Published in final edited form as:

J Subst Abuse Treat. 2019 June ; 101: 12–17. doi:10.1016/j.jsat.2019.03.002.

What happens to agreement over time? A longitudinal study of self-reported substance use compared to saliva toxicological testing among subsidized housing residents

Alexis Rendon, Eun-Young Mun, Emily Spence-Almaguer, Scott T. Walters

University of North Texas Health Science Center, School of Public Health, Department of Health Behavior and Health Systems, Fort Worth, TX, United States of America

Abstract

The agreement between self-reported and toxicologically verified substance use provides important information about the validity of self-reported use. While some studies report aggregate agreement across follow-up points, only a few have examined the agreement at each time point separately. An overall rate of agreement across time may miss changes that occur as people progress through a research study. In this study, a sample of 644 adults (43.8% male, 32.6% White, 57.0% Black, 90.2% ages 36+) residing in subsidized housing was used to determine the agreement between self-reported use and saliva toxicological testing for marijuana, cocaine, PCP, amphetamine, and methamphetamine at three different time points. Agreement between saliva toxicological testing and self-report ranged between 84.2% and 94.3% for different substances over time. Higher rates of agreement were found for cocaine than had been reported by previous studies. Statistically significant differences in the odds ratios of concordance over time (baseline, 6-month, and 12-month follow-up) were found for marijuana and the combined category for PCP, amphetamine, and methamphetamine. Our findings suggest that oral fluid drug tests generally withstand community field assessments and result in relatively high levels of agreement for marijuana, cocaine, PCP, amphetamine, and methamphetamine use, when compared to self-report. Because of the ease of sample collection and low chance of adulteration, we conclude that saliva testing is a viable method for toxicological confirmation of substance use behavior in this setting.

Keywords

Substance use; oral fluid drug test; self-report; timeline follow-back; agreement; subsidized housing

1. Introduction

In substance use research, concordance refers to the agreement between self-reported drug use and toxicological confirmation. There are several methods of collecting self-reported substance use and toxicological substance use. Aimed at decreasing recall bias, the Timeline Follow-back (TLFB) method utilizes a visual calendar to enhance recall of substance use (Sobell & Sobell, 1992). There are multiple types of toxicology media: sweat, blood, urine, saliva, and hair, with different time windows of detection and varying degrees of invasiveness (Dolan, Rouen, & Kimber, 2004). Compared to urinalysis, considered to be the

toxicological testing standard in substance use research, saliva sample testing has low invasiveness and similar levels of sensitivity (Bennett, Davies, & Thomas, 2003; Quintela & Crouch, 2006). Thus, oral fluid testing has been established as an acceptable indicator of substance use (Cone, 2012; Neale & Robertson, 2003) and is as accurate as urinalysis for detecting the presence of opiates and methadone (Bennett, Davies, & Thomas, 2003).

Self-reported substance use and toxicological testing attempt to capture the same information within a specified time frame, although they do not always agree. Comparing self-report data with toxicological testing results in four distinct substance use groups: concordant users (self-report and toxicology both positive), concordant non-users (self-report and toxicology both negative), and two types of discordant users. Discordant users either underreport (self-report negative, toxicology positive) or overreport (self-report positive, toxicology negative). Studies often present concordant users and concordant non-users or “matches” between self-report and toxicological testing as a percent agreement (Hjorthøj, Hjorthøj, & Nordentoft, 2012).

Studies examining the agreement between self-report and toxicological confirmation show that some groups tend to underreport substance use (Digiusto, Seres, Bibby, & Batey, 1996; Goldfinger et al., 1996; Harrison, 1997; Napper, Fisher, Johnson, & Wood, 2010; Polcin, 2016; Schumacher et al., 1995; Sloan, Bodapati, & Tucker, 2004). There are a few instances of overreporting substance use, although this tends to be rarer (Haddock et al., 2009; McDowell et al., 2005). This means that, for most groups, estimates that rely on self-report probably underestimate the true amount of substance use, depending on the substance and population studied (Hjorthøj et al., 2012; Schumacher et al., 1995). Agreement between self-report and urinalysis ranges between 70%–87% for cocaine (Carroll et al., 2004; Elkashef et al., 2005; McDowell et al., 2005; Pettinati et al., 2008a; Pettinati et al., 2008b; Somoza et al., 2008; Stasiewicz et al., 2008) and 86%–98% for marijuana (Babor and The Marijuana Treatment Project Research Group, 2004; Godley, Godley, Dennis, Funk, & Passetti, 2002; Stasiewicz et al., 2008). A meta-analysis of substance use concordance conducted by Hjorthøj et al. (2012) included randomized controlled trials and cohort designs. Most studies with multiple follow-up points have combined toxicological confirmation data across follow-ups. For example, Carroll et al. (2004) combined weekly urine tests over 12 weeks to examine the overall agreement across the duration of the program. This methodology of collapsing follow-up concordance is frequently used in studies that report agreement between self-report and toxicological confirmatory tests (Carroll et al., 2006; Kiluk, Nich, Babuscio, & Carroll, 2010; McDowell et al., 2005; Morgenstern, Hogue, Dasaro, Kuerbis, & Dauber, 2008; Pettinati et al., 2008a; Pettinati et al., 2008b). While this provides valuable information about the average agreement post-intervention, combining data from all assessment points makes it challenging to examine how well the two modes converge over time. For instance, it is unclear whether concordance goes up or down, or tends to be stable over time.

As mentioned, while it is common to combine all follow-up time points for an overall agreement rate (Carroll et al., 2004; Carroll et al., 2006; Kiluk et al., 2010; McDowell et al., 2005; Morgenstern et al., 2008; Pettinati et al., 2008a; Pettinati et al., 2008b), this does not allow for the study of agreement between self-report and biological toxicological

confirmation over time. Of the few studies that have examined concordance over time, findings are mixed. Some studies find that concordance increases over time, while other studies find that concordance decreases over time (Babor and The Marijuana Treatment Project Research Group, 2004; Clark, Zyambo, Li, & Cropsey, 2016; Dillon, Turner, Robbins, & Szapocznik, 2005; Rohsenow, Martin, Eaton, & Monti, 2007). For instance, Babor and The Marijuana Treatment Project Research Group (2004) found that agreement was highest at baseline (94%) with 91% agreement at 4 months and 92% at 9 months. This study compared self-report to urine toxicological results for 450 adults seeking treatment for marijuana use. Similarly, when comparing the agreement between self-report and urine toxicological confirmation, Rohsenow et al. (2007) found that the proportion of false negatives (discordant: false negative self-report) increased slightly from 3 months (7%) to 6 months (9%).

The present study examined the agreement between self-reported substance use and saliva toxicological testing administered in field settings to determine if the rate of agreement changed over 12 months. Our sample was gathered from a population of adults with mental health problems living in subsidized housing. Unstable housing and substance use are reciprocal public health problems. While substance use disorders pose a significant barrier to maintaining stable housing (Willenbring, Whelan, Dahlquist, & O'Neal, 1990), the lack of stable, affordable housing is a barrier to substance use treatment and abstinence (Zerger, 2002). This analysis compared self-reported use with a saliva toxicological test for marijuana, cocaine, PCP, amphetamine, and methamphetamine at three time points spaced 6 months apart to determine if the agreement between the two modes changed over time. Compared to baseline, we hypothesized that agreement would increase at the 6-month and 12-month follow-ups due to study participants becoming more comfortable with reporting previously denied substance use.

2. Methods

2.1. Participants

This study uses data ($N=644$ at baseline, 43.8% male, 32.6% White, 57.0% Black, 90.2% age 36+) from m.chat, a technology-assisted health coaching program. The program enrolled low-income, uninsured or Medicaid-eligible adults participating in subsidized housing programs (Rendon, Livingston, Suzuki, Hill, & Walters, 2017; Walters, Spence-Almaguer, Hill, & Abraham, 2015). The m.chat program used monthly coaching meetings to help clients set and achieve goals related to diet, exercise, recreation and leisure, medication management, and substance use. Rolling recruitment and data collection began in November 2014 and continued through June 2017. Participants were recruited from Tarrant County, Texas from housing lists provided by housing agencies via case manager referrals, and word-of-mouth referrals.

Study participants were at least 18 years old and either Medicaid enrolled or low-income and uninsured at the time of screening. In addition, participants self-reported at least one of the following mental health problems: having been prescribed medication for psychological or emotional problems, receiving a pension for a psychiatric disability, experiencing hallucinations in the past year, or scoring >9 on the 9-item Patient Health Questionnaire

(PHQ-9) depression screener (Kroenke, Spitzer, & Williams, 2001). Participants were told at the start of the interview that they would be asked about their drug use and would submit a saliva drug test. All participants signed consent forms and were given assurances of confidentiality. Exclusion criteria included (1) not receiving subsidized housing, (2) any physical or sensory impairment that would substantially limit program participation or prevent accurate assessment of their health status, (3) non-English speaking, and (4) limited autonomy or decision-making capabilities (e.g., substantially neurologically or cognitively impaired). The project was approved by a local Institutional Review Board.

2.2. Substance use assessments and procedures

Semi-structured interviews were conducted by research assistants at baseline, 6 months and 12 months. Assessments were conducted in the field at participants' homes, other public places with semi-private interview space (e.g., libraries, recreation centers), or the project office. At each assessment, participants were asked about their substance use in the past 90 days via the TLFB assessment (Sobell & Sobell, 1992) and were given an oral fluid test collected via the Quantisal® oral fluid collection device (U.S. Patent No. 5479937, 1994). Samples were stored in a refrigerator before they were mailed for external laboratory testing via enzyme immunoassay (EIA). Positive EIA samples were further confirmed with gas chromatography/mass spectrometry at the external lab. The results indicated either a positive or negative detection of opiates, oxycodone, barbiturates, marijuana, cocaine, phencyclidine (PCP), amphetamine, and methamphetamine.

The present study focused on the detection of marijuana, cocaine, PCP, amphetamine, and methamphetamine. We excluded opiates and barbiturates from analysis because most participants reported having a prescription for opiates or painkillers and we were unable to differentiate prescribed opiate use from misuse. No participants tested positive for barbiturates or self-reported any barbiturate use. Due to the small number of PCP users, this substance was grouped together with amphetamines and methamphetamines for a larger stimulant category. The method of toxicological testing provided a maximum capture window of 72 h (3 days) for cocaine, PCP, amphetamine, and methamphetamine and up to 120 h (5 days) for marijuana, depending on severity of use (Redwood Toxicology Laboratory, 2015). Although self-reported substance use was collected for the past 90 days, only self-reported data from the corresponding drug-detection window was used to examine agreement. For example, only the previous 3 days (72 h) of self-reported use were used to compare to the toxicological test results for cocaine, PCP, amphetamine, and methamphetamine. For these analyses, the widest time interval was used to reduce false negatives. The drug testing cutoff concentrations were as follows: marijuana 1 ng/mL, cocaine 4 ng/mL, PCP 5 ng/mL, amphetamine 15 ng/mL, and methamphetamine 15 ng/mL. See Table 1 for descriptive statistics of the sample. Positive use in Table 1 indicates use based on either a toxicological test or self-reported use within the past 90 days.

3. Results

This sample was primarily composed of adults aged 36 and older (90.2%), a majority of whom (76.1%) reported an income of less than \$1000 per month (Table 1). At baseline, the

median age was 52.2 years and median monthly income was \$755. Any illicit substance use remained relatively stable between baseline (41.0%), 6 months (38.8%), and 12 months (42.1%). Men (55.0%) were more likely to use illicit substances compared to women (40.8%). Illicit substance use was associated with race at baseline, with 28.8% of Blacks, 13.5% of Whites, 3.1% of other races, and 1.6% of Hispanics having used illicit substances. Rates of substance use also varied across different types of substances. For instance, marijuana use decreased from baseline (30.6%) to 6 months (23.1%) and then increased at 12 months (25.8%) (Table 1).

Agreement between self-reported substance use and saliva toxicological testing was examined using data from the baseline, 6-month, and 12-month follow-ups (Table 2). More cases of substance use were detected positive via toxicological test than self-report, with a single exception of marijuana use at baseline. Raw agreement (i.e., estimated as the ratio of the agreement cell counts divided by total N ; see von Eye & Mun, 2005) estimates were excellent across all substance use types and over time. Raw agreement improved for marijuana use (0.85, 0.91, and 0.93, respectively across time) and was consistent for cocaine use (0.85, 0.91, and 0.89, respectively over time) and for the amphetamine, methamphetamine, and PCP group (0.94, 0.84, and 0.94, respectively over time). We further examined interrater agreement with Cohen's kappa (see Table 2). Overall, Cohen's kappa estimates across three types were 0.50 ($SE=0.03$, $t=19.26$, $p<0.01$), 0.32 ($SE=0.04$, $t=14.81$, $p<0.01$), and 0.24 ($SE=0.04$, $t=11.73$, $p<0.01$), respectively. This suggests that there was 50%, 32%, and 24% greater agreement than what was expected based on chance (i.e., accounting for main effects or marginal differences), and that moderate (Cohen's kappa between 0.4 and 0.6) to fair (between 0.2 and 0.4) agreement exists based on the interpretation proposed by Landis and Koch (1977). Across time, the agreement between two measures increased for marijuana (Cohen's kappa=0.41, 0.50, and 0.67), decreased for cocaine (Cohen's kappa=0.40, 0.17, 0.17), and was somewhat inconsistent for the amphetamine, methamphetamine, and PCP group (Cohen's kappa=0.46, 0.00, and 0.35; see Table 2 for greater details).

We further examined odds ratios (see Table 3) because odds ratio provides a marginal free test of association, unlike other measures of association for 2 by 2 tables, such as Pearson's Chi-square or the standard normal z , and also because the natural logarithm of an odds ratio value has a nice interpretation (see von Eye & Mun, 2003). We found that the odds ratios of concordance between two modes of substance use assessment were significantly different over time for marijuana and the combined amphetamine, methamphetamine, and PCP group (Chi-square for 2 $dfs=18.94$ and 22.66 , respectively) but not for cocaine (Chi-square for 2 $dfs=5.12$, $p=0.08$). The tests of odds ratios and the tests of homogeneity of odds ratios across time were generally consistent with the findings reported in Table 2 and provided statistical inferences about changes in agreement.

4. Discussion

This study examined rates of agreement between self-report and a toxicological drug test at baseline, 6 months, and 12 months among a sample of adults with mental health problems who were residing in subsidized housing. This study was unique because the parent program

was not a substance use treatment study, nor was substance use required for program eligibility.

In this study, it was not possible to examine individual-level factors due to the small *N* for positive-positive and discordant cases at follow-up (Table 2). Thus, we focused on rates of agreement over time in the current study. These individual level factors were assessed in our previous study that focused only on baseline assessments. Previously, we found that substance use concordance at baseline decreased with age and varied by race (Rendon et al., 2017). In the current study, we found that agreement (raw agreement) was generally high across time: 85%–93% for marijuana, 85%–91% for cocaine, and 84%–94% for the combined PCP, amphetamine, and methamphetamine group. The agreement rate of marijuana fell within previously reported rates of between 86% and 98% (Babor and The Marijuana Treatment Project Research Group, 2004; Godley et al., 2002; Stasiewicz et al., 2008). The agreement for cocaine was higher than previous studies which ranged between 70% and 87% (Carroll et al., 2004; Elkashef et al., 2005; McDowell et al., 2005; Pettinati et al., 2008a; Pettinati et al., 2008b; Somoza et al., 2008; Stasiewicz et al., 2008). The rates of agreement tended to improve over time for marijuana. The rates of agreement for cocaine were consistently high. For the combined amphetamine, methamphetamine, and PCP group, agreement was also high at baseline and at the 12-month follow-up. Using odds ratios and Cohen's kappa, we further evaluated agreement between the two modes of substance use reports. This interrater agreement approach goes beyond the raw agreement approach in the sense that marginal differences or main effects are taken into account. We found that there was a moderate level of agreement for marijuana use, which increased over time. For cocaine, the rate of agreement was moderate to fair, which tended to decrease over time. For the amphetamine, methamphetamine, and PCP group, the agreement over time was inconsistent but this result should be interpreted with a caveat that there were very few cases in some of the critical cells.

With regard to discordant cases, there may be a number of reasons why self-reported substance use and toxicological test did not match. First, concentrations that were lower than the minimum detection points set by the laboratory would be reported as a negative test result, explaining some false negative cases. Second, there may be real-life problems with using this collection device in the field as used in this study. In a study to validate the collection devices we used, samples were collected from laboratory solutions of substance concentrations, not from actual substance users (Quintela & Crouch, 2006). These samples were stored in the dark at room temperature for 12 h prior to analysis. Because our samples were collected in the field, they were undoubtedly exposed to higher or lower temperatures than would be ideal during vehicle transport back to the office. In addition, samples were typically held for 1–2 weeks before mailing for lab testing in order to maximize “batch” shipping, and thus reduce overall shipping costs. No samples in this study were analyzed within 12 h (the period used by the lab for norming purposes). Sample degradation may have led to imprecise test results, leading to lower rates of agreement. Finally, reduced recall of substance use behavior may also affect the agreement rate. Participants may be unable to accurately remember their substance use or may unconsciously or consciously censor their behavior when asked about their substance use behaviors, and thus may have underreported their use. On the aggregate, given that the toxicological test procedure was the same over

time, the greater agreement or stable agreement over time may be interpreted as generally accurate reporting behavior on the participants' side. Our study did find cases of disagreement—both false positives and false negatives. However, our findings suggest that oral fluid withstands community field assessments and results in relatively high levels of agreement for marijuana, cocaine, PCP, amphetamine, and methamphetamine use. Ease of sample collection and low chance of adulteration lead us to conclude that saliva testing is a viable method for toxicological confirmation of substance use behavior in field settings such as ours.

4.1. Limitations

Our results may not generalize to other populations who are less motivated to make behavior changes or who are motivated to explicitly conceal their behavior (e.g., mandated or criminal justice clients). Further, due to low rates of PCP, amphetamine, and methamphetamine use, there were very few cases in many sub-tables in Table 2. For these low-rate behaviors, we would need to examine larger samples to improve estimation precision. Still, findings from this research may help those who want to make accurate estimations about the extent of substance use in low-income housing populations.

5. Conclusion

Agreement between saliva toxicological testing and self-report in a subsidized housing population ranged between 84.2% and 94.3% for different substances over time. We found significant differences in concordance across time for marijuana and the combined PCP, amphetamine, and methamphetamine group, but not for cocaine. Future studies should consider examining the agreement between self-report and saliva toxicological testing for other types of substances.

Acknowledgement

This study was supported by CMS DSRIP 138980111.2.6 and in part by R01 AA019511. This work was supported by a Medicaid 1115 Waiver to the State of Texas by Centers for Medicare and Medicaid Services (CMS). CMS had no role in the study design, collection, analysis or interpretation of the data, writing the manuscript, or the decision to submit the paper for publication.

References

- Babor TF, & The Marijuana Treatment Project Research Group (2004). Brief treatments for cannabis dependence: Findings from a randomized multisite trial. *Journal of Consulting and Clinical Psychology*, 72(3), 455–466. 10.1037/0022-006X.72.3.455. [PubMed: 15279529]
- Bennett GA, Davies E, & Thomas P (2003). Is oral fluid analysis as accurate as urinalysis in detecting drug use in a treatment setting? *Drug and Alcohol Dependence*, 72(3), 265–269. 10.1016/j.drugalcdep.2003.08.001. [PubMed: 14643943]
- Carroll KM, Ball SA, Nich C, Martino S, Frankforter TL, Farentinos C, & Woody GE (2006). Motivational interviewing to improve treatment engagement and outcome in individuals seeking treatment for substance abuse: A multisite effectiveness study. *Drug and Alcohol Dependence*, 81(3), 301–312. 10.1016/j.drugalcdep.2005.08.002. [PubMed: 16169159]
- Carroll KM, Fenton LR, Ball SA, Nich C, Frankforter TL, Shi J, & Rounsaville BJ (2004). Efficacy of disulfiram and cognitive behavior therapy in cocaine-dependent outpatients. *Archives of General Psychiatry*, 61(3), 264–272. 10.1001/archpsyc.61.3.264. [PubMed: 14993114]

- Clark CB, Zyambo CM, Li Y, & Cropsey KL (2016). The impact of non-concordant self-report of substance use in clinical trials research. *Addictive Behaviors*, 58, 74–79. 10.1016/j.addbeh.2016.02.023. [PubMed: 26921721]
- Cone EJ (2012). Oral fluid results compared to self-reports of recent cocaine and heroin use by methadone maintenance patients. *Forensic Science International*, 215(1–3), 88–91. 10.1016/j.forsciint.2011.01.046. [PubMed: 21377302]
- Digiusto E, Seres V, Bibby A, & Batey R (1996). Concordance between urinalysis results and self-reported drug use by applicants for methadone maintenance in Australia. *Addictive Behaviors*, 21(3), 319–329. [PubMed: 8883483]
- Dillon FR, Turner CW, Robbins MS, & Szapocznik J (2005). Concordance among biological, interview, and self-report measures of drug use among African American and Hispanic adolescents referred for drug abuse treatment. *Psychology of Addictive Behaviors*, 19(4), 404–413. 10.1037/0893-164X.19.4.404. [PubMed: 16366812]
- Dolan K, Rouen D, & Kimber J (2004). An overview of the use of urine, hair, sweat and saliva to detect drug use. *Drug and Alcohol Review*, 23(2), 213–217. 10.1080/09595230410001704208. [PubMed: 15370028]
- Elkashef A, Holmes TH, Bloch DA, Shoptaw S, Kampman K, Reid MS, ... Vocci F (2005). Retrospective analyses of pooled data from CREST I and CREST II trials for treatment of cocaine dependence. *Addiction*, 100(Suppl. 1), 91–101. 10.1111/j.1360-0443.2005.00986.x. [PubMed: 15730353]
- Godley MD, Godley SH, Dennis ML, Funk R, & Passeti LL (2002). Preliminary outcomes from the assertive continuing care experiment for adolescents discharged from residential treatment. *Journal of Substance Abuse Treatment*, 23(1), 21–32. 10.1016/S0740-5472(02)00230-1. [PubMed: 12127465]
- Goldfinger SM, Schutt RK, Seidman LJ, Turner WM, Penk WE, & Tolomiczenko GS (1996). Self-report and observer measures of substance abuse among homeless mentally ill persons in the cross-section and over time. *The Journal of Nervous and Mental Disease*, 184(11), 667–672. 10.1097/00005053-199611000-00003. [PubMed: 8955679]
- Goldstein AS, Gavojdea S, & Zogg DF (1994). U.S. Patent No. 5479937 Oral collection device and method for immunoassay.
- Haddock G, Cross Z, Beardmore R, Tarrier N, Lewis S, Moring J, & Barrowclough C (2009). Assessing illicit substance use in schizophrenia: The relationship between self-report and detection by hair analysis. *Schizophrenia Research*, 114(1–3), 180–181. 10.1016/j.schres.2009.07.020. [PubMed: 19713083]
- Harrison L (1997). The validity of self-reported drug use in survey research: An overview and critique of research methods. *NIDA Research Monograph*, 167, 17–36. [PubMed: 9243555]
- Hjorthøj CR, Hjorthøj AR, & 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(3), 225–233. 10.1016/j.addbeh.2011.11.025. [PubMed: 22143002]
- Kiluk BD, Nich C, Babuscio T, & Carroll KM (2010). Quality versus quantity: Acquisition of coping skills following computerized cognitive-behavioral therapy for substance use disorders. *Addiction*, 105(12), 2120–2127. 10.1111/j.1360-0443.2010.03076.x. [PubMed: 20854334]
- Kroenke K, Spitzer RL, & Williams JBW (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606–613. 10.1046/j.1525-1497.2001.016009606.x. [PubMed: 11556941]
- Landis JR, & Koch GG (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. 10.2307/2529310. [PubMed: 843571]
- McDowell D, Nunes EV, Seracini AM, Rothenberg J, Vosburg SK, Ma GJ, & Petkova E (2005). Desipramine treatment of cocaine-dependent patients with depression: A placebo-controlled trial. *Drug and Alcohol Dependence*, 80(2), 209–221. 10.1016/j.drugalcdep.2005.03.026. [PubMed: 15913920]
- Morgenstern J, Hogue A, Dasaro C, Kuerbis A, & Dauber S (2008). Characteristics of individuals screening positive for substance use in a welfare setting: Implications for welfare and substance-

use disorders treatment systems. *Journal of Studies on Alcohol and Drugs*, 69(4), 561–570. 10.15288/jsad.2008.69.561. [PubMed: 18612572]

Napper LE, Fisher DG, Johnson ME, & Wood MM (2010). The reliability and validity of drug users' self-reports of amphetamine use among primarily heroin and cocaine users. *Addictive Behaviors*, 35(4), 350–354. 10.1016/j.addbeh.2009.12.006. [PubMed: 20053503]

Neale J, & Robertson M (2003). Comparisons of self-report data and oral fluid testing in detecting drug use amongst new treatment clients. *Drug and Alcohol Dependence*, 71(1), 57–64. 10.1016/S0376-8716(03)00053-X. [PubMed: 12821206]

Pettinati HM, Kampman KM, Lynch KG, Suh JJ, Dackis CA, Oslin DW, & O'Brien CP (2008a). Gender differences with high-dose naltrexone in patients with co-occurring cocaine and alcohol dependence. *Journal of Substance Abuse Treatment*, 34(4), 378–390. 10.1016/j.jsat.2007.05.011. [PubMed: 17664051]

Pettinati HM, Kampman KM, Lynch KG, Xie H, Dackis C, Rabinowitz AR, & O'Brien CP (2008b). A double blind, placebo-controlled trial that combines disulfiram and naltrexone for treating co-occurring cocaine and alcohol dependence. *Addictive Behaviors*, 33(5), 651–667. 10.1016/j.addbeh.2007.11.011. [PubMed: 18079068]

Polcin DL (2016). Co-occurring substance abuse and mental health problems among homeless persons: Suggestions for research and practice. *Journal of Social Distress and the Homeless*, 25(1), 1–10. 10.1179/1573658X15Y.0000000004. [PubMed: 27092027]

Quintela O, & Crouch DJ (2006). Recovery of drugs of abuse from the immunoanalysis Quantisal oral fluid collection device. *Journal of Analytical Toxicology*, 30, 614–616. 10.1093/jat/30.8.614. [PubMed: 17132261]

Redwood Toxicology Laboratory (2015). Laboratory testing reference guide. https://www.redwoodtoxicology.com/docs/resources/drug_classification.pdf.

Rendon A, Livingston M, Suzuki S, Hill W, & Walters S (2017). What's the agreement between self-reported and biochemical verification of drug use? A look at permanent supportive housing residents. *Addictive Behaviors*, 70, 90–96. 10.1016/j.addbeh.2017.02.011. [PubMed: 28214742]

Rohsenow DJ, Martin RA, Eaton CA, & Monti PM (2007). Cocaine craving as a predictor of treatment attrition and outcomes after residential treatment for cocaine dependence. *Journal of Studies on Alcohol and Drugs*, 68(5), 641–648. 10.15288/jsad.2007.68.641. [PubMed: 17690796]

Schumacher JE, Milby JB, Raczynski JM, Caldwell E, Engle M, Carr J, & Michael M (1995). Validity of self-reported crack cocaine use among homeless persons in treatment. *Journal of Substance Abuse Treatment*, 12(5), 335–339. 10.1016/0740-5472(95)02009-5. [PubMed: 8583516]

Sloan JJ, Bodapati MR, & Tucker TA (2004). Respondent misreporting of drug use in self-reports: Social desirability and other correlates. *Journal of Drug Issues*, 34(2), 269–292. 10.1177/002204260403400202.

Sobell LC, & Sobell MB (1992). Timeline follow-back In Litten RZ, & Allen JP (Eds.). *Measuring alcohol consumption* Totowa, NJ, USA: Humana Press 10.1007/978-1-4612-0357-5_3.

Somoza E, Somoza P, Lewis D, Li SH, Winhusen T, Chiang N, ... Elkashef A (2008). The SRPHK1 outcome measure for cocaine-dependence trials combines self-report, urine benzoylecgonine levels, and the concordance between the two to determine a cocaine-use status for each study day. *Drug and Alcohol Dependence*, 93(1–2), 132–140. 10.1016/j.drugalcdep.2007.09.007. [PubMed: 18029115]

Stasiewicz PR, Vincent PC, Bradizza CM, Connors GJ, Maisto SA, & Mercer ND (2008). Factors affecting agreement between severely mentally ill alcohol abusers' and collaterals' reports of alcohol and other substance abuse. *Psychology of Addictive Behaviors*, 22(1), 78–87. 10.1037/0893-164X.22.1.78. [PubMed: 18298233]

von Eye A, & Mun EY (2003). Characteristics of measures for 2×2 tables. *Understanding Statistics*, 2(4), 243–266. 10.1207/S15328031US0204_02.

von Eye A, & Mun EY (2005). *Analyzing rater agreement: Manifest variable methods*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Walters ST, Spence-Almaguer E, Hill W, & Abraham S (2015). Integrating health coaching and technology with vulnerable clients. *Social Work Today*, 6–7 Retrieved from <https://www.socialworktoday.com/archive/090915p6.shtml>.

- Willenbring ML, Whelan JA, Dahlquist JS, & O'Neal ME (1990). Community treatment of the chronic public inebriate I: Implementation. *Alcoholism Treatment Quarterly*, 7(1), 79–97. 10.1300/J020V07N01_07.
- Zerger S (2002). Substance abuse treatment: What works for homeless people? A review of the literature. Retrieved from National Health Care for the Homeless Council <https://www.nhchc.org/wp-content/uploads/2012/02/SubstanceAbuseTreatmentLitReview.pdf>.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Participant Demographics

		Baseline <i>N</i> = 644		Follow-up 1 <i>N</i> = 516		Follow-up 2 <i>N</i> = 361	
Demographics		<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Age	Young Adult (18 – 35 years)	63	9.78%	42	8.14%	23	6.37%
	Middle Adult (36 – 55 years)	365	56.68%	280	54.26%	198	54.85%
	Older Adult (56+ years)	215	33.39%	194	37.60%	140	38.78%
Sex	Male	282	43.79%	224	43.41%	156	43.21%
	Female	362	56.21%	292	56.59%	205	56.79%
Race	Black/African American	367	56.99%	287	55.62%	193	53.46%
	White	210	32.61%	175	33.91%	139	38.50%
	Hispanic	37	5.75%	33	6.40%	17	4.71%
	Other/Multi-racial	28	4.35%	29	5.62%	11	3.05%
	Don't Know/Refused	2	0.31%	2	0.39%	1	0.28%
Monthly Income	\$0	32	4.97%	18	3.49%	3	0.83%
	\$1 - \$499	102	15.84%	61	11.82%	30	8.31%
	\$500 - \$999	356	55.28%	270	52.33%	257	71.19%
	\$1,000+	153	23.76%	163	31.59%	68	18.84%
	Refused	1	0.16%	4	0.78%	3	0.83%
Illicit Substances (Self-Report or Saliva Test)	Any	264	40.99%	200	38.76%	152	42.11%
	Marijuana	197	30.59%	119	23.06%	93	25.76%
	Cocaine	156	24.22%	75	14.53%	62	17.17%
	Amphetamine, Methamphetamine, PCP	91	14.13%	48	9.30%	36	9.97%

Table 2.

Agreement Between Timeline Follow-back (TLFB) and the Saliva Toxicological Test across Substance Groups at Baseline, 6 months, and 12 Months

Baseline				Follow-up 1				Follow-up 2			
Marijuana	Test+	Test-	Row Total	Test+	Test-	Row Total	Test+	Test-	Row Total		
	TLFB+	46 (44.2%)	58 (55.8%)	104	TLFB+	30 (57.7%)	22 (42.3%)	52	TLFB+	31 (96.9%)	32
	TLFB-	36 (6.7%)	500 (93.3%)	536	TLFB-	26 (5.7%)	430 (94.3%)	456	TLFB-	25 (7.7%)	325
	82	558	640	56	452	508	56	301	357		
	Cohen's kappa (95% CI)			0.41 (0.31, 0.51)			Cohen's kappa (95% CI)			0.67 (0.55, 0.78)	
Cocaine	Test+	Test-	Row Total	Test+	Test-	Row Total	Test+	Test-	Row Total		
	TLFB+	41 (87.2%)	6 (12.8%)	47	TLFB+	6 (46.2%)	7 (53.8%)	13	TLFB+	5 (55.6%)	9
	TLFB-	87 (14.8%)	501 (85.2%)	588	TLFB-	41 (8.3%)	454 (91.7%)	495	TLFB-	35 (10.1%)	348
	128	507	635	47	461	508	40	317	357		
	Cohen's kappa (95% CI)			0.40 (0.31, 0.50)			Cohen's kappa (95% CI)			0.17 (0.02, 0.32)	
PCP, Amphetamine, Methamphetamine											
	Test+	Test-	Row Total	Test+	Test-	Row Total	Test+	Test-	Row Total		
	TLFB+	17 (81.0%)	4 (19.0%)	21	TLFB+	1 (14.3%)	6 (85.7%)	7	TLFB+	7 (77.8%)	9
	TLFB-	32 (5.2%)	582 (94.8%)	614	TLFB-	73 (14.8%)	420 (85.2%)	493	TLFB-	21 (6.0%)	348
	49	586	635	74	426	500	28	329	357		
	Cohen's kappa (95% CI)			0.46 (0.31, 0.61)			Cohen's kappa (95% CI)			0.35 (0.16, 0.55)	

Table 3.

Odds Ratios of Concordance Between Timeline Follow-back (TLFB) and the Saliva Toxicological Test across Substance Groups at Baseline, 6 months, and 12 Months

Time Point		Substance Group (OR, 95% CI)					
		Marijuana		Cocaine		Amphetamine, Methamphetamine, PCP	
	Baseline	11.02	(6.59, 18.42)	39.35	(16.22, 95.48)	77.30	(24.58, 243.11)
	6-Month	22.55	(11.45, 44.42)	9.49	(3.05, 29.57)	0.96	(0.11, 8.08)
	12-Month	372.00	(48.72, 2840.18)	11.18	(2.87, 43.57)	54.50	(10.65, 278.79)
Breslow-Day Test for Homogeneity of the Odds Ratio $df = 2$		Chi-Square	p -value	Chi-Square	p -value	Chi-Square	p -value
		18.94	<.01	5.12	0.08	22.66	<.01