Assessment methods and schedules for collecting daily process data on substance use related health behaviors: A randomized control study

Anne Buu[a], **Lynn S. Massey**[b], **Maureen A. Walton**[b], **James A. Cranford**[b], **Marc A. Zimmerman**[c], and **Rebecca M. Cunningham**[d]

[a]Department of Health Behavior and Biological Sciences, University of Michigan, 400 North Ingalls, Ann Arbor, MI 48109, USA

[b]Addiction Research Center, Department of Psychiatry, University of Michigan, 4250 Plymouth Road, Ann Arbor, MI 48109, USA

[c]Department of Health Behavior and Health Education, University of Michigan, 1415 Washington Heights, Ann Arbor, MI 48109

[d]Department of Emergency Medicine, University of Michigan, 2800 Plymouth Rd, Bldg 10-G080, Ann Arbor, MI 48109

Abstract

Background—Interactive voice response (IVR) and short message service (SMS) systems have been used to collect daily process data on substance use. Yet, their relative compliance, use patterns, and user experiences are unknown. Furthermore, recent studies presented the potential of a hybrid weekly protocol requiring recall of behaviors in past week right after the weekend, in order to reduce the concerns about low compliance and measurement reactivity associated with daily data collection and also provide high quality data on the peak of use.

Methods—This study randomized substance users to four (2×2) assessment groups with different combinations of assessment methods (IVR or SMS) and schedules (daily or weekly). The compliance rates and use patterns during the experimental period of 90 days and user experiences reported after the period were compared across the groups.

Correspondence: Anne Buu, 400 North Ingalls, Ann Arbor, MI 48109, USA, Tel.: +1 (734) 764-9008, fax: +1(734) 647-2416, buu@umich.edu.

Results—When IVR was assigned, the weekly schedule generated a higher compliance rate than the daily schedule. When SMS was used, however, the assessment schedule did not have an effect on compliance. While both the daily and weekly surveys via IVR can be completed within a short time, the weekly survey administered via SMS took much longer than its daily counterpart. Such an increased time consumption may offset the benefit of a less frequent assessment schedule.

Conclusions—IVR is a better choice for delivering the hybrid protocol of weekly collection of daily process data because of its higher compliance rate, shorter duration, and lower likelihood of interruption during data collection.

1. Introduction

Research studies with daily process designs involving data collection from participants once per day over a defined period have increased dramatically in the last decade (Gunthert and Wenze, 2012). These designs have the advantage of eliminating retrospection bias and minimizing selectivity in describing experiences. More importantly, they have greater ecological validity because behavioral processes are assessed in real time and in their natural contexts (Reis, 2012). For example, alcohol consumption usually takes place in social settings with important antecedents and consequences such as moods and marital interactions, which can be effectively captured by daily process data (Cranford et al., 2010). Because of these advantages, such designs have been adopted more frequently in substance use research.

In spite of these advantages, daily process designs require a higher cost and heavier participant burden than retrospective interviews or surveys. Yet, such costs and burdens could be reduced by collecting data using participants' own mobile phones through interactive voice response (IVR) or short message service (SMS) systems (Conner and Lehman, 2012). IVR systems administering surveys with prerecorded audio and recording participants' responses into databases automatically have been commonly adopted to collect daily process data in the substance abuse field (Yang et al., 2015). Recently, SMS has also become a popular research tool because of its popularity (Suffoletto et al., 2012). Researchers, however, have not conducted a randomized control study comparing these two assessment methods in terms of compliance, use patterns, and user experiences. This is a critical gap in the literature because such comparison can inform future applications of these methods in various settings.

Daily process designs unavoidably involve self-monitoring of the target behavior which is an active component of some cognitive-behavioral interventions for substance use disorders (Simpson et al., 2005). The potential measurement reactivity (i.e., reducing the target behavior due to self-awareness) is undesirable for studies that aim to investigate the association between the target behavior and its antecedent or consequence (Yang et al., 2015). Another major drawback of daily process designs is low compliance that tends to

result in nonrandom missing data and biased samples (Leigh, 2000). A possible way to address both the issues of measurement reactivity and low compliance is to implement a less intensive assessment schedule. Simpson et. al. (2005) randomized a treatment sample into daily or weekly IVR monitoring and found no significant difference in the percent of calls made. Yet, such findings were limited by the short study duration (28 days) and an ongoing addiction treatment that may have promoted compliance in both groups. Another study with a small community sample completing both daily and weekly IVR for 128 days found a high correlation between the two reports of drinking (Tucker et al., 2007). The correlation, however, was likely to be inflated because the daily protocol may have facilitated the recall in the weekly protocol. Moreover, a feasibility study using SMS to collect data from young adults (Kuntsche and Robert, 2009) only on Saturday and Sunday afternoons (to minimize participant burden but still enable the maximum capture of high-risk drinking that usually occurs on Friday and Saturday nights) was able to reach a retention rate of 75% over 4 weekends. Taken together, previous studies presented the potential of a *hybrid protocol* that requires recall of behaviors in past 7 days right after the weekend, to reduce concerns about low compliance and measurement reactivity associated with daily data collection and also provide high quality data on the peak of use (weekend). Furthermore, it is unknown whether the differences in compliance, use patterns, and user experiences between daily and weekly assessment schedules vary across assessment methods (IVR vs. SMS).

This study aims to address the current knowledge gaps by randomizing substance users to four (2×2) assessments groups with different combinations of assessment methods (IVR or SMS) and schedules (daily or weekly). The compliance rates and use patterns during the experimental period of 90 days and user experiences reported after the period are compared across the groups. The results have important implications for designing future studies that collect daily process data on substance use related health behaviors.

2. Material and Methods

2.1. Study Sample and Procedures

This study is a randomized control study that re-contacted participants who previously enrolled in a natural history study, the Flint Youth Injury (FYI) Study, of 14 to 24 year-olds with recent drug use who sought care in an Emergency Department in Flint, Michigan (see Bohnert et al., 2015). Study procedures were approved and conducted in compliance with the Institutional Review Boards for the University of Michigan and Hurley Medical Center. A Certificate of Confidentiality was also obtained from the National Institutes of Health.

The recruitment period was from March 2014 to January 2016. Of the 600 subject pool, 103 were excluded because they did not agree to be re-contacted for future studies, were in jail, or had died. Remaining participants were sent a "welcome back postcard" and contacted (e.g., phone, home visit, social media) during the recruitment period. After providing consent for the daily process study, 331 participants self-administered a 30-minute computerized assessment including demographic information and conventional measures of substance use related risk behaviors/problems in past six months, followed by a 20–30 minute staff-administered timeline follow-back interview, which used a calendar and landmark events to facilitate participants' recall of substance use related behaviors for each

day in the past 90 days (see Buu et al., 2014). Participants received $20 cash for completing the baseline assessment and also offered the options to participate in urine drug screening ($5 cash) and HIV testing ($5 cash). After the assessment, they were randomized into four experimental groups (IVR daily, IVR weekly, SMS daily, and SMS weekly) and received a 10-minute training session for the assigned group.

Participants in the daily groups reported daily by IVR/SMS about behaviors on the previous day for 90 days. The weekly groups retrospectively reported about their behaviors in the previous 7 days on Sunday or Monday after the baseline; for those whose baseline was on a Sunday, Monday, or Tuesday, the duration was 13 weeks, whereas the others had the duration of 14 weeks. This protocol ensured that the IVR/SMS data collection fully covered the 90 days after baseline for both the daily and weekly groups. All participants were instructed to call or text the computer system to take a short survey between 8 am and 11:59 pm using their own mobile device, to ensure that they had sufficient time to complete the assessment before the data collection was closed at 1 am. The IVR/SMS system automatically sent a call or text reminder for those who had not completed their survey at 2 pm daily (or on Sunday for the weekly group). Participants had the option to take their survey from that call or text message. Research staff monitored compliance and contacted participants with incomplete Sunday surveys via phone call, text, email, and/or Facebook messaging on Monday to remind them to complete it before 11:59 pm on Mondays; the staff also verified there were no technical issues that needed to be addressed. For participants in the daily groups, the staff contacted them after missing two consecutive surveys. Non-compliant participants were contacted 2–3 times per week using the same methods described above. After the experimental period ended, a post assessment was conducted with a brief conventional measure of substance use related risk behaviors in past 90 days and a satisfaction questionnaire about participants' experiences with IVR/SMS (both were self-administered), followed by a staff-administered 90-day timeline follow-back interview. Participation in the post assessment received $25 cash and were offered the option to participate in urine drug screening ($5 cash).

The original subject payment for daily IVR/SMS was $1 per survey with an extra $10 per month for completing 75% of their daily surveys; the one for weekly IVR/SMS was $7 per survey with an extra $10 per month for completing 75% of their weekly surveys. This payment structure applied to the 87 participants recruited during March 2014 to September 2014 (Cohort 1). Due to concerns about low compliance, our research team later implemented a higher payment structure for the 244 participants recruited during January 2015 to January 2016 (Cohort 2). The new subject payment was $4 per daily survey and $27 per weekly survey (the extra payment for 75% compliance was dropped). We increased to this amount based on feedback from participants in Cohort 1, given the length of time required to complete assessments. In order to investigate the effect of this change, we conducted an additional set of analysis to compare these two cohorts in terms of their compliance rates, use patterns, and user experiences.

Among the 331 participants recruited to participate in IVR/SMS data collection, 24 people have never provided any data during the experimental period and thus were not included in the analysis. They were not significantly different (p>0.05) from the rest of the 307

participants on the average age, race (e.g., percentage of Black), percentage of smokers, percentage of drinkers, percentage of marijuana users, percentage of illicit drug users, and percentage of prescription drug misusers. They did have a higher percentage of males (79% vs. 50%; p<.01) and a lower percentage receiving public assistance (39% vs. 66%; p<.05). Among the 307 participants included in statistical analysis, 81 were assigned to the IVR daily group, 76 to the IVR weekly, 81 to the SMS daily, and 69 to the SMS weekly.

2.2. Measures

2.2.1 Baseline assessment—The baseline assessment consisted of demographic information and conventional measures of substance use (Chung et al., 2000; Humeniuk et al., 2008; Saunders et al., 1993), violence (Straus et al., 1996), and sexual risk behaviors (Bearman and Jones, 1997; Darke et al., 1991; Ward et al., 1990). For each substance (e.g., alcohol), questions included quantity, frequency, and problems in past 6 months, with ordinal response scales. Participants were categorize as current users or nonusers of: alcohol, nicotine, marijuana, other illicit drugs (cocaine, methamphetamine, inhalants, hallucinogens, or street opioids), and prescription drugs (misusing simulants, sedatives, or opioids).

2.2.2 IVR/SMS assessment—IVR questions under the same assessment schedule were exactly the same as those for SMS which had the limitation of 160 characters per question. The daily survey asked about the behaviors yesterday with the total of 19 items modified from the conventional measures in the baseline assessment: one item for alcohol use, one item for marijuana use, one item for illicit drug use, 3 items for prescription drug use, 8 items for fighting, and 5 items for sexual behaviors. If the participant reported no fighting or sexual behaviors (the first item under each domain), they were redirected to the same amount of follow-up questions about physical activity and diet (i.e., 11 filler questions), in order to prevent them from underreporting for the purpose of saving time. In the weekly survey, for each item of the daily survey, the participants were asked to recall their behaviors during the past week from Saturday to Sunday (backward). This resulted in the total of 133 questions (with 14 filler questions). Participants responded to each question with numbers (e.g., l="yes"; 0="no"). In addition to responses, the system recorded the date, time, duration to complete the survey, the number of times logging in the system, and whether they completed the entire survey.

2.2.3 Post assessment—The post assessment was designed to serve two purposes: (1) the conventional measure with items mirroring the IVR/SMS questions could provide data to examine the validity of retrospective reports of health risk behaviors in past 90 days, using the daily/weekly data as gold standards; and (2) the satisfaction survey (Stone et al., 2003; Lim et al., 2010) generated data for evaluating participants' experiences with the assessment methods and schedules during the experimental period. The satisfaction survey had 2 items on ease of use $(a = 0.79)$, 4 items on privacy concern $(a = 0.88)$, 3 items on burden of participation $(a = 0.33)$, 5 items on difficulty in remembering $(a = 0.69)$, and 4 items on embarrassment about questions $(a = 0.85)$. All the questions were responded with a 5-point ordinal scale (from 1="not at all" to 5="extremely"). The average score across all the items under each domain was used as the composite score.

2.3. Statistical Analysis

The four experimental groups were compared on demographic information, substance use at baseline, compliance rates, use patterns, and user experiences. When the outcome variable was continuous, an F test of ANOVA was conducted. If the overall test was significant (p<.05), the Tukey's Studentized Range Test controlling the Type I experimentwise error rate at 0.05 was adopted for pairwise comparison. When the outcome variable was binary, an overall chi-squared test (or Fisher's exact test if some cells had expected counts less than 5) was conducted first; if it was significant, pairwise comparisons were carried out using a Bonferroni correction to control the Type I experimentwise error rate at 0.05. Furthermore, an additional set of analysis was conducted to compare Cohort 1 and Cohort 2 on all the outcome variables in order to examine the effects of increasing subject payment; an independent sample *t* test was used for continuous outcomes whereas a chi-squared test (or Fisher's exact test) was adopted for binary outcomes.

3. Results

3.1. Demographics and Substance Use

The 307 people participating in IVR/SMS assessment were on average 24 years old (range=18–29) with the median annual household income of $0-4,999. About 50% of them were male; 60% Black; 26% White; and 66% under public assistance. Based on self-report substance use in past 6 months at baseline, 69% used alcohol, 67% used nicotine, 73% used marijuana, 12% used other illicit drugs, and 18% misused prescription drugs. The four experimental groups were not significantly different on baseline demographic or substance use variables (see Table 1). We also compared the two cohorts on these variables and found they were only significantly different on alcohol use: 81% for Cohort 1; and 64% for Cohort 2 (p<.01).

We investigated whether the four groups differed on self-reported substance use after the experimental period. At post assessment, the quantity and frequency of both alcohol and marijuana use in the past 3 months were inquired. For each of these 4 outcomes, we conducted an ANCOVA analysis using the group assignment as the predictor while controlling for the corresponding consumption variable at baseline; no differences were found on any of the outcomes (p>.05).

3.2. Group Differences in Compliance

The compliance rate for each participant was calculated as the number of calls/texts actually made divided by the total number supposed to be completed for the group that he/she was assigned to. Table 2 shows that the average compliance rate was 71% for IVR weekly, followed by 65% for SMS weekly, 57% for SMS daily, and 50% for IVR daily. The distributions of compliance rates by experimental groups are also displayed in a side-by-side box plot in Figure 1. The boxes covered the first quartile (the bottom line) to the third quantile (the top line), with the median indicated by the central line. Because the box lengths were approximately equal, the 4 experimental groups had about the same variance. In comparison to the other groups, the IVR weekly group had an outstanding performance with the majority of participants having compliance rates higher than 80% (the median was over

80%; the third quartile and the maximum were both at the 100%). Based on the finding of multiple comparison, when the assessment schedule was fixed, we found no difference between the IVR and SMS groups. Further, when IVR was assigned, the compliance rate for the weekly group was significantly higher than that for the daily group. Yet, when SMS was used, the assessment schedule did not have an effect on compliance. Moreover, the proportions of participants who had never provided any IVR/SMS data were compared across the four groups. The only significant difference was between SMS weekly (15%) and IVR daily (1%).

The effect of increasing subject payment on compliance was also examined. The compliance rate for Cohort 1 was 48% (SD=33%) and that for Cohort 2 was 65% (SD=29%). The t test result indicated a significant increase in compliance (t=4.34, df=305, p<.0001), but the two cohorts did not differ on the proportion of participants who had never provided any IVR/SMS data (X^2=1.58, df=1, p>.05). Furthermore, we examined the interaction between payment condition and assessment condition by conducting a three-way ANOVA, but we did not find any significant interactions (i.e., cohort× method and cohort× schedule; p>>.05).

Given the lengthy 90-day experimental period, we generated a figure to show the temporal change in average compliance across the 12 weeks by the two daily groups. Figure 2 shows that the SMS group maintained a higher average compliance rate than the IVR group within each week. Both groups had the highest rates of decrease in average compliance before Week 5. From Week 5 to Week 10, both groups only decreased slightly. However, between Week 10 and Week 12, the IVR group declined rapidly again whereas the SMS group remained stable. Further, the SMS group was able to persist with over 50% of compliance by Week 9 but the IVR group was only able to achieve that by Week 5. Such week-by-week comparison was, however, not feasible for the weekly groups because they were only assessed once per week. Instead, their monthly compliance rates were calculated. For the IVR weekly group, the average compliance rates from the first month to the third month were 84%, 69%, and 63%, respectively; for the SMS weekly group, the corresponding rates were 79%, 61%, and 57%.

3.3. Group Differences in Use Patterns

Table 2 shows the results of group comparison on use patterns. We used the median of the values from all the calls/texts made by the participant as his/her measure. The average duration to complete the survey varied significantly across groups with 50 minutes for SMS weekly, followed by 15 minutes for SMS daily, 7 minutes for IVR weekly, and 2 minutes for IVR daily. Relatedly, SMS weekly had a lower rate of completion in a single login (75%) than all the other groups (97–100%). SMS weekly also had a lower percentage of complete assessment (91%) than IVR daily (100%). Further, IVR weekly was more likely to take the survey during evening time (17%) than IVR daily (4%). Additionally, IVR daily had a higher tendency to take the survey by responding to reminder calls than the two SMS groups. Moreover, no cohort effect was found on any of the variables of use patterns.

3.4. Group Differences in User Experiences

Table 2 shows the means of user experiences (score range: 1–5) for each of the experimental groups in 5 domains: ease of use, privacy concern, burden of participation, difficulty in remembering, and embarrassment about questions. Across the groups, participants found that the technology was quite easy to use (3.81–4.28) and they had almost no concern about privacy (1.02–1.14), low burden of participation (1.36–1.57), very little difficulty in remembering their behaviors (1.14–1.21), and almost no embarrassment about answering the questions (1.05–1.13). We did not find group differences in any of the domains (p>.05). Furthermore, the cohort effect on user experiences was examined by conducting t tests on the 5 composite scores. The only significant difference (p<0.01) was that Cohort 2 was more concerned about privacy (M=1.12; SD=0.41) than Cohort 1 (M=1.02; SD=0.06).

4. Discussion

The results showed that when IVR was assigned, the weekly schedule generated a higher compliance rate than the daily schedule. However, when SMS was used, the assessment schedule did not have an effect on compliance. This may be explained by the group differences in use patterns. While both the daily and weekly surveys via IVR can be completed within less than 10 minutes, the weekly survey administered via SMS took much longer than its daily counterpart (50 vs. 15 minutes). Such an increased time consumption may offset the benefit of a less frequent assessment schedule. Unlike IVR that required continuous attention and responses to the prerecorded audio, SMS allowed delays between receiving and sending texts and, thus, encouraged multi-tasking while taking the survey. This was particularly an issue for SMS weekly that involved 133 questions. The higher likelihood of interruption during data collection for this particular group was also demonstrated by its lower percentage of completion within a single login attempt (75%), as well as its lower rate of completing the entire survey (91%). Furthermore, probably because IVR required constant interactions with the computer system, the weekly survey was more likely to be taken during evening time than the daily survey (17% vs. 4 %) when the participant had more free time. The IVR groups also had a higher tendency to take the survey during the reminder call in comparison to the SMS groups.

This is the first randomized control study that examined the effects of assessment methods (IVR or SMS) and schedules (daily or weekly) on compliance, use patterns and user experiences. The study was designed to test the feasibility of the *hybrid protocol* that requires recall of substance related health behaviors weekly right after the weekend (the peak of use). Although this paper has provided some evidence to support the feasibility of adopting the hybrid protocol of weekly collection of daily process data to promote higher compliance, our analysis did not evaluate the *quality* of the weekly reports on substance use related health behaviors. Such investigation requires sophisticated analysis on the daily process data. Another study limitation is that our investigation did not include web-based daily protocols (e.g., text message with a web-link to a survey) which has yielded higher compliance (e.g., Testa et al., 2015). Such protocols, however, require ownership of a smart phone or access to internet and thus may not be practical for low-income participants like our study sample.

The research team's decision to increase subject payment during the course of recruitment may have potentially influenced participants' behaviors other than compliance, but the analysis comparing the two cohorts did not find any differences on use patterns or demographic variables. On the other hand, this change provided the opportunity to examine whether increasing subject payment promoted higher compliance. The result of this study supported this hypothesis. Simpson et al. (2005) compared two small samples receiving different levels of incentive but did not find differences in compliance. These contradictive results may be explained by the magnitude of change/difference in payment – our study adopted a more drastic change. Although compliance was better with the larger payment schedule, this change was confounded with the removal of the bonuses. Future studies randomizing increasing incentive amount could further example the cost and benefits of various payment amounts and schedules, including bonuses. In many settings, however, increasing payment is not feasible. A recent study (Lindsay et al., 2014) examined the feasibility of a more cost-effective approach to promote higher compliance: the prize-based contingency-management (CM) approach. The study compared two groups of participants enrolled in a cocaine treatment study requiring daily IVR calls, with one group earning $1 for each call and the other earning one draw per call from a "prize bowl" with varying awards. The study found that odds of calling were 4.7 times greater in the prize-CM group than in the fixed dollar CM group, providing one strategy for boosting compliance when the funding is limited.

In conclusion, this study makes a unique contribution to the literature on daily process of substance use by conducting the first randomized control study to examine the effects of assessment methods and schedules on compliance, use patterns and user experiences. The finding indicates that IVR is a better choice for delivering the hybrid protocol of weekly collection of daily process data because of its higher compliance rate, shorter duration, and lower likelihood of interruption during data collection, in comparison to SMS among low-income participants. When daily data collection is feasible or desirable (e.g., delivering intervention), IVR is also recommended because of its shorter duration and higher probability of capturing participants' responses during reminder calls, particularly among low-income samples who may have less reliable SMS systems.

Acknowledgments

References

Bohnert KM, Walton MA, Ranney M, Bonar EE, Blow FC, Zimmerman MA, Booth BM, Cunningham RM. Understanding the service needs of assault-injured, drug-using youth presenting for care in an urban Emergency Department. Addict Behav. 2015; 41:97–105. [PubMed: 25452051]

Buu A, Li R, Walton MA, Yang H, Zimmerman MA, Cunningham RM. Changes in substance use-related health risk behaviors on the Timeline Follow-back interview as a function of length of recall period. Subst Use Misuse. 2014; 49:1259–1269. [PubMed: 24601785]

Chung T, Colby SM, Barnett NP, Rohsenow DJ, Spirito A, Monti PM. Screening adolescents for problem drinking: performance of brief screens against DSM-IV alcohol diagnoses. J Stud Alcohol. 2000; 61:579–587. [PubMed: 10928728]

Conner, TS., Mehl, MR. Handbook of research methods for studying daily life. Guilford Press; New York, NY, USA: 2012.

Cranford JA, Tennen H, Zucker RA. Feasibility of using interactive voice response to monitor daily drinking, moods, and relationship processes on a daily basis in alcoholic couples. Alcohol Clin Exp Res. 2010; 34:499–508. [PubMed: 20028351]

Darke S, Hall W, Heather N, Ward J, Wodak A. The reliability and validity of a scale to measure HIV risk-taking behaviour among intravenous drug users. AIDS. 1991; 5:181–186. [PubMed: 2031690]

Gunthert KC, Wenze SJ. Daily Diary Methods. Handbook Of Research Methods For Studying Daily Life. Guilford Press New York, NY, USA. 2012:144–159.

Harris, KM., Halpern, CT., Whitsel, E., Hussey, J., Tabor, J., Entzel, P., Udry, JR. The national longitudinal study of adolescent health: Research design. 2009. Available at http://www.cpc.unc.edu/projects/addhealth/design

Humeniuk R, Ali R, Babor TF, Farrell M, Formigoni ML, Jittiwutikarn J, De Lacerda RB, Ling W, Marsden J, Monteiro M, Nhiwatiwa S, Pal H, Poznyak V, Simon S. Validation of the alcohol, smoking and substance involvement screening test (ASSIST). Addiction. 2008; 103:1039–1047. [PubMed: 18373724]

Kuntsche E, Robert B. Short message service (SMS) technology in alcohol research— A feasibility study. Alcohol Alcohol. 2009; 44:423–428. [PubMed: 19482879]

Leigh BC. Using daily reports to measure drinking and drinking patterns. J Subst Abuse. 2000; 12:51–65. [PubMed: 11288474]

Lim MSC, Sacks-Davis R, Aitken CK, Hocking JS, Hellard ME. Randomised controlled trial of paper, online and SMS diaries for collecting sexual behaviour information from young people. J Epidemiol Community Health. 2010; 64:885–889. [PubMed: 19767322]

Lindsay JA, Minard CG, Hudson S, Green CE, Schmitz JM. Using prize-based incentives to enhance daily interactive voice response (IVR) compliance: A feasibility study. J Subst Abuse Treat. 2014; 46:74–77. [PubMed: 24029622]

Reis, HT. Handbook Of Research Methods For Studying Daily Life. Guilford Press; New York, NY, USA: 2012. Why Researchers Should Think "Real-World": A Conceptual Rationale; p. 3-21.

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; 88:791–804. [PubMed: 8329970]

Simpson TL, Kivlahan DR, Bush KR, McFall ME. Telephone self-monitoring among alcohol use disorder patients in early recovery: A randomized study of feasibility and measurement reactivity. Drug Alcohol Depend. 2005; 79:241–250. [PubMed: 16002033]

Stone AA, Broderick JE, Schwartz JE, Shiffman S, Litcher-Kelly L, Calvanese P. Intensive momentary reporting of pain with an electronic diary: Reactivity, compliance, and patient satisfaction. Pain. 2003; 104:343–351. [PubMed: 12855344]

Straus MA, Hamby SL, Boney-McCoy S, Sugarman DB. The revised conflict tactics scales (CTS2) development and preliminary psychometric data. J Fam Issues. 1996; 17:283–316.

Suffoletto B, Callaway C, Kristan J, Kraemer K, Clark DB. Text-message-based drinking assessments and brief interventions for young adults discharged from the emergency department. Alcohol Clin Exp Res. 2012; 36:552–560. [PubMed: 22168137]

Testa M, Parks KA, Hoffman JH, Crane CA, Leonard KE, Shyhalla K. Do drinking episodes contribute to sexual aggression perpetration in college men? J Stud Alcohol Drugs. 2015; 76:507–515. [PubMed: 26098025]

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. J Stud Alcohol Drugs. 2007; 68:538–542. [PubMed: 17568958]

Ward, J., Darke, S., Hall, W. National Drug and Alcohol Research Centre. University of New South Wales Sydney; 1990. The HIV risk-taking behaviour scale (HRBS) manual.

Yang H, Cranford JA, Li R, Buu A. Two-stage model for time-varying effects of discrete longitudinal covariates with applications in analysis of daily process data. Stat Med. 2015; 34:571–581. [PubMed: 25395042]

Highlights

- The weekly schedule had a higher compliance rate than the daily one for Interactive voice response (IVR)

- The assessment schedule did not have an effect on compliance for short message service (SMS)

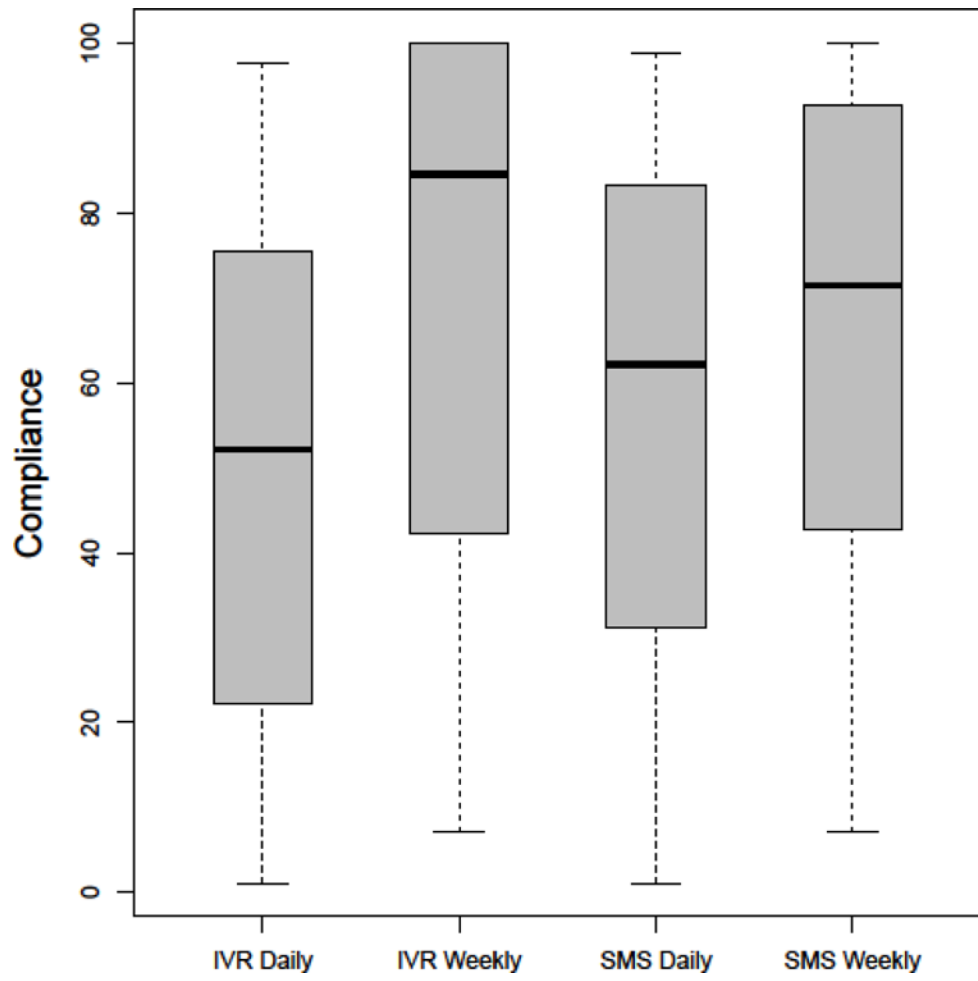- IVR is recommended for the hybrid weekly collection of daily process data

Figure 1.
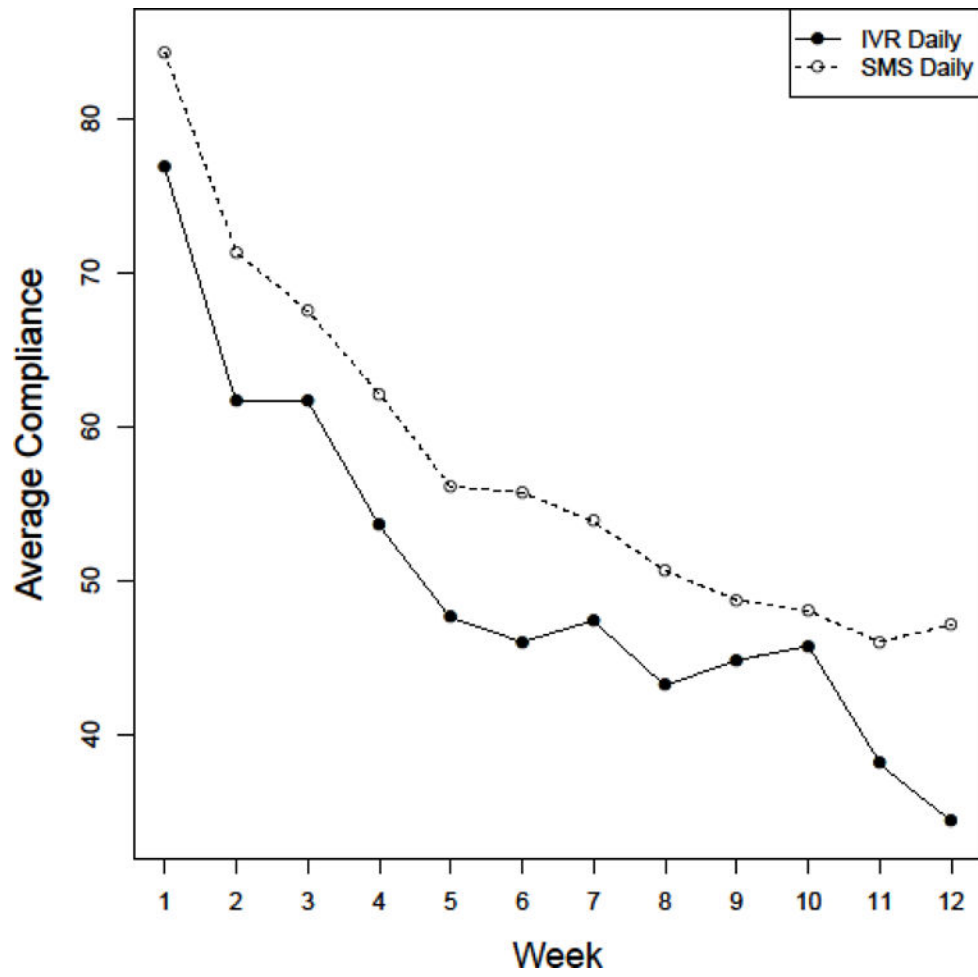The distributions of compliance rates by experimental groups.

Figure 2.
The week-by-week change in average compliance by the assessment method.

Table 1

Descriptive statistics of demographic and substance use variables at baseline by randomization groups (N=307).

	1.IVR Daily (N=81)	2.IVR Weekly (N=76)	3.SMS Daily (N=81)	4.SMS Weekly (N=69)	Statistical test
Age[a]	24.49 (2.35)	24.32 (2.69)	24.33 (2.32)	24.45 (2.33)	$F_{(3,303)}$=0.10, p=0.96
Male	45.68%	59.21%	51.85%	40.58%	$X^2_{(3)}$=5.72, p=0.13
Black	61.73%	52.63%	62.96%	62.32%	$X^2_{(3)}$=2.27, p=0.52
Public assistance	70.51%	55.41%	67.09%	69.57%	$X^2_{(3)}$=4.80, p=0.19
Nicotine use	70.37%	73.68%	58.02%	65.22%	$X^2_{(3)}$=4.98, p=0.17
Alcohol use	77.78%	60.53%	66.67%	69.57%	$X^2_{(3)}$=5.65, p=0.13
Marijuana use	72.84%	73.68%	67.90%	76.81%	$X^2_{(3)}$=1.56, p=0.67
Other illicit drug use	13.58%	9.21%	7.41%	17.39%	$X^2_{(3)}$=4.33, p=0.23
Prescription drug misuse	19.75%	21.05%	12.35%	20.29%	$X^2_{(3)}$=2.61, p=0.46

[a]Mean (standard deviation)

Table 2

Group comparison on compliance, use patterns, and user experiences (N=307).

	1.IVR Daily (N=81)	2.IVR Weekly (N=76)	3.SMS Daily (N=81)	4.SMS Weekly (N=69)	Group difference [b]
Compliance					
Rate [a]	50.21 (30.70)	71.46 (30.84)	57.12 (28.81)	64.95 (30.61)	1–2, 1–4, 2–3
Use patterns					
Duration [a]	2.20 (0.66)	7.44 (2.69)	14.66 (7.57)	49.81 (21.80)	All possible pairs
Single login	100%	97.37%	98.77%	75.36%	1–4, 2–4, 3–4
Completion of assessment	100%	100%	97.53%	91.30%	1–4
Evening responses	3.7%	17.11%	14.81%	11.59%	1–2
Response to reminders	64.20%	46.05%	38.27%	31.88%	1–3, 1–4,
User experiences					
Ease of use [a]	4.28 (0.97)	4.08 (1.24)	3.81 (1.23)	4.02 (1.17)	None
Privacy concern [a]	1.14 (0.43)	1.10 (0.36)	1.10 (0.40)	1.02 (0.08)	None
Burden of participation [a]	1.47 (0.68)	1.36 (0.53)	1.57 (0.72)	1.57 (0.61)	None
Difficulty in remembering [a]	1.14 (0.32)	1.21 (0.31)	1.17 (0.37)	1.15 (0.30)	None
Embarrassment about questions [a]	1.13 (0.38)	1.12 (0.47)	1.11 (0.37)	1.05 (0.18)	None

[a] Mean (standard deviation)

[b] Type I experimentwise error rate was controlled at the level of 0.05.