Title

Correlates of Validity of Self-Reported Methamphetamine Use among a Sample of Dependent Adults

Journal

Substance Use & Misuse, 53(10)

ISSN

1082-6084

Authors

Rowe, Christopher
Vittinghoff, Eric
Colfax, Grant
et al.

Publication Date

2018-08-24

DOI

10.1080/10826084.2018.1432649

Peer reviewed

HHS Public Access

Author manuscript

Subst Use Misuse. Author manuscript; available in PMC 2019 August 24.

Published in final edited form as:

Subst Use Misuse. 2018 August 24; 53(10): 1742–1755. doi:10.1080/10826084.2018.1432649.

Correlates of validity of self-reported methamphetamine use among a sample of dependent adults

Christopher Rowe, MPH[a], **Eric Vittinghoff, PhD**[b], **Grant Colfax, MD**[c], **Philip O. Coffin, MD**[a,b], and **Glenn-Milo Santos, PhD**[a,b]

[a]San Francisco Department of Public Health, 25 Van Ness Avenue, Suite 500, San Francisco, CA 94102, U.S.A.

[b]University of California, San Francisco, 500 Parnassus Avenue, San Francisco, CA 94143, U.S.A.

[c]County of Marin Department of Health and Human Services, 3240 Kerner Boulevard, San Rafael, CA 94901, U.S.A.

Abstract

BACKGROUND—Self-reported data is widely used in substance use research, yet few studies have assessed the validity of self-reported methamphetamine use compared to biological assays.

OBJECTIVES—We sought to assess the validity and correlates of validity of self-reported methamphetamine use compared to urine toxicology (UTOX).

METHODS—Using a sample of methamphetamine-dependent individuals enrolled in a randomized controlled pharmacotherapy trial in the United States (n=327 visits among 90 participants), we calculated sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and the kappa coefficient of self-reported methamphetamine use in the past three days compared to UTOX, as well as the NPV of self-reported methamphetamine use over an extended recall period of one month. We used multivariable logistic regression models to assess correlates of concordance between self-reported methamphetamine use and UTOX.

RESULTS—The sensitivity of self-reported methamphetamine use in the past three days was 86.7% (95%CI: 81.4%−91.4%), the specificity was 85.3% (77.7–91.3), the PPV was 91.5% (86.9–94.8), and the NPV was 78.0% (69.4–86.1), compared to UTOX (kappa=0.71). The NPV over the extended recall period was 70.6% (48.0–85.7). In multivariable analyses, validity of self-reported methamphetamine was higher for older participants but lower during follow-up compared to baseline and when polysubstance use or depressive symptoms were reported.

CORRESPONDING AUTHOR: Christopher Rowe, chris.rowe@sfdph.org, 25 Van Ness, Suite 500, San Francisco, CA 94102, U.S.A., Telephone: 415-437-6283, Fax: 415-431-4039.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

PUBLISHING STATEMENT

This manuscript has not been published elsewhere and has not been submitted simultaneously for publication elsewhere.

DISCLAIMER

The authors are solely responsible for the content of this article, which does not necessarily represent the official views of the San Francisco Department of Public Health.

CONCLUSIONS/IMPORTANCE—Our sample of methamphetamine-dependent adults reported recent methamphetamine use with high validity compared to UTOX. Validity increased with age but decreased when participants reported depressive symptoms or polysubstance use as well as later in the study timeline and during longer recall periods.

INTRODUCTION

Methamphetamine use is associated with considerable mortality and morbidity and is increasing globally (Darke, Kaye, McKetin, & Duflou, 2008; Lineberry & Bostwick, 2006; United Nations Office on Drugs and Crime, 2016). In 2015, nearly 15 million Americans were estimated to have used methamphetamine in their lifetime and the number of current users increased over 60% from 2010 to 2014 (Center for Behavioral Health Statistics and Quality, 2011, 2015, 2016). Also in 2015, methamphetamine was the most common drug involved in drug-related criminal offenses in 27 states (United States Sentencing Commission, 2015). Increasing global prevalence and negative sequelae highlight a continued need for research on the most valid measures for methamphetamine use.

Participant self-report is a widely used data collection method in substance use related research. Compared to biological measures such as blood, hair, or urine assays, self-reported data is cheaper, easier and faster to obtain, and presents a lower burden on study participants. As such, studies that have aimed to discern trends, identify correlates and risk factors, and evaluate interventions have all relied on self-reported substance use behaviors. Parallel efforts have sought to assess the validity of self-reported substance use compared to biological measures (Brown, Kranzler, & Del Boca, 1992; Harrison, 1997; Ledgerwood, Goldberger, Risk, Lewis, & Price, 2008; Miller et al., 2015; Napper, Fisher, Johnson, & Wood, 2010; Rendon, Livingston, Suzuki, Hill, & Walters, 2017; Secades-Villa & Fernandez-Hermida, 2003). Compared to other more commonly used substances like cannabis, cocaine, or opioids, relatively few studies have assessed the validity of self-reported methamphetamine use compared to biochemical assays (Hjorthoj, Hjorthoj, & Nordentoft, 2012). Moreover, the majority of studies that sought to validate self-reported methamphetamine or amphetamine use either relied on small samples or low prevalence of use (Chen, Fang, Shyu, & Lin, 2006; Gryczynski, Schwartz, Mitchell, O'Grady, & Ondersma, 2014; Haddock et al., 2009; Ledgerwood et al., 2008), were conducted among distinct populations (Kab et al., 2012; Rendon et al., 2017), or did not widely explore demographic or behavioral differences in the validity of self-reported data (Napper et al., 2010). As societies continue to grapple with methamphetamine use, it is essential that research utilizing self-reported substance use data considers both the overall validity of such data as well as potential differences in reporting accuracy across subgroups.

We assessed the validity of self-reported methamphetamine use compared to urine toxicology among a sample of methamphetamine-dependent individuals enrolled in a randomized controlled pharmacotherapy trial in San Francisco, California. Furthermore, we explored a wide range of individual correlates of validity, including demographics, clinical and psychosocial characteristics, and additional substance use behaviors.

METHODS

Study Sample

This analysis examines data from a randomized, double-blind, placebo-controlled trial testing the efficacy of aripiprazole in decreasing methamphetamine use among dependent adults. The study has been described in detail elsewhere (Coffin et al., 2013); the primary finding was that aripiprazole, compared to placebo, had no significant effect on methamphetamine use. Active recruitment of participants was conducted at clinics, community-based organizations, and on the street in select neighborhoods; passive recruitment involved online advertisements and posted flyers. Eligibility criteria included methamphetamine dependence (as assessed by the Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders IV-TR, or SCID); an interest in stopping or reducing methamphetamine use; age 18–60 years; urine positivity for methamphetamine use at screening; no acute medical or psychiatric illness; and no clinically significant abnormalities during baseline safety laboratory monitoring. Participants were excluded if they exhibited major depression or bipolar disorder (as assessed by SCID); had a history of psychiatric medication in 4 weeks prior to screening; or, for HIV-positive individuals, had a CD4 cell count below 200 cells/μl. Participants were randomized 1:1 to receive either aripiprazole or identical placebo and followed for a total of 12 weeks after enrollment. Participants provided full informed consent and all study procedures and materials were approved by the University of California San Francisco Committee on Human Research.

Measures

Weekly study visits were scheduled for all participants, during which urine was collected and screened for methamphetamine metabolites and 30-minute substance use counseling was provided. Urine was screened using VERDICT®-II screening assays (MedTox Diagnoistics, Inc., Burlington, NC) with all screening results confirmed by two staff members before being documented. Behavioral information was collected by audio computer-assisted self-interviews (ACASI) at baseline and the 4-, 8-, and 12-week visits (i.e. monthly). The ACASI collected information on methamphetamine and other substance use, sexual activity, and mental health in the past four weeks at baseline and since the last monthly ACASI at the 4-, 8-, and 12-week visits (referred to as a “recall period”). At baseline, the ACASI interview also collected basic sociodemographic and clinical information.

Basic sociodemographic and clinical data included date of birth (which we converted to age at enrollment), race, gender, HIV status, annual income, and education history. Sexual orientation was determined from the gender of the participant (male or female) and the gender of all sexual partners reported during the duration of the study (male, female, transmale, or transfemale) as well as a two questions from a pre-screening telephone survey that asked whether the participant had sex with men or women while feeling the effects of meth in the past three months. Sexual orientation was defined as either reporting only heterogender sexual partners or reporting any same-gender or non-cisgender (i.e., transmale or transfemale) sexual partners, which we refer to as non-heterogender partners.

Frequency of methamphetamine use as well the timing of the most recent episode of methamphetamine use in the past four weeks was collected at baseline and since the last monthly ACASI at the monthly follow-up visits. Because urine assays generally detect methamphetamine metabolites from one to three days after use, the timing of the most recent episode was converted to binary self-report of any methamphetamine use in the past three days.

Participants also reported use of alcohol, marijuana, poppers (i.e., amyl nitrate), crack cocaine, powder cocaine, heroin, MDMA (i.e., ecstasy), ketamine, and other hallucinogens during each recall period. To assess use of substances other than methamphetamine (i.e., polysubstance use) among participants during each recall period, we considered use of any of the listed substances as polysubstance use and use of crack cocaine, powder cocaine, heroin, MDMA, ketamine, or other hallucinogens as polysubstance use other than alcohol, marijuana, and poppers, consistent with prior studies (Patterson, Semple, Zians, & Strathdee, 2005; Rowe, Santos, McFarland, & Wilson, 2015).

Participants also completed the Center for Epidemiologic Studies Depression Scale (CES-D) to assess symptoms of depression and depressive disorder in the one week prior to each monthly ACASI (Radloff, 1977). CES-D scores range from 0 to 60, with higher scores representing greater symptoms of depression. We used the standard cutoff to dichotomize CES-D scores into 0–15 as not having depressive symptoms, and 16 or greater as having depressive symptom, which has been shown to be valid in multiple populations (Lewinsohn, Seeley, Roberts, & Allen, 1997; Shean & Baldwin, 2008; Whooley, Avins, Miranda, & Browner, 1997).

Analysis

Primary Validity Measures—Our primary analysis involved assessing the validity of self-reported methamphetamine use in the past three days compared to urine toxicology, with up to four observations per participant (at baseline and the 4-, 8-, and 12-week visits). We calculated the sensitivity, specificity, positive predictive value, and negative predictive value for the entire sample as well as for subgroups of participants and their recall periods by age, race, gender, sexual orientation, HIV status, income, and education; we also calculated these validity measures for subgroups of recall periods by whether or not any polysubstance use or polysubstance use other than alcohol, marijuana, and poppers was reported, whether or not the participant had depressive symptoms based on CES-D during each recall period, visit number (i.e., baseline visit, 4-week visit, 8-week visit, or 12-week visit), and all follow-up visits in aggregate (i.e., 4-, 8-, and 12-week visits). We also calculated kappa coefficient using the three-day recall for the entire sample and for all participant and recall period subgroups. Data across study visits were aggregated in order to make full use of all available study data and increase the sample size available for analyses. To account for multiple recall periods per participant, we calculated 95% confidence intervals for all validity measures that included data from multiple visits using bias-corrected percentile bootstrap with 1000 replications and resampling by participant; for measures that do not include data from multiple visits, we calculated confidence intervals in the same way but did not resample by participant.

Negative Predictive Value with Extended Recall—To incorporate the weekly urine toxicology results between the baseline, 4-, 8-, and 12-week visits, and examine the validity of self-reported methamphetamine use over an extended recall period (i.e., greater than three days), we calculated the negative predictive value of reporting no methamphetamine use in the time between monthly ACASI surveys for periods that had at least one urine screen. This analysis excludes all baseline visit data because there were no urine screens in the weeks prior to the baseline visit. Seven additional visits/recall periods were excluded because, due to staff error, ACASI surveys asked about behaviors since a date that was different than the date of the participant's last monthly ACASI. This extended recall NPV was calculated for the entire sample.

Correlates of Validity—We used two different approaches to assess correlates of validity of self-reported methamphetamine use compared to urine toxicology. For the first approach (the stratified approach), we used logistic regression models fit with generalized estimating equations (GEE) to account for multiple recall periods per participant. We created four main stratified model structures to correspond to sensitivity, specificity, PPV, and NPV: (1) in the sensitivity model, the dependent variable was positive self-report of methamphetamine use in the past three days and only reporting periods with positive urine toxicology were included; (2) in the specificity model, the dependent variable was negative self-report of methamphetamine use in the past three days and only reporting periods with negative urine toxicology were included; (3) in the PPV model, the dependent variable was positive urine toxicology and only reporting periods with positive self-report of methamphetamine use in the past three days were included; (4) in the NPV model, the dependent variable was negative urine toxicology and only reporting periods with negative self-report of methamphetamine use in the past three days were included. The independent variables were a given participant or recall period characteristic (e.g. race, reporting of polysubstance use during the recall period) and all were included as indicator variables. We conducted Wald tests to assess the overall significance of each characteristic in each model.

In our second approach for assessing correlates of validity (the concordance approach), we generated a composite binary variable for each recall period indicating concordance between methamphetamine self-report and urine toxicology results, whether positive or negative. We used separate GEE logistic regression models assessing odds of concordance with a single participant or recall period characteristic as the independent variable. Again, we conducted Wald tests to assess the overall significance of the coefficients.

Because separate regression models were fit for each participant and recall period characteristic, we fit final multivariable models for the stratified approach and the concordance approach. For all five of the final multivariable models (four stratified models and one concordance model), covariates included age, race, gender, and any additional characteristics with an overall Wald test p-value < 0.25 in the appropriate bivariate model. Because reports of polysubstance use other than alcohol, marijuana, and poppers represent a subset of reports of any polysubstance use, only one of these variables, the one with the lower p-value, was retained in cases where both had Wald test p-values < 0.25. If both visit number and overall visit type (i.e., baseline versus all follow-up visits) had overall Wald test p-values < 0.25 in the bivariate models, we conducted additional Wald tests to test for

differences between the coefficients for the 4-, 8-, and 12-week visits; if there were no significant differences between the coefficients for these follow-up visits ($p < 0.05$), only overall visit type was retained.

Age, race, gender, income, education, polysubstance use, visit number, and visit type were assessed as potential correlates of validity based on prior literature showing associations between these characteristics and validity of self-reported substance use in different settings (Fendrich, Mackesy-Amiti, & Johnson, 2008; Langendam, van Haastrecht, & van Ameijden, 1999; Rendon et al., 2017; Schuler, Lechner, Carter, & Malcolm, 2009; White et al., 2014). Sexual orientation, HIV status, and depression were included as exploratory candidates based on elevated rates of substance use among sexual minorities (Medley et al., 2016), HIV-positive individuals (Gurung et al., 2017), and those suffering from depression (Swendsen & Merikangas, 2000), which make the validity of self-reported substance use among these groups of particular interest.

Although the treatment being assessed in the parent pharmacologic trial showed no effect on methamphetamine use, we fit bivariate stratified and concordance models with the treatment arm (aripiprazole versus placebo) for each participant. As with all other characteristics assessed in the biviariate models, treatment arm was included in the appropriate multivariable models if the bivariate Wald test *p*-value was < 0.25.

RESULTS

Our study sample includes 327 visits with both valid urine toxicology results and ACASI survey data among 90 study participants. The participants were racially/ethnically diverse (50% white, 19% black or African American, 17% Hispanic/Latino, and 14% any other race); most (88%) were male and reported non-heterogender sexual partners (66%) (Table 1). Numbers and percentages of methamphetamine use by self-report in the past three days and urine-positivity on the day of reporting; and numbers and percentages of self-reported methamphetamine abstinence during the entire period between monthly ACASI surveys and universal urine-negativity for all intervening weekly urine screens in the same period are presented in the Supplemental Table in the Appendix.

All validity measures for the entire sample and each subgroup are presented in Table 2. The overall sensitivity of self-reported methamphetamine use in the past three days was 86.7% (95%CI: 81.4% - 91.4%), the specificity was 85.3% (77.7% - 91.3%), the PPV was 91.5% (86.9% - 94.8%), and the NPV was 78.0% (69.4% - 86.1%). The kappa coefficient for the overall sample was 0.71 (95% CI: 0.62 – 0.78). The NPV of self-reported methamphetamine use over the extended recall period between monthly ACASI surveys was 70.6% (48.0% - 85.7%) among the entire sample (data not shown).

In the multivariable stratified models (Table 3), certain older age groups were associated with greater sensitivity and PPV (Sensitivity Model: Age 30–39 vs. Age 20–29 OR=7.3, 95%CI=1.4–39.2, p=0.020; PPV Model: Age 40–49 vs. Age 20–29 OR=5.8, 95%CI=1.2–29.5, p=0.033), reporting polysubstance use other than alcohol, marijuana, and poppers during reporting periods was associated with lower PPV (OR=0.2, 95%CI=0.1–0.6,

p=0.004), and follow-up reporting periods were associated with lower sensitivity and higher specificity compared to baseline reporting periods (Sensitivity Model: OR=0.3, 95%CI=0.1–0.8, *p*=0.024; PPV Model: OR=10.1, 95%CI=2.1–48.7, *p*=0.004). In the concordance model, there was a lower odds of concordance between self-report and urine toxicology results during recall periods in which the participant's CES-D score was 16 or greater (OR=0.4, 95%CI=0.2–0.9, *p*=0.024).

Treatment arm was not significantly associated with validity outcomes in the multivariable stratified models assessing sensitivity (p=0.124 in bivariate model; p=0.200 in multivariable model) and NPV (p=0.110 in bivariate model; p=0.185 in multivariable model) (Data not shown).

DISCUSSION

In our sample of methamphetamine-dependent adults participating in a pharmacotherapy trial, we found that self-reported methamphetamine use in the past three days had a relatively high sensitivity compared to urine toxicology. We also found that there were no significant differences in reporting accuracy between most demographic subgroups; however, we did find significant associations between reporting accuracy and age, polysubstance use other than alcohol, marijuana, and poppers, type of study visit (follow-up vs. baseline), and presence of depressive symptoms, which may have important implications for studies that rely solely on self-reported substance use behaviors.

Prior studies that have examined the sensitivity of self-reported methamphetamine use compared to a biological assay have varied widely, ranging from 35% to 100% (Chen et al., 2006; Gryczynski et al., 2014; Haddock et al., 2009; Ledgerwood et al., 2008). This variability may be partially explained by low prevalence of methamphetamine use among the samples studied, a limitation that is bypassed in our study of methamphetamine-dependent individuals. An analysis of a large sample of drug users that included 223 amphetamine-urine-positive individuals found a sensitivity of 61% for self-reported amphetamine in the last two days compared to urine toxicology (Napper et al., 2010); however, restricting the self-report window to two days, which is shorter than the urine detection window of one to three days for amphetamine, may artificially deflate the calculated sensitivity. It should also be noted that there may be important differences in the validity of self-reported substance use between participants in observational settings (Akinci, Tarter, & Kirisci, 2001; Fendrich & Johnson, 2005; Zaldivar Basurto et al., 2009) and those in treatment settings (Clark, Zyambo, Li, & Cropsey, 2016; Dillon, Turner, Robbins, & Szapocznik, 2005; Schuler et al., 2009; Wilcox, Bogenschutz, Nakazawa, & Woody, 2013), with the latter more likely to report with higher sensitivity. Participants in treatment settings, such as the clinical trial from which our sample was derived, will have personally acknowledged their substance use and desire to obtain treatment upon enrollment and may perceive less stigma in this type of setting compared to participants in more natural, observational research settings. These distinctions may reduce the presence of social desirability bias in the reporting of substance use behaviors in treatment-related research settings relative to observational settings. As such, it is important to interpret our findings within the context of a clinical trial evaluating a substance use treatment.

The high sensitivity among our sample suggests that self-reported methamphetamine use can be highly valid in clinical trial settings with a limited recall window. Despite this high sensitivity, it is important to consider the limitations of short recall windows in substance use research (i.e. three or six days in the present study). Although some studies may be interested in substance use behaviors that occur only a short time prior to study visits, many seek to collect information on behaviors over longer periods of time before or between study visits. In our study, the lower negative predictive value when examining the extended recall period between monthly ACASI surveys compared to the shorter three-day window suggests that longer recall duration may adversely affect the validity of self-reported data. As a result, study designs that rely solely on self-reported data and have the ability to follow participants may benefit from minimizing recall windows for reporting of substance use behaviors, using such tools as ecological momentary assessment or other daily data collection techniques (Rendina, Ventuneac, Mustanski, Grov, & Parsons, 2016; Shiffman, 2009). Indeed, a previous study conducted by the same team as the present study found more prevalent substance use reported by daily text message compared to surveys with greater recall periods (Rowe et al., 2015).

In contrast to prior studies that have assessed correlates of validity of self-reported use of other substances, we found few sociodemographic differences in the validity of self-reported methamphetamine use. For example, several studies found that Black or African American participants were more likely to underreport use of marijuana, cocaine, and opiates, relative to white participants (Fendrich & Johnson, 2005; Fendrich, Johnson, Wislar, Hubbell, & Spiehler, 2004; Richardson, Fendrich, & Johnson, 2003; White et al., 2014). Our conflicting findings suggest that demographic differences in validity may be substance- or sample-specific and not generalizable to how broader demographic groups report all substance use behaviors. Moreover, it should be noted that these earlier studies included cross-sectional household surveys and a cohort study targeting urban MSM, which are likely to capture different individuals than our pharmacotherapy trial among methamphetamine users interested in reducing their use. The results of our stratified multivariable models do, however, suggest that certain older age groups of methamphetamine-dependent adults may report substance use with greater validity compared to younger age groups, though the findings were not consistent across models or age groups. Prior studies in different populations and assessing different substances have had mixed findings linking age to the validity of self-reported substance use (Ledgerwood et al., 2008; Rendon et al., 2017), further highlighting the importance of delineating findings by the substance and population being studied.

We also found that recall periods in which greater depressive symptoms were reported suffered from poorer validity compared to recall periods with absent or less severe depressive symptoms. This finding is consistent with a study that found depression severity to negatively modify the validity of self-reported medication adherence (Gonzalez et al., 2013) as well as a broader psychology literature that has linked depression to cognitive impairment and memory deficits (Rock, Roiser, Riedel, & Blackwell, 2014). The association between depression and the validity of self-reported substance use data is particularly important given documented links between depression and substance use itself (Swendsen & Merikangas, 2000). For example, if a particular intervention simultaneously targets or affects

both an individual's substance use and depressive symptoms, potential changes in validity of self-reported substance use data as a result of changes in depressive symptoms should be considered when evaluating the intervention.

Compared to baseline reporting periods, follow-up reporting periods had both lower sensitivity and higher specificity of self-reported methamphetamine use. This suggests that study participants under-reported methamphetamine use at later stages in the study. This is consistent with multiple studies that have assessed the validity of self-reported substance use compared to biological assays in the context of behavioral or pharmacologic intervention trials and found reductions in sensitivity over time (Clark et al., 2016; Schuler et al., 2009; Tassiopoulos et al., 2004, 2006). It is plausible that such under-reporting could result from concerns of social desirability among participants, which may be enhanced as participants develop relationships with study counselors and other study staff over the duration of a study. Social desirability concerns among research participants has previously been associated with under-reporting of substance use behaviors (Johnson & Fendrich, 2005; Welte & Russell, 1993); however, this has the potential to be particularly problematic for studies that track substance use behaviors over time, as relationships between research staff and participants evolve. Because our study was not set up to specifically examine the presence or impact of social desirability bias, further research is needed to explore the association between social desirability concerns and the validity of substance use reporting in the context of longitudinal trials. Alternatively, because active methamphetamine use was part of the study's inclusion criteria, participants may have been more open to disclosing their methamphetamine use at baseline in order to ensure eligibility.

Our study has several limitations. First, due to our use of convenience sampling to recruit participants into a pharmacotherapy trial, our findings may not be generalizable to the broader population of methamphetamine-dependent adults. Second, eligibility criteria included both self-reported methamphetamine use as well as methamphetamine-positive urine toxicology, which may introduce selection bias towards a more accurate or honest sample of participants compared to the general population of methamphetamine-dependent adults. Third, our study design included weekly urine screens, which precludes our ability to objectively assess methamphetamine use without gaps for the entirety of follow-up; more specifically, the biological detection window of one to three days for urine toxicology leaves an undetectable window of four to six days between weekly visits. As such, the primary focus of our analysis was on the most recent episode of methamphetamine use, with a secondary analysis in which we calculated the negative predictive value of methamphetamine reporting over an extended recall period. Researchers should take care to select the biological assay (e.g., urine, hair, saliva, fingernail, sweat) with detection windows that best meets the needs of their particular research aims. Fourth, because our participants were aware that their urine would be screened for methamphetamine metabolites, the validity of their self-reported data may have been enhanced; our findings should be interpreted with caution in the context of studies that may rely only on self-reported data not coupled with a biological assay.

Ultimately, our sample of methamphetamine-dependent adults participating in a pharmacotherapy trial reported recent methamphetamine use with high validity compared to

urine toxicology. Validity decreased during longer recall periods and later in the study as well as when participants reported either polysubstance use other than alcohol, marijuana, and poppers, or clinically significant depressive systems; validity was higher for older age groups compared to the youngest group. The validity of self-reported methamphetamine use has been understudied relative to other more commonly used substances, highlighting the importance of novel contributions to the literature. Our findings suggest that self-reported methamphetamine use data among dependent adults in treatment-related trial settings is likely reliable when examining limited recall windows of three to six days; however, researchers should exercise caution when relying solely on self-reported data over longer recall periods, in longitudinal trials, or in studies in which polysubstance use or depressive symptoms are hypothesized to vary greatly or change over time.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

FUNDING

The study was funded by National Institute on Drug Abuse grant #1 R01 DA023387–01.

REFERENCES

Akinci IH, Tarter RE, & Kirisci L (2001). Concordance between verbal report and urine screen of recent marijuana use in adolescents. Addict Behav, 26(4), 613–619. [PubMed: 11456081]

Brown J, Kranzler HR, & Del Boca FK (1992). Self-reports by alcohol and drug abuse inpatients: factors affecting reliability and validity. Br J Addict, 87(7), 1013–1024. [PubMed: 1322747]

Center for Behavioral Health Statistics and Quality. (2011). 2010 National Survey on Drug Use and Health: Detailed Tables. Rockville, MD.: Substance Abuse and Mental Health Service Administration.

Center for Behavioral Health Statistics and Quality. (2015). 2014 National Survey on Drug Use and Health: Detailed Tables. Rockville, MD.: Substance Abuse and Mental Health Service Administration.

Center for Behavioral Health Statistics and Quality (2016). 2015 National Survey on Drug Use and Health: Detailed Tables. Rockville, MD.: Substance Abuse and Mental Health Service Administration.

Chen WJ, Fang CC, Shyu RS, & Lin KC (2006). Underreporting of illicit drug use by patients at emergency departments as revealed by two-tiered urinalysis. Addict Behav, 31(12), 2304–2308. doi: 10.1016/j.addbeh.2006.02.015 [PubMed: 16564643]

Clark CB, Zyambo CM, Li Y, & Cropsey KL (2016). The impact of non-concordant self-report of substance use in clinical trials research. Addict Behav, 58, 74–79. doi: 10.1016/j.addbeh.2016.02.023 [PubMed: 26921721]

Coffin PO, Santos GM, Das M, Santos DM, Huffaker S, Matheson T, … Colfax GN (2013). Aripiprazole for the treatment of methamphetamine dependence: a randomized, double-blind, placebo-controlled trial. Addiction, 108(4), 751–761. doi: 10.1111/add.12073 [PubMed: 23186131]

Darke S, Kaye S, McKetin R, & Duflou J (2008). Major physical and psychological harms of methamphetamine use. Drug Alcohol Rev, 27(3), 253–262. doi: 10.1080/09595230801923702 [PubMed: 18368606]

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. Psychol Addict Behav, 19(4), 404–413. doi: 10.1037/0893-164X.19.4.404 [PubMed: 16366812]

Fendrich M, & Johnson TP (2005). Race/ethnicity differences in the validity of self-reported drug use: results from a household survey. J Urban Health, 82(2 Suppl 3), iii67–81. doi: 10.1093/jurban/jti065 [PubMed: 15933333]

Fendrich M, Johnson TP, Wislar JS, Hubbell A, & Spiehler V (2004). The utility of drug testing in epidemiological research: results from a general population survey. Addiction, 99(2), 197–208. [PubMed: 14756712]

Fendrich M, Mackesy-Amiti ME, & Johnson TP (2008). Validity of self-reported substance use in men who have sex with men: comparisons with a general population sample. Ann Epidemiol, 18(10), 752–759. doi: 10.1016/j.annepidem.2008.06.001 [PubMed: 18693041]

Gonzalez JS, Schneider HE, Wexler DJ, Psaros C, Delahanty LM, Cagliero E, & Safren SA (2013). Validity of medication adherence self-reports in adults with type 2 diabetes. Diabetes Care, 36(4), 831–837. doi: 10.2337/dc12-0410 [PubMed: 23204245]

Gryczynski J, Schwartz RP, Mitchell SG, O'Grady KE, & Ondersma SJ (2014). Hair drug testing results and self-reported drug use among primary care patients with moderate-risk illicit drug use. Drug Alcohol Depend, 141, 44–50. doi: 10.1016/j.drugalcdep.2014.05.001 [PubMed: 24932945]

Gurung S, Ventuneac A, Cain D, Mirzayi C, Ferraris C, Rendina HJ, … Parsons JT (2017). Alcohol and substance use diagnoses among HIV-positive patients receiving care in NYC clinic settings. Drug Alcohol Depend, 180, 62–67. doi: 10.1016/j.drugalcdep.2017.07.034 [PubMed: 28881318]

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. Schizophr Res, 114(1–3), 180–181. doi: 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 Res Monogr, 167, 17–36. [PubMed: 9243555]

Hjorthoj CR, Hjorthoj AR, & Nordentoft M (2012). Validity of Timeline Follow-Back for self-reported use of cannabis and other illicit substances--systematic review and meta-analysis. Addict Behav, 37(3), 225–233. doi: 10.1016/j.addbeh.2011.11.025 [PubMed: 22143002]

Johnson T, & Fendrich M (2005). Modeling sources of self-report bias in a survey of drug use epidemiology. Ann Epidemiol, 15(5), 381–389. doi: 10.1016/j.annepidem.2004.09.004 [PubMed: 15840552]

Kab V, Evans J, Sansothy N, Stein E, Claude-Couture M, Maher L, … Young Women's Study, C. (2012). Testing for amphetamine-type stimulant (ATS) use to ascertain validity of self-reported ATS use among young female sex workers in Cambodia. Addict Sci Clin Pract, 7, 11. doi: 10.1186/1940-0640-7-11 [PubMed: 23186171]

Langendam MW, van Haastrecht HJ, & van Ameijden EJ (1999). The validity of drug users' self-reports in a non-treatment setting: prevalence and predictors of incorrect reporting methadone treatment modalities. Int J Epidemiol, 28(3), 514–520. [PubMed: 10405858]

Ledgerwood DM, Goldberger BA, Risk NK, Lewis CE, & Price RK (2008). Comparison between self-report and hair analysis of illicit drug use in a community sample of middle-aged men. Addict Behav, 33(9), 1131–1139. doi: 10.1016/j.addbeh.2008.04.009 [PubMed: 18547737]

Lewinsohn PM, Seeley JR, Roberts RE, & Allen NB (1997). Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults. Psychol Aging, 12(2), 277–287. [PubMed: 9189988]

Lineberry TW, & Bostwick JM (2006). Methamphetamine abuse: a perfect storm of complications. Mayo Clin Proc, 81(1), 77–84. doi: 10.4065/81.1.77 [PubMed: 16438482]

Medley G, Lipari RN, Bose J, Cribb DS, Kroutil LA, & McHenry G (2016). Sexual Orientation and Estimates of Adult Substance Use and Mental Health: Results from the 2015 National Survey on Drug Use and Health NSDUH Data Review: Substance Abuse and Mental Health Services Administration.

Miller P, Curtis A, Jenkinson R, Droste N, Bowe SJ, & Pennay A (2015). Drug use in Australian nightlife settings: estimation of prevalence and validity of self-report. Addiction, 110(11), 1803–1810. doi: 10.1111/add.13060 [PubMed: 26189494]

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. Addict Behav, 35(4), 350–354. doi: 10.1016/j.addbeh.2009.12.006 [PubMed: 20053503]

Patterson TL, Semple SJ, Zians JK, & Strathdee SA (2005). Methamphetamine-using HIV-positive men who have sex with men: correlates of polydrug use. J Urban Health, 82(1 Suppl 1), i120–126. doi: 10.1093/jurban/jti031 [PubMed: 15738313]

Radloff LS (1977). The CES-D scale: A self-report depression scale for research in the general population. Applied psychological measurement, 1(3), 385–401.

Rendina HJ, Ventuneac A, Mustanski B, Grov C, & Parsons JT (2016). Prospective Measurement of Daily Health Behaviors: Modeling Temporal Patterns in Missing Data, Sexual Behavior, and Substance Use in an Online Daily Diary Study of Gay and Bisexual Men. AIDS Behav, 20(8), 1730–1743. doi: 10.1007/s10461-016-1359-0 [PubMed: 26992392]

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. Addict Behav, 70, 90–96. doi: 10.1016/j.addbeh.2017.02.011 [PubMed: 28214742]

Richardson J, Fendrich M, & Johnson TP (2003). Neighborhood effects on drug reporting. Addiction, 98(12), 1705–1711. [PubMed: 14651502]

Rock PL, Roiser JP, Riedel WJ, & Blackwell AD (2014). Cognitive impairment in depression: a systematic review and meta-analysis. Psychol Med, 44(10), 2029–2040. doi: 10.1017/S0033291713002535 [PubMed: 24168753]

Rowe C, Santos GM, McFarland W, & Wilson EC (2015). Prevalence and correlates of substance use among trans female youth ages 16–24 years in the San Francisco Bay Area. Drug Alcohol Depend, 147, 160–166. doi: 10.1016/j.drugalcdep.2014.11.023 [PubMed: 25548025]

Schuler MS, Lechner WV, Carter RE, & Malcolm R (2009). Temporal and gender trends in concordance of urine drug screens and self-reported use in cocaine treatment studies. J Addict Med, 3(4), 211–217. doi: 10.1097/ADM.0b013e3181a0f5dc [PubMed: 20209029]

Secades-Villa R, & Fernandez-Hermida JR (2003). The validity of self-reports in a follow-up study with drug addicts. Addict Behav, 28(6), 1175–1182. [PubMed: 12834660]

Shean G, & Baldwin G (2008). Sensitivity and specificity of depression questionnaires in a college-age sample. J Genet Psychol, 169(3), 281–288. doi: 10.3200/GNTP.169.3.281-292 [PubMed: 18788328]

Shiffman S (2009). Ecological momentary assessment (EMA) in studies of substance use. Psychol Assess, 21(4), 486–497. doi: 10.1037/a0017074 [PubMed: 19947783]

Swendsen JD, & Merikangas KR (2000). The comorbidity of depression and substance use disorders. Clin Psychol Rev, 20(2), 173–189. [PubMed: 10721496]

Tassiopoulos K, Bernstein J, Heeren T, Levenson S, Hingson R, & Bernstein E (2004). Hair testing and self-report of cocaine use by heroin users. Addiction, 99(5), 590–597. doi: 10.1111/j.1360-0443.2004.00685.x [PubMed: 15078233]

Tassiopoulos K, Bernstein J, Heeren T, Levenson S, Hingson R, & Bernstein E (2006). Predictors of disclosure of continued cocaine use. Addict Behav, 31(1), 80–89. doi: 10.1016/j.addbeh.2005.04.005 [PubMed: 15908136]

United Nations Office on Drugs and Crime. (2016). World Drug Report.

United States Sentencing Commission. (2015). 2015 Sourcebook of Federal Sentencing Statistics.

Welte JW, & Russell M (1993). Influence of socially desirable responding in a study of stress and substance abuse. Alcohol Clin Exp Res, 17(4), 758–761. [PubMed: 8214409]

White D, Rosenberg ES, Cooper HL, del Rio C, Sanchez TH, Salazar LF, & Sullivan PS (2014). Racial differences in the validity of self-reported drug use among men who have sex with men in Atlanta, GA. Drug Alcohol Depend, 138, 146–153. doi: 10.1016/j.drugalcdep.2014.02.025 [PubMed: 24629628]

Whooley MA, Avins AL, Miranda J, & Browner WS (1997). Case-finding instruments for depression. Two questions are as good as many. J Gen Intern Med, 12(7), 439–445. [PubMed: 9229283]

Wilcox CE, Bogenschutz MP, Nakazawa M, & Woody G (2013). Concordance between self-report and urine drug screen data in adolescent opioid dependent clinical trial participants. Addict Behav, 38(10), 2568–2574. doi: 10.1016/j.addbeh.2013.05.015 [PubMed: 23811060]

Zaldivar Basurto F, Garcia Montes JM, Flores Cubos P, Sanchez Santed F, Lopez Rios F, & Molina Moreno A (2009). Validity of the self-report on drug use by university students: correspondence between self-reported use and use detected in urine. Psicothema, 21(2), 213–219. [PubMed: 19403073]

Table 1:

Participant and Visit characteristics by urine positivity and self-reported methamphetamine use among the entire study sample (n=327 visits among 90 participants)

	n	(%)
Entire Sample	90	
Participant Characteristics		
Age		
20–29	18	(20.0)
30–39	23	(25.6)
40–49	36	(40.0)
50+	13	(14.4)
Race		
White	45	(50.0)
Black or African American	17	(18.9)
Hispanic/Latino	15	(16.7)
Other	13	(14.4)
Gender		
Female	11	(12.2)
Male	79	(87.8)
Sexual orientation		
Only heterogender partners	29	(32.2)
Non-heterogender partners	59	(65.6)
HIV status		
Negative	63	(70.0)
Positive	27	(30.0)
Income		
No income	18	(20.0)
Under $30,000	57	(63.3)
$30,000 and above	12	(13.3)
Education		
High school or less	40	(44.4)
Some college	36	(40.0)
College or more	14	(15.6)
Substance Use in Four Weeks Prior to Baseline		
Methamphetamine use frequency		
Less than once per week	8	(8.9)
1–2 days per week	20	(22.2)
3–6 days per week	43	(47.8)
Everyday	19	(21.1)
Polysubstance use*	81	(90.0)

	n	(%)
Polysubstance use other than alcohol, marijuana, and poppers†	39	(43.3)
Use of specific substances		
Alcohol	57	(63.3)
Marijuana	59	(65.6)
Poppers	19	(21.1)
Crack cocaine	19	(21.1)
Powdered cocaine	11	(12.2)
Heroin	11	(12.2)
MDMA	13	(14.4)
Ketamine	4	(4.4)
Other hallucinogens	4	(4.4)
Monthly Visit Information		
Visits included in analysis		
Baseline	87	(96.7)
Week 4	80	(88.9)
Week 8	79	(87.8)
Week 12	81	(90.0)

*Polysubstance use is defined as any use of alcohol, marijuana, poppers, crack cocaine, powdered cocaine, heroin, MDMA, ketamine, or other hallucinogens

†Polysubstance use other than alcohol, marijuana, or poppers is defined as any use of crack cocaine, powdered cocaine, heroin, MDMA, ketamine, or other hallucinogens.

Table 2:

Validity measures of self-reported methamphetamine use compared to urine screen among study sample (n=327 visits among 90 participants)

	Three Day Recall[a]									
	Sensitivity		Specificity		PPV		NPV		Kappa	
	%	(95% CI)[c]	%	(95% CI)[c]	%	(95% CI)[c]	%	(95% CI)[c]	κ	(95% CI)[c]
Entire Sample	86.7%	(81.4% – 91.4%)	85.3%	(77.7% – 91.3%)	91.5%	(86.9% – 94.8%)	78.0%	(69.4% – 86.1%)	0.71	(0.62 – 0.78)
Participant Characteristics										
Age										
20–29	73.3%	(51.5% – 88.9%)	85.7%	(76.3% – 94.9%)	81.5%	(61.5% – 96.0%)	78.9%	(60.9% – 92.0%)	0.59	(0.37 – 0.78)
30–39	94.6%	(85.1% – 100%)	76.9%	(50.0% – 92.7%)	89.8%	(79.7% – 97.0%)	87.0%	(58.3% – 100%)	0.74	(0.57 – 0.89)
40–49	87.1%	(79.0% – 93.7%)	89.7%	(75.8% – 97.8%)	95.3%	(91.1% – 98.9%)	74.5%	(57.9% – 88.2%)	0.73	(0.59 – 0.84)
50+	84.4%	(68.0% – 96.6%)	87.5%	(73.3% – 100%)	93.1%	(80.0% – 100%)	73.7%	(50.0% – 94.4%)	0.69	(0.44 – 0.88)
Race										
White	87.6%	(79.7% – 93.4%)	71.7%	(55.2% – 85.4%)	88.4%	(81.3% – 93.6%)	70.2%	(52.5% – 84%)	0.59	(0.44 – 0.73)
Black or African American	82.1%	(61.8% – 95.5%)	92.3%	(82.8% – 100%)	94.1%	(78.8% – 100%)	77.4%	(59.3% – 94.9%)	0.72	(0.55 – 0.87)
Hispanic/Latino	87.9%	(73.1% – 97.3%)	95.7%	(86.4% – 100%)	96.7%	(87.0% – 100%)	84.6%	(68.8% – 96.4%)	0.82	(0.64 – 0.96)
Other	88.5%	(76.0% – 100%)	95.2%	(75.0% – 100%)	95.2%	(86.7% – 100%)	87.0%	(62.5% – 97.0%)	0.83	(0.66 – 0.96)
Gender										
Female	73.9%	(50.0% – 89.5%)	83.3%	(75.0% – 100%)	85.0%	(58.3% – 100%)	71.4%	(50.0% – 89.5%)	0.56	(0.33 – 0.81)
Male	88.3%	(82.7% – 93.3%)	85.7%	(77.3% – 92.4%)	92.2%	(87.9% – 95.6%)	79.2%	(68.5% – 88.2%)	0.73	(0.64 – 0.81)
Sexual orientation										
Only hetero-gender partners	81.0%	(69.2% – 90.7%)	84.2%	(73.5% – 92.9%)	88.7%	(76.5% – 95.7%)	74.4%	(57.1% – 87.8%)	0.64	(0.48 – 0.79)
Same-gender partners	89.4%	(83.5% – 94.4%)	84.7%	(73.9% – 92.3%)	92.5%	(87.7% – 96.5%)	79.2%	(68.3% – 88.9%)	0.73	(0.64 – 0.82)
HIV status										
Negative	86.7%	(80.5% – 92.9%)	80.0%	(70.0% – 87.7%)	87.9%	(81.7% – 93.1%)	78.2%	(66.2% – 87.5%)	0.66	(0.55 – 0.76)
Positive	86.8%	(73.9% – 94.5%)	100.0%	(N/A[d])	100.0%	(N/A[d])	77.5%	(59.4% – 90.9%)	0.80	(0.67 – 0.91)
Income										
No income	87.2%	(73.7% – 95.9%)	76.2%	(56.3% – 89.5%)	87.2%	(73.1% – 96.0%)	76.2%	(53.8% – 91.7%)	0.63	(0.38 – 0.82)
Under $30,000	85.9%	(78.5% – 91.5%)	91.0%	(81.7% – 97.0%)	94.3%	(88.7% – 98.3%)	78.9%	(67.4% – 87.5%)	0.75	(0.66 – 0.83)

	Three Day Recall[a]									
	Sensitivity		Specificity		PPV		NPV		Kappa	
	%	(95% CI)[c]	%	(95% CI)[c]	%	(95% CI)[c]	%	(95% CI)[c]	κ	(95% CI)[c]
$30,000 and above	88.2%	(73.1% – 100%)	60.0%	(20.0% – 80.0%)	88.2%	(77.4% – 97.0%)	60.0%	(16.7% – 100%)	0.48	(0.15 – 0.76)
Education										
High school or less	86.6%	(74.7% – 93.3%)	91.2%	(83.0% – 97.0%	93.4%	(86.9% – 97.8%)	82.5%	(71.0% – 91.9%)	0.77	(0.64 – 0.86)
Some college	84.6%	(76.7% – 92.5%)	80.5%	(63.0% – 91.3%)	90.6%	(81.3% – 96.4%)	70.2%	(51.2% – 85.7%)	0.63	(0.47 – 0.77)
College or more	92.1%	(82.4% – 100%)	77.8%	(50.0% – 94.1%)	89.7%	(81.1% – 97.1%)	82.4%	(60.0% – 100%)	0.71	(0.48 – 0.88)
Visit Characteristics										
Polysubstance use										
Not Reported	85.3%	(71.1% – 96.6%	95.2%	(78.9% – 100.0%)	96.7%	(85.7% – 100.0%)	80.0%	(57.9% – 93.3%)	0.78	(0.59 – 0.93)
Reported	87.0%	(80.9% – 92.3%)	83.2%	(75.3% – 90.7%)	90.6%	(86.1% – 94.9%)	77.5%	(65.6% – 86.0%)	0.69	(0.59 – 0.77)
Polysubstance use other than alcohol, marijuana and poppers										
Not Reported	88.7%	(83.0% – 91.8%)	91.2%	(83.3% – 98.0%)	96.2%	(92.6% – 99.3%)	76.5%	(63.0% – 86.6%)	0.76	(0.64 – 0.84)
Reported	82.9%	(72.2% – 91.0%)	79.7%	(67.4% – 89.4%)	82.9%	(72.7% – 91.5%)	79.7%	(66.7% – 89.4%)	0.63	(0.47 – 0.75)
CES-D score										
<16	89.7%	(82.6% – 94.8%)	89.7%	(86.6% – 97.9%)	95.5%	(90.5% – 98.9%)	84.0%	(73.9% – 91.9%)	0.81	(0.71 – 0.89)
≥16	82.8%	(74.3% – 90.2%)	75.0%	(59.5% – 87.0%)	86.5%	(77.2% – 93.3%)	69.2%	(53.3% – 82.6%)	0.57	(0.40 – 0.71)
Visit type										
Baseline	93.8%	(86.5% – 98.5%)	63.6%	(40.0% – 82.6%)	88.4%	(79.4% – 94.4%)	77.8%	(50.0% – 94.4%)	0.61	(0.40 – 0.80)
Follow-Up (4-, 8-, and 12-week visits)	83.6%	(76.3% – 90.2%)	90.4%	(82.8% – 95.4%)	93.1%	(88.1% – 96.8%)	78.0%	(67.6% – 86.2%)	0.72	(0.62 – 0.80)
4-week visit	83.3%	(71.4% – 92.6%)	84.6%	(69.0% – 96.4%)	91.8%	(83.7% – 98.1%)	71.0%	(53.3% – 86.7%)	0.65	(0.45 – 0.81)
8-week visit	87.8%	(78.0% – 95.8%)	96.7%	(87.0% – 100%)	97.7%	(90.9% – 100%)	82.9%	(69.4% – 93.9%)	0.82	(0.68 – 0.92)
12-week visit	79.1%	(65.9% – 90.0%)	89.5%	(78.9% – 97.6%)	89.5%	(77.5% – 97.4%)	79.1%	(65.9% – 89.4%)	0.68	(0.52 – 0.83)

[a] Three day recall measures the validity of self-reporting methamphetamine use in the three days prior to a urine screen compared to the results of the urine screen.

[c] 95% Confidence intervals calculated using bias-corrected percentile bootstrap with 1000 replications and resampling by participant to account for multiple visits per participant, except in the case of the baseline, 4-, 8-, and 12-week visit types, which were calculated without resampling by participant.

[d] 95% Confidence intervals not calculable when validity measure equals 100%.

Table 3:

Multivariable GEE logistic regression models assessing predictors of validity of self-reported methamphetamine use compared to urinalysis (n=327 visits among 90 participants)

	Stratified Validity Measure Models[a]																			
	Sensitivity (n=209 visits among 73 participants)				Specificity (n=109 visits among 48 participants)				Positive Predictive Value (n=199 visits among 77 participants)				Negative Predictive Value (n=127 visits among 53 participants)				As Predictor of Concordance[b]			
	OR	(95% CI)	*p*-value	Wald *p*-value	OR	(95% CI)	*p*-value	Wald *p*-value	OR	(95% CI)	*p*-value	Wald *p*-value	OR	(95% CI)	*p*-value	Wald *p*-value	OR	(95% CI)	*p*-value	Wald *p*-value
Participant Characteristics																				
Age																				
20–29		Reference				Reference				Reference				Reference				Reference		
30–39	7.3	(1.4 – 39.2)	0.020		0.9	(0.1 – 6.9)	0.941		2.0	(0.4 – 9.9)	0.373		1.4	(0.3–7.3)	0.725		1.7	(0.6 – 4.6)	0.292	
40–49	2.9	(0.9 – 9.9)	0.087	0.106	0.8	(0.1 – 5.6)	0.836	0.940	5.8	(1.2 – 29.5)	0.033	0.110	0.5	(0.2–1.8)	0.318	0.564	1.6	(0.7 – 3.9)	0.258	0.613
50+	2.9	(0.6 – 14.1)	0.178		0.5	(0.0 – 5.9)	0.560		6.7	(0.9 – 47.4)	0.057		0.6	(0.1–2.7)	0.477		1.8	(0.6 – 5.2)	0.291	
Race																				
White		Reference				Reference				Reference				Reference				Reference		
Black or African American	0.6	(0.2 – 1.9)	0.366		1.8	(0.2 – 19.1)	0.616		2.5	(0.4 – 17.2)	0.342		1.5	(0.4–5.1)	0.550		1.3	(0.5 – 3.1)	0.556	
Hispanic/Latino	1.0	(0.2 – 4.3)	0.991	0.772	3.6	(0.3 – 42.6)	0.307	0.348	9.4	(1.0 – 87.6)	0.050	0.178	2.1	(0.5–9.3)	0.316	0.623	2.0	(0.7 – 5.9)	0.187	0.334
Other	1.2	(0.3 – 5.7)	0.799		7.8	(0.7 – 87.6)	0.095		2.8	(0.3 – 24.6)	0.345		2.3	(0.5–10.6)	0.275		2.3	(0.8 – 7.3)	0.141	
Gender																				
Female		Reference				Reference				Reference				Reference				Reference		
Male	2.0	(0.5 – 7.9)	0.313	0.313	0.3	(0.0 – 4.3)	0.403	0.403	4.1	(0.7 – 25.8)	0.130	0.130	1.1	(0.3–4.1)	0.908	0.908	1.2	(0.5 – 3.4)	0.676	0.676
Sexual orientation																				
Only hetero-gender partners		Reference				-		-		-		-		-		-		Reference		
Non-heterogender partners	2.4	(0.7 – 7.6)	0.144	0.144		-				-				-			1.6	(0.7 – 3.6)	0.259	0.259
HIV status																				
Negative		-		-		-		-		-		-		-		-		Reference		
Positive		-				-				-				-			1.4	(0.6 – 3.3)	0.449	0.449
Income																				
No income		-				Reference				-				-				-		
Under $30,000		-		-	5.6	(0.6 – 52.1)	0.132	0.059		-		-		-		-		-		-
$30,000 and above		-			0.3	(0.0 – 3.2)	0.304			-				-				-		
Education																				
High school or less		-				Reference				-				-				-		
Some college		-		-	0.4	(0.1 – 2.2)	0.303	0.569		-		-		-		-		-		-
College or more		-			0.4	(0.0 – 5.4)	0.507			-				-				-		
Polysubstance use other than alcohol, marijuana, and poppers																				
Not Reported		-		-		Reference				Reference				-		-		Reference		
Reported		-			0.5	(0.1 – 2.3)	0.335	0.335	0.2	(0.1 – 0.6)	0.004	0.004		-			0.6	0.3 – 1.1)	0.097	0.097

	Stratified Validity Measure Models[a]																As Predictor of Concordance[b]			
	Sensitivity (n=209 visits among 73 participants)				Specificity (n=109 visits among 48 participants)				Positive Predictive Value (n=199 visits among 77 participants)				Negative Predictive Value (n=127 visits among 53 participants)							
	OR	(95% CI)	*p*-value	Wald *p*-value	OR	(95% CI)	*p*-value	Wald *p*-value	OR	(95% CI)	*p*-value	Wald *p*-value	OR	(95% CI)	*p*-value	Wald *p*-value	OR	(95% CI)	*p*-value	Wald *p*-value
CES-D score																				
<16		-				Reference				Reference				Reference				Reference		
≥16		-		-	0.3	(0.0 – 2.1)	0.231	0.231	0.3	(0.1 – 1.0)	0.057	0.057	0.5	(0.2–1.2)	0.107	0.107	0.4	(0.2 – 0.9)	0.024	0.024
Visit type																				
Baseline		Reference				Reference				-				-				-		
Follow-Up	0.3	(0.1 – 0.8)	0.024	0.024	10.1	(2.1 – 48.7)	0.004	0.004		-		-		-		-		-		-

[a]Multivariable GEE logistic regression models assessing participant and recall period characteristics as predictors of: (1) sensitivity, the odds of self reported methamphetamine use in the past three days among reporting periods with methamphetamine-positive urine; (2) specificity, the odds of a no self-reported methamphetamine use in the past three days among reporting periods with methamphetamine-negative urine; (3) positive predictive value, the odds of methamphetamine-positive urine among reporting periods with self-reported methamphetamine use in the past three days; (4) negative predictive value, the odds of methamphetamine-negative urine among reporting periods with no self-reported methamphetamine use in the past three days.

[b]Multivariable GEE logistic regression model assessing participant and recall period characteristics as predictors of concordance between self-reported methamphetamine use in the past three days and urinalysis results.