level assessments were collected during participants’ baseline appointments and were used in this study to explore Timeline’s validity.

Alcohol Use Disorders Identification Test (AUDIT).: The AUDIT is 10-item survey that assesses alcohol use and related problems over the past year (27) and has excellent psychometric properties (28). We used the first three items from the AUDIT, which assess frequency and quantity of alcohol use and frequency of binge drinking, to compare them with drinking reports submitted via Timeline.

Drug Abuse Screening Test (DAST-10).: The DAST is a 10-item survey that assesses drug-related problems (29) and has good psychometric properties (30). We used the DAST total score to compare with drug use reports submitted using Timeline.

Sexual Behavior Survey (SBS; 31).: A version of the SBS, modified to suit men who have sex with men, was used to assess number of male sex partners and anal sex partners in the past year. The SBS assessed how frequently participants used condoms during insertive and receptive anal sex in the past year, rated on a 1 (*always*) to 5 (*never*) scale.

Procedures

Participants were screened online. After consenting, those who met basic eligibility criteria were oriented to study procedures through in-person visits or video calls. During these appointments, participants completed baseline surveys and a Timeline of their alcohol use, drug use, and sexual behavior over the 30 days prior to enrollment. Next, research staff guided participants in downloading the MetricWire app for daily diary data collection. Participants were instructed to complete an assessment on their smartphone once each day for 30 days, as soon as possible after waking. Push notifications reminded them to do so each day at 9:00 a.m, though participants could complete these assessments within a 12 hour window. At the end of the 30-day period, participants completed another Timeline assessment, this time assessing the same 30-day period as the daily diary assessments.

Compensation was provided based on daily diary response rates. Participants enrolling in person were paid \$2 for each diary assessment they submitted, plus a bonus of \$10 for every

week they submitted 100% of these assessments, for a total possible payment of \$100. Those enrolling via video call were paid \$1.40 for each morning assessment, plus a bonus of \$10 for every week they submitted 100% of reports, for a total possible payment of \$82. All procedures were approved by the Institutional Review Board of [masked for review].

Analysis Plan

We first reviewed both the daily diary and Timeline responses for any missing data, and then matched daily diary responses (based on the date and time they were submitted) to Timeline-collected data. We then calculated descriptive statistics for demographics and all behavioral data of interest. To explore the validity of data collected using Timeline, we estimated pairwise correlations between Timeline variables and baseline assessments of similar constructs/behaviors, and between Timeline variables and the total/average of each behavior reported via daily diary. Finally, we compared the extent to which day-level reports of alcohol use, drug use, and sexual behavior agreed across the daily diary and Timeline assessments using Fleiss' kappa for absolute agreement (32). All statistical analyses were conducted in Stata 13 (Stata Corp., 2013).

RESULTS

Fifteen participants completed the study, 13 of whom were enrolled via video call appointments (83%). Participants provided 98.2% of daily diary assessments within 12 hours of being prompted across the 30-day study period, and all participants completed both baseline and follow-up versions of Timeline. After dropping lagged days, this resulted in available data for 439 days. Table 2 shows bivariate correlations between Timeline-generated summary variables and relevant items from baseline surveys. These correlations ranged from moderate to strong ($r = 0.50\text{--}0.59$), suggesting that Timeline variables are generally related to items from other surveys that assess the same or similar concepts. While the correlation between frequency of drinking as measured by the AUDIT and number of drinking days reported via Timeline showed a more modest relationship ($r=0.41$), this may be due to slight differences between the two variables: the AUDIT asks participants to indicate how often they typically drank in a given month over the past year, while the Timeline asks participants to indicate the specific days on which they drank in the past month. Similarly, a modest correlation ($r=0.34$) was observed between the number of male sex partners participants reported over the past year on the SBS versus the number of sex events reported on Timeline, but these variables assess slightly different constructs (i.e., number of *partners* versus number of *events*) as well as different recall intervals. These findings suggest that assessing these behaviors with the Timeline meets conventional standards for concurrent validity.

Participant mean totals for each behavior reported via daily diary and Timeline are provided in Table 3, with bivariate correlations between overall reports of each behavior for the two assessment methods. These results suggest overall engagement in each behavior was similar across methods, but that participants reported fewer behaviors on Timeline than they did via daily diary. Although this difference was modest for most behaviors, total number of drinking days, "heavy" drinking days, and drug use days differed more substantially, with

participants reporting an average of 3.7 fewer drinking days, 1.9 fewer “heavy” drinking days, and 1.7 fewer drug use days via Timeline than they did through daily diaries. This may explain the modest bivariate correlations observed between the two methods for these behaviors ($r = 0.55\text{--}0.69$), when compared with others ($r = 0.87\text{--}0.92$). Figure 3 provides scatterplots of the total number of each behavior, as reported on Timeline (x axis) and via daily diary (y axis). The diagonal reference line represents what would be perfect agreement between the two methods. Across behaviors, total values appear to be largely grouped above this reference line, offering further support for the conclusion that participants reported slightly more behaviors via daily diary than they did through Timeline.

Finally, Table 4 shows the day-level agreement between the two methods for each behavior, with Fleiss’ kappa statistics for the average agreement between the two methods on specific days. These results suggest that there was substantial agreement about whether “heavy” drinking, drug use, and any sexual activity occurred on a specific day, when averaging day-level agreement across the entire 30-day period. Reports of these behaviors agreed on more than 80% of days, and agreed more often than might be expected by chance alone ($p < .05$). However, overall average agreement between the methods was considerably poorer when comparing reports about more moderate levels of alcohol use, as well as for specific types of sex events, like whether or not anal sex and/or condomless anal sex (CAS) occurred on that day. While reports across the two methods agreed more often than might be expected by chance for any alcohol use on a given day, the methods agreed less than 70% of the time. On days when sex was reported via at least one of the methods, reports of anal sex and CAS agreed no more often than might be expected by chance, and about 45% and 51% of the time, respectively. Importantly, for most behaviors, agreement across the methods was substantially worse and likely below an acceptable level ($< 80\%$) for days that occurred 15–30 days before participants recalled them on Timeline. Agreement was improved (above this 80% threshold in most cases) for days that participants recalled no more than 14 days after they had passed. Two exceptions were days on which any anal sex or CAS had occurred, which agreed on 53% and 60% of days across methods for days that participants recalled < 14 days later. In the final week of the study period, however, agreement for these two behaviors rose to 62% and 83%, respectively.

DISCUSSION

Overall, this study’s results suggest that Timeline can yield accurate data about respondents’ total alcohol use, drug use, and sexual behavior during a 30-day recall window, and that these data are valid when compared to common survey instruments that require participants to report aggregated behaviors over broad time intervals. This is consistent with past studies showing the TLFB is broadly valid for assessing these risk behaviors when compared with standardized survey measures (4, 5, 8). As such, Timeline may provide a viable alternative to these survey measures in studies that require more precision and details about specific behaviors that have occurred over more constrained intervals of time. However, our findings showed participants tended to underreport behaviors on Timeline compared with a more intensive assessment method that minimizes recall biases (e.g., daily diary). While correspondence between the two methods was generally high for total reports of most behaviors (especially sexual behaviors), participants reported an average of 31% fewer

drinking days, 34% fewer “heavy” drinking days, and 24% fewer drug use days on the Timeline than they did via daily diary. These results are consistent with past studies showing that participants may underreport these behaviors on the TLFB when compared with daily assessments (17, 18, 20). Our results further suggest that Timeline may produce less accurate data about *timing* of behaviors when compared to daily assessment methods, especially when requiring participants to recall behaviors that occurred between 15 and 30 days in the past. This is consistent with other similar studies showing that day-level correspondence between alcohol, drug, and sexual behaviors reported on a TLFB and daily assessment methods was poor (21).

Together, these findings suggest researchers should consider several factors when determining whether using Timeline is appropriate. First, Timeline may be best suited for when researchers are interested in participants’ *overall involvement* in a given health behavior over a constrained time interval (≈ 30 days or fewer), rather than the *specific timing* of behaviors. For example, studies exploring the co-occurrence of multiple behaviors, such as alcohol/drug use with sexual risk behavior or dual use of alcohol and marijuana, might consider whether a more intensive method that minimizes recall bias is more appropriate, while weighing participant burden. For studies like these, a more intensive assessment method may provide a stronger degree of evidence for observed associations.

Second, our results suggest that choice of recall interval may affect report accuracy. Correspondence between Timeline and daily diary reports was improved when participants recalled their behavior 14 or fewer days after they occurred. This is consistent with previous research suggesting that repeated 7-day TLFB enhances accuracy of self-reported alcohol use compared to 30-day TLFB (33). Thus, researchers should minimize recall window length whenever possible, especially in studies of populations in which risk behaviors are high-frequency events (e.g., heavy drinkers, heavy drug users, those at high risk for HIV), since specific details of these behaviors might be more difficult to recall at broader intervals.

Finally, our findings suggest Timeline may be more appropriate for studies focusing on certain behaviors and outcomes. While correspondence between the two methods in participants’ overall engagement in various sexual behaviors, typical quantity of alcohol use, and drug use were reasonable, correspondence between total number of drinking days and “heavy” drinking days were more modest when assessed at 30-day intervals. As such, researchers might consider using more intensive methods for studies focused on “binge” drinking or how often participants drank, or choose a shorter recall interval.

Limitations

While this study has a number of key strengths, several limitations are important to note. First, the study involved a relatively small sample. While most participants reported a high frequency of each behavior of interest during the study period, conducting the same analyses with more participants would likely have produced more stable and certain results. Second, since we compared daily diary reports over 30 days with a Timeline completed at the end of that period, the correspondence between these two methods could be inflated because submitting daily reports may have improved recall on Timeline. Thus, the accuracy of Timeline may be lower without a daily diary procedure. However, participants were *not*

explicitly informed that they would be asked to recall these behaviors at the end of the 30-day period in an effort to minimize this effect. Finally, this study was focused on men who have sex with men recruited from dating apps and social media, so these findings may not generalize to other populations. However, focusing on a heavy drinking, high-risk subset of this population provides an especially stringent test of Timeline's validity: for these participants, alcohol use and sexual behaviors were inherently high-frequency events, meaning that a variety of risk behaviors occurred during a relatively brief study period (i.e., 30 days). Since recalling these events may be more difficult when they occur frequently, it is possible that accuracy might be improved among lower-risk populations and that these results represent the degree of accuracy expected with challenging recall. Future Timeline-based research should aim to establish its validity in a larger sample, its ability to capture valid data on behaviors beyond those presented, and focus on other populations of interest

CONCLUSION

Timeline is a flexible and interactive web application developed to collect detailed risk behavior data using an approach similar to the Timeline Followback. Results from a study of heavy drinking, high-risk MSM ($N = 15$) supported Timeline's validity for assessing alcohol/drug use, and sexual behavior when compared with popular survey measures. Findings suggested that Timeline yields overall estimates of these behaviors that are comparable to more intensive methods (e.g., daily diary) but that underreporting is higher using Timeline compared with daily diary methods. The day-level correspondence between these methods was poor, especially for days that occurred 14 days or more before recall. Researchers should consider these findings when determining whether Timeline is appropriate for their study goals.

Acknowledgements

This manuscript was supported by R21MH109374 from the National Institute on Mental Health, and R34AA023478, P01AA019072, K08AA024056, and L30AA023336 from the National Institute on Alcohol Abuse and Alcoholism.

APPENDIX

TABLE 1

Demographic and behavioral characteristics of the study sample ($N = 15$)

Characteristics	Mean (SD) or N (%)
Age (Range: 18 – 31, $M \pm SD$)	24.2 (3.8)
Race	
White	12 (80.0)
Asian	2 (13.3)
Multiracial	1 (6.7)
Ethnicity (Hispanic or Latino)	2 (13.3)
Currently in Exclusive Relationship ^a	1 (6.7)
College degree	10 (66.7)

Characteristics	Mean (SD) or N (%)
Low income ^b	2 (13.3)
Unemployed	0 (0.0)
Gay or bisexual identity	15 (100.0)
Alcohol-related problems (AUDIT ^c ≥ 8)	10 (66.7)
Other drug-related problems (DAST ^d > 3)	4 (26.7)
Injection drug use, past 3 months	0 (0.0)
Total # of sex events	7.1 (4.5)
Total number of condomless anal sex (CAS) events	2.1 (2.5)

Note.

^aRepresents participants who reported currently being in a sexually exclusive, monogamous relationship with one partner.

^bRepresents those with a household annual income <\$30,000/year.

^cAlcohol Use Disorders Identification Test (AUDIT).

^dDrug Abuse Screening Test (DAST-10).

TABLE 2

Bivariate correlations of Timeline variables with relevant individual-level measures of similar constructs.

Variable	Timeline Variable ^a	r
<i>Alcohol/Drug Use</i>		
Drinking frequency ^b	Drinking days	0.41 *
Drinking quantity ^b	Avg. drinks/drinking day	0.50 *
Frequency of heavy drinking ^b	“Heavy” drinking days	0.59 *
Drug problems (DAST total) ^c	Drug use days	0.83 *
<i>Sexual Behavior</i>		
# male partners ^d	Any sex days	0.34
# male anal sex partners ^d	Anal sex days	0.52 *
Condom use frequency ^{d,e}	CAS days	0.52 *

Note.

* $p < .05$.

^aRepresents the total count of days on which each event occurred, as assessed using Timeline.

^bAs assessed using the Alcohol Use Disorders Identification Test (AUDIT).

^cMeasured using the Drug Abuse Screening Test (DAST-10).

^dDuring the past year, as assessed using the Sexual Behavior Survey.

^eRepresents ratings of condom use frequency during insertive and/or receptive anal sex with male partners in the past year.

TABLE 3

Means and standard deviations of total alcohol/drug use and sexual behavior reported via Ecologic Momentary Assessment and Timeline Followback over 30 days

Variable	Daily Diary		Timeline		r ^a
	M	SD	M	SD	
<i>Alcohol/Drug Use</i>					
Drinking days	11.7	7.2	8.0	5.1	0.55

Variable	Daily Diary		Timeline		r^a
Avg. drinks/drinking day	4.7	1.8	4.7	2.3	0.88
Heavy drinking days	5.5	5.6	3.6	4.5	0.58
Drug use days	7.2	10.7	5.5	9.1	0.69
<i>Sexual Behavior</i>					
Any sex days	3.3	3.3	2.8	3.3	0.88
Anal sex days	2.3	2.2	1.8	2.4	0.87
Condomless anal sex days	1.3	1.8	1.2	2.1	0.92

^aAll values are $p < .05$.

TABLE 4

Absolute agreement and Fleiss' kappa for drug/alcohol use and sex events reported on specific days across daily diary and Timeline

Variable	Absolute agreement (%)	k	SE	p
<i>Drug/Alcohol use</i>				
Drinking day	69.5	0.34	.05	< .001
Weeks 1 & 2	62.9	0.24	.07	<.001
Weeks 3 & 4	75.9	0.44	.07	<.001
“Heavy” drinking day	80.5	0.28	.05	<.001
Weeks 1 & 2	74.3	0.19	.07	.003
Weeks 3 & 4	86.5	0.40	.06	<.001
Drug use day	81.2	0.46	.05	<.001
Weeks 1 & 2	79.5	0.42	.07	<.001
Weeks 3 & 4	82.9	0.50	.06	<.001
<i>Sexual behaviors</i>				
Any sex	84.5	0.20	.05	<.001
Weeks 1 & 2	84.3	0.18	.07	.003
Weeks 3 & 4	84.7	0.21	.07	.009
Anal sex^a	45.6	-0.14	.11	.895
Weeks 1 & 2	38.5	-0.22	.16	.932
Weeks 3 & 4	52.5	-0.03	.16	.568
Condomless anal sex (CAS)^a	51.0	-0.06	.14	.682
Weeks 1 & 2	42.9	-0.14	.18	.784
Weeks 3 & 4	60.0	0.02	.20	.468

^aReflects agreement about the type of sex that occurred on days when participants reported having sex via either method.

References

1. Sobell L, Sobell M. Timeline Follow-back: A technique for assessing self-reported alcohol consumption In: Litten R, Allen J, editors. Measuring alcohol consumption: Psychosocial and biological methods. Totowa, NJ: Humana Press; 1992 p. 41–72.
2. Sobell LC, Maisto SA, Sobell MB, Cooper AM. Reliability of alcohol abusers' self-reports of drinking behavior. Behaviour research and therapy. 1979;17(2):157–60. [PubMed: 426744]
3. Shiffman S, Stone AA, Hufford MR. Ecological Momentary Assessment. Annual Review of Clinical Psychology. 2008;4(1):1–32.

4. Carey MP, Carey K, Maisto S, Gordon C, Weinhardt L. Assessing sexual risk behaviour with the Timeline Followback (TLFB) approach: continued development and psychometric evaluation with psychiatric outpatients. *International journal of STD & AIDS*. 2001;12(6):365–75. [PubMed: 11368817]
5. Fals-Stewart W, O'Farrell TJ, Freitas TT, McFarlin SK, Rutigliano P. The timeline followback reports of psychoactive substance use by drug-abusing patients: psychometric properties. *Journal of consulting and clinical psychology*. 2000;68(1):134. [PubMed: 10710848]
6. Fishburne JW, Brown JM. How do college students estimate their drinking? Comparing consumption patterns among quantity-frequency, graduated frequency, and timeline follow-back methods. *Journal of Alcohol and Drug Education*. 2006;50(1):15.
7. Roy M, Dum M, Sobell LC, Sobell MB, Simco ER, Manor H, et al. Comparison of the quick drinking screen and the alcohol timeline followback with outpatient alcohol abusers. *Substance use & misuse*. 2008;43(14):2116–23. [PubMed: 18825590]
8. Carey KB. Reliability and validity of the time-line follow-back interview among psychiatric outpatients: A preliminary report. *Psychology of Addictive Behaviors*. 1997;11(1):26.
9. Robinson SM, Sobell LC, Sobell MB, Leo GI. Reliability of the Timeline Followback for cocaine, cannabis, and cigarette use. *Psychology of addictive behaviors*. 2014;28(1):154. [PubMed: 23276315]
10. Weinhardt LS, Carey MP, Maisto SA, Carey KB, Cohen MM, Wickramasinghe SM. Reliability of the timeline follow-back sexual behavior interview. *Annals of Behavioral Medicine*. 1998;20(1): 25–30. [PubMed: 9755348]
11. Carey KB, Carey MP, Maisto SA, Henson JM. Temporal stability of the timeline followback interview for alcohol and drug use with psychiatric outpatients. *Journal of studies on alcohol*. 2004;65(6):774–81. [PubMed: 15700516]
12. Hjorthøj CR, Hjorthøj AR, Nordentoft M. Validity of timeline follow-back for self-reported use of cannabis and other illicit substances—systematic review and meta-analysis. *Addictive behaviors*. 2012;37(3):225–33. [PubMed: 22143002]
13. Hjorthøj CR, Fohlmann A, Larsen AM, Arendt M, Nordentoft M. Correlations and agreement between delta-9-tetrahydrocannabinol (THC) in blood plasma and timeline follow-back (TLFB)-assisted self-reported use of cannabis of patients with cannabis use disorder and psychotic illness attending the CapOplus randomized clinical trial. *Addiction*. 2012;107(6):1123–31. [PubMed: 22151583]
14. Simons JS, Wills TA, Emery NN, Marks RM. Quantifying alcohol consumption: self-report, transdermal assessment, and prediction of dependence symptoms. *Addictive behaviors*. 2015;50:205–12. [PubMed: 26160523]
15. Tucker JA, Foushee HR, Black BC, Roth DL. Agreement between prospective interactive voice response self-monitoring and structured retrospective reports of drinking and contextual variables during natural resolution attempts. *Journal of studies on alcohol and drugs*. 2007;68(4):538–42. [PubMed: 17568958]
16. Carney MA, Tennen H, Affleck G, Del Boca FK, Kranzler HR. Levels and patterns of alcohol consumption using timeline follow-back, daily diaries and real-time" electronic interviews". *Journal of studies on alcohol*. 1998;59(4):447–54. [PubMed: 9647427]
17. Searles JS, Helzer JE, Walter DE. Comparison of drinking patterns measured by daily reports and timeline follow back. *Psychology of Addictive Behaviors*. 2000;14(3):277. [PubMed: 10998953]
18. Searles JS, Helzer JE, Rose GL, Badger GJ. Concurrent and retrospective reports of alcohol consumption across 30, 90 and 366 days: interactive voice response compared with the timeline follow back. *Journal of studies on alcohol*. 2002;63(3):352–62. [PubMed: 12086136]
19. Simpson CA, Xie L, Blum ER, Tucker JA. Agreement between prospective interactive voice response telephone reporting and structured recall reports of risk behaviors in rural substance users living with HIV/AIDS. *Psychology of Addictive Behaviors*. 2011;25(1):185. [PubMed: 21443312]
20. Schroder KE, Johnson CJ, Wiebe JS. Interactive voice response technology applied to sexual behavior self-reports: A comparison of three methods. *AIDS and Behavior*. 2007;11(2):313–23. [PubMed: 16841191]

21. Wray TB, Kahler CW, Monti PM. Using ecological momentary assessment (EMA) to study sex events among very high-risk men who have sex with men (MSM). *AIDS and Behavior*. 2016;20(10):2231–42. [PubMed: 26746212]
22. Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL. Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. *Science*. 1998;280(5365):867–73. [PubMed: 9572724]
23. Dolezal C, Marhefka SL, Santamaria EK, Leu C-S, Brackis-Cott E, Mellins CA. A comparison of audio computer-assisted self-interviews to face-to-face interviews of sexual behavior among perinatally HIV-exposed youth. *Archives of sexual behavior*. 2012;41(2):401–10. [PubMed: 21604065]
24. Maisto SA, Conigliaro JC, Gordon AJ, McGinnis KA, Justice AC. An experimental study of the agreement of self-administration and telephone administration of the Timeline Followback interview. *Journal of studies on alcohol and drugs*. 2008;69(3):468–71. [PubMed: 18432391]
25. Rueger SY, Trela CJ, Palmeri M, King AC. Self-administered web-based timeline followback procedure for drinking and smoking behaviors in young adults. *Journal of Studies on Alcohol and Drugs*. 2012;73(5):829–33. [PubMed: 22846247]
26. Sobell LC, Brown J, Leo GI, Sobell MB. The reliability of the Alcohol Timeline Followback when administered by telephone and by computer. *Drug and alcohol dependence*. 1996;42(1):49–54. [PubMed: 8889403]
27. Saunders JB, Aasland OG, Babor TF, De la Fuente JR, Grant M. Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption—II. *Addiction*. 1993;88(6):791–804. [PubMed: 8329970]
28. de Meneses-Gaya C, Zuardi AW, Loureiro SR, Crippa JAS. Alcohol Use Disorders Identification Test (AUDIT): An updated systematic review of psychometric properties. *Psychology & Neuroscience*. 2009;2(1):83.
29. Skinner HA. The drug abuse screening test. *Addictive behaviors*. 1982;7(4):363–71. [PubMed: 7183189]
30. Yudko E, Lozhkina O, Fouts A. A comprehensive review of the psychometric properties of the Drug Abuse Screening Test. *Journal of substance abuse treatment*. 2007;32(2):189–98. [PubMed: 17306727]
31. Gordon CM, Carey MP, Carey KB. Effects of a drinking event on behavioral skills and condom attitudes in men: implications for HIV risk from a controlled experiment. *Health Psychology*. 1997;16(5):490. [PubMed: 9302547]
32. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychological bulletin*. 1971;76(5):378.
33. Hoepfner BB, Stout RL, Jackson KM, Barnett NP. How good is fine-grained Timeline Followback data? Comparing 30-day TLFB and repeated 7-day TLFB alcohol consumption reports on the person and daily level. *Addictive behaviors*. 2010;35(12):1138–43. [PubMed: 20822852]

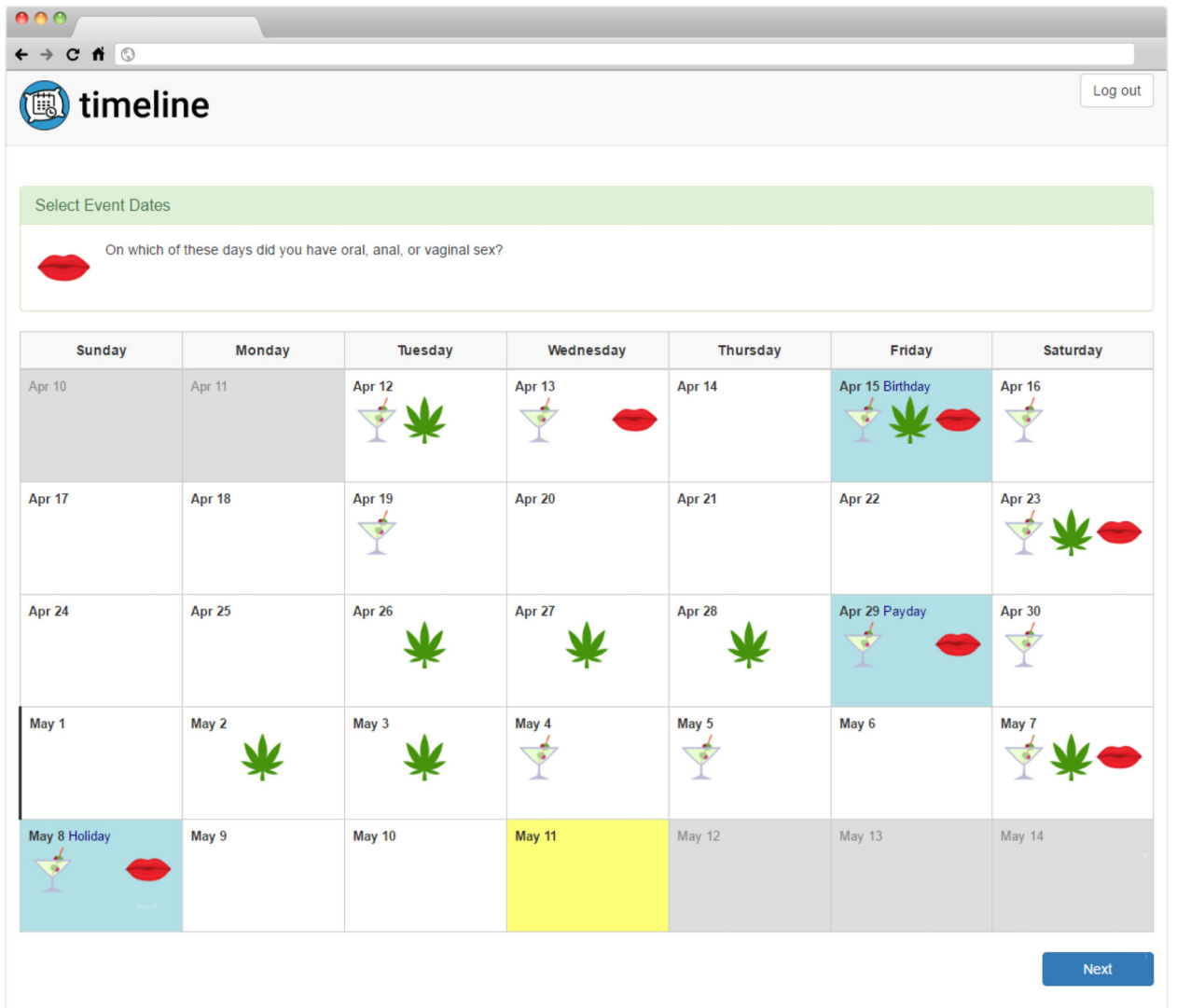


Figure 1.
Timeline calendar view screen.

The screenshot shows a web browser window with the 'timeline' application. The interface includes a calendar on the left, a header with the 'timeline' logo and a 'Log out' button, and a main content area for 'Day 27 • May 8th, 2016'. The calendar highlights the 8th of May. The survey questions are as follows:

1.) How many standard drinks did you have?

2.) Over how many **hours** were you drinking (from the first drink you had to your last)?

One Standard Drink Equals...

- Beer = 12 oz.
- Wine = 5 oz.
- Liquor = 1 - 1.5 oz.

3.) How many partners did you engage in oral, anal, or vaginal sex with on this day?
 1 2 3 4

For partner 1...

3a.) Was this the first time you had sex with this partner?
 Yes No

3b.) Was this person a casual or a committed sex partner? (A committed partner is someone with whom you have a commitment to have sex with only that person. A casual partner is someone you have little romantic commitment to and haven't made a commitment to have sex with only that person.)
 Casual Committed

Figure 2.
 Timeline detail view screen.

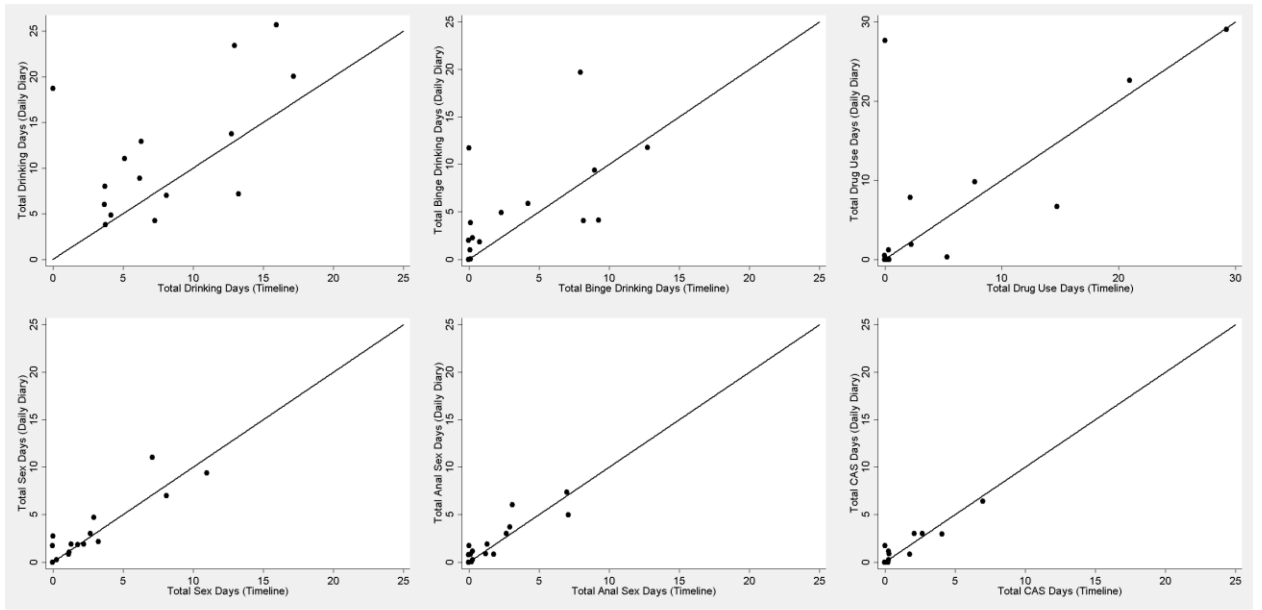


Figure 3.
Scatterplots of the total number of alcohol/drug use and sex days reported via daily diary versus Timeline