Addictive Behaviors 38 (2013) 2568–2574

Contents lists available at SciVerse ScienceDirect

Addictive Behaviors

Concordance between self-report and urine drug screen data in adolescent opioid dependent clinical trial participants

Claire E. Wilcox [a,*], Michael P. Bogenschutz [a,b], Masato Nakazawa [c], George Woody [d]

[a] University of New Mexico Department of Psychiatry, 1 University of New Mexico, MSC 09-5030, Albuquerque, NM 87131, USA
[b] Center on Alcoholism, Substance Abuse, and Addictions, University of New Mexico CASAA, 2650 Yale SE, MSC 11-6280, Albuquerque, NM 87106, USA
[c] Heritage College of Osteopathic Medicine, Ohio University, Grosvenor Hall, Athens, Ohio 45701, USA
[d] Perelman School of Medicine at the University of Pennsylvania and Treatment Research Institute, 500 Public Ledger Bldg. 150 South Independence Mall (W), Philadelphia, PA 19106, USA

HIGHLIGHTS

- We examine concordance between urine drug screen results and drug use self report.
- We examine factors that influence self-report validity.
- Self report is a valid measure of drug use.
- Adding urine tests improves detection of drug use.

ARTICLE INFO

Keywords:
Concordance
Treatment research
Self-report
Urine drug screen
Adolescent
Substance use disorder

ABSTRACT

Objective measures of drug use are very important in treatment outcome studies of persons with substance use disorders, but obtaining and interpreting them can be challenging and not always practical. Thus, it is important to determine if, and when, drug-use self-reports are valid. To this end we explored the relationships between urine drug screen results and self-reported substance use among adolescents and young adults with opioid dependence participating in a clinical trial of buprenorphine–naloxone. In this study, 152 individuals seeking treatment for opioid dependence were randomized to a 2-week detoxification with buprenorphine–naloxone (DETOX) or 12 weeks of buprenorphine–naloxone (BUP), each with weekly individual and group drug counseling. Urine drug screens and self-reported frequency of drug use were obtained weekly, and patients were paid $5 for completing weekly assessments. At weeks 4, 8, and 12, more extensive assessments were done, and participants were reimbursed $75. Self-report data were dichotomized (positive vs. negative), and for each major drug class we computed the kappa statistic and the sensitivity, specificity, positive predictive value, and negative predictive value of self-report using urine drug screens as the "gold standard". Generalized linear mixed models were used to explore the effect of treatment group assignment, compensation amounts, and participant characteristics on self-report. In general, findings supported the validity of self-reported drug use. However, those in the BUP group were more likely to under-report cocaine and opioid use. Therefore, if used alone, self-report would have magnified the treatment effect of the BUP condition.

1. Introduction

In treatment studies of patients with substance use disorders, obtaining valid drug use outcome data can be challenging. Urine test data are often used as primary outcome measures because self-reported drug use data can be invalid (Lavori et al., 1999; Winhusen et al., 2003). However, objective measures come with their own complications including high cost, varying and sometimes narrow windows for detection, and inaccuracy (Lavori et al., 1999; Winhusen et al., 2003). Rather than simply dismissing self-report data, it may be more useful to identify factors that influence the accuracy of self-reports and characteristics of individuals that are more, or less, likely to give accurate reports, as, often, self-report is an adequate measure of substance use (Babor, Steinberg, Anton, & Del Boca, 2000; Brown, Kranzler, & Del Boca, 1992; Del Boca & Darkes, 2003; Del Boca & Noll, 2000). Certain study design factors may increase accuracy, such as more rigorous information-gathering methods or not having contingencies for drug use (Darke, 1998; Del Boca & Noll, 2000; Sherman & Bigelow, 1992). Ongoing examination of this issue is important because of continuing changes in the nature,

* Corresponding author at: University of New Mexico Department of Psychiatry, 1 University of New Mexico, MSC 09-5030, Albuquerque, NM 87131, USA. Tel.: +1 505 272 2223.
E-mail addresses: cewilcox@salud.unm.edu (C.E. Wilcox), mbogenschutz@salud.unm.edu (M.P. Bogenschutz), nakazawa@ohio.edu (M. Nakazawa).

0306-4603/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.addbeh.2013.05.015

distribution, and demography of drug use. More research is needed to identify factors that may influence concordance between self-reports and urine toxicology test results, and how they vary in different populations. This study extends concordance examination to a new population: youthful opioid abusers being treated with a relatively new pharmacotherapy (buprenorphine) for opioid dependence (Woody et al., 2008).

The unique benefits of self-report procedures are flexibility, adaptability, relatively low cost, efficiency, portability, and the possibility of collecting data through a variety of technologies such as telephone, computer and even video (Del Boca & Noll, 2000). Some have found that self-reports are as sensitive as, and may sometimes be more sensitive than, objective measures when data are collected with clear instructions to respondents combined with methods to improve their motivation and facilitate cognitive processing (Babor et al., 2000; Del Boca & Noll, 2000). For example, a study of psychiatric patients in an emergency department showed that for marijuana, self-report was more sensitive than urinalysis (Perrone, De Roos, Jayaraman, & Hollander, 2001). Urine assay procedures can be inaccurate, increasing the relative validity of self-reports (Akinci, Tarter, & Kirisci, 2001; Brown et al., 1992; Jain, 2004; Magura, Goldsmith, Casriel, Goldstein, & Lipton, 1987; Perrone et al., 2001; Sherman & Bigelow, 1992; Solbergsdottir, Bjornsson, Gudmundsson, Tyrfingsson, & Kristinsson, 2004; Zanis, McLellan, Cnaan, & Randall, 1994).

Under-reporting of drug use may vary according to drug class, though there is little consensus on which classes are more affected (Brown et al., 1992; Darke, 1998; Falck, Siegal, & Carlson, 1992; Magura et al., 1987; Perrone et al., 2001; Sherman & Bigelow, 1992; Solbergsdottir et al., 2004; Zanis et al., 1994). Over-reporting use (reporting positive when urine screen is negative) may also occur but is less frequent than under-reporting, and findings of over-reporting may be due to the inaccuracy of the assay procedure (Akinci et al., 2001; Brown et al., 1992; Jain, 2004; Magura et al., 1987; Perrone et al., 2001; Sherman & Bigelow, 1992; Solbergsdottir et al., 2004; Zanis et al., 1994). Contingencies also affect the validity of self-reports. For example, patients applying for methadone treatment may over-report opioid use because they are afraid that they will not qualify for treatment or that the physician will not prescribe a dose that prevents withdrawal (Digiusto, Seres, Bibby, & Batey, 1996; Sherman & Bigelow, 1992), while those on methadone treatment may under-report to avoid disapproval, termination of treatment, or loss of take-home privileges. Other contextual factors may also affect self-report accuracy, such as whether interviewers are para-professionals or professionals, the way questions are asked, whether strategies to enhance recall are used, conditions under which the data are obtained (treatment, research, occupational), perceived confidentiality, and whether the patient directly enters self-reports into a computer or provides them during an interview with a clinician or research technician (Del Boca & Noll, 2000; Digiusto et al., 1996; Schumacher et al., 1995; Sherman & Bigelow, 1992).

Finally, patient factors can also influence the validity of self-report. For example, pregnancy is associated with more under-reporting, likely related to fear of losing custody or criminal retribution (Marques, Tippetts, & Branch, 1993). Employment, African-American race, diagnosis of histrionic personality disorder and cognitive deficits have been associated with under-reporting, whereas diagnoses of dependent personality, passive-aggressive personality or axis I affective disorders have been associated with less under-reporting (Babor et al., 2000; Del Boca & Noll, 2000; Fendrich, Mackesy-Amiti, Johnson, Hubbell, & Wislar, 2005). Some (Solbergsdottir et al., 2004) but not all studies (Kilpatrick, Howlett, Sedgwick, & Ghodse, 2000) have found younger age to be negatively correlated with under-reporting; adolescents may be especially influenced by social pressure of peers, characteristics of the adult examiner, and perceived threat to confidentiality (Schwarz, 1999). Factors that have not reliably been predictive include gender, past criminality, and antisocial personality disorder (Digiusto et al., 1996; Magura et al., 1987).

In view of these inconsistent findings on the validity of self-reports, we conducted a secondary analysis of self-report and urine test data from a randomized trial of buprenorphine–naloxone treatment for opioid addicted youth done by the NIDA Clinical Trials Network (CTN) (Woody et al., 2008). Although the primary outcome was opioid use as measured by urine test results at weeks 4, 8 and 12, weekly self-report and urine test data were collected on use of cocaine, opiates, amphetamines, benzodiazepines, and cannabis. These data allowed us to explore predictors of concordance between urine drug tests and self-reports. Consistent with existing evidence, we hypothesized that concordance would be reasonably high for most drugs, and that self-report would be more specific than sensitive since patients tend to under-report more than over-report. We also hypothesized that self-reported positives would be lower in the BUP group than in the DETOX group regardless of drug screen results due to greater engagement in treatment and desire to please the providers. Finally, in an exploratory analysis, we evaluated other subject factors that were previously shown to be associated with the validity of self-reported drug use.

2. Materials and methods

2.1. Participants and outcomes

In the parent study, 152 subjects aged 15–21 seeking treatment for opioid dependence were randomized to a 2-week detoxification with buprenorphine–naloxone (DETOX; N = 78), or 12 weeks of buprenorphine–naloxone (BUP; N = 74), with a dose taper beginning in week 9 and ending in week 12, each with weekly individual and group drug counseling (Woody et al., 2008). Subjects were paid $5 for weekly assessments which included urine drug screen and self-report of drug use, and $75 for more extensive assessments at weeks 4, 8 and 12. Weekly assessments took approximately 30 min, and monthly assessments (weeks 4, 8, and 12) took approximately 90 min. Participants were asked “In the past week how many days did you use: [heroin, methadone, other opiates, benzodiazepines, cocaine, amphetamines, methamphetamines, cannabis?]” A dichotomous self-report response was created as follows: for cocaine, cannabis, and benzodiazepines, if participants indicated non-zero days of use, the response was coded as “1” for each drug; otherwise, as “0”. For amphetamines, participants' responses to methamphetamine and amphetamine were first combined, and non-zero responses in either group were coded as “1”. Similarly, for opioids, participants' responses to heroin, methadone, and other opiates were first combined, and non-zero responses were coded as “1”. The same questions were used for the more extensive monthly assessments.

The urinalyses for drugs of abuse were performed on site utilizing the SureStep drug screen card (which tests for all drugs noted above except oxycodone but does include a test for tricyclic antidepressants) and the Rapid One OXY on-site urine drug screen for oxycodone. Cutoffs in ng/ml were as follows: amphetamines (1000 ng/ml), barbiturates (300 ng/ml), benzodiazepines (300 ng/ml), cocaine (benzoylecgonine) (300 ng/ml), methadone (300 ng/ml), methamphetamine (1000 ng/ml), morphine (hydrocodone, hydromorphone, heroin) (2000 ng/ml), phencyclidine (PCP) (25 ng/ml), tetrahydrocannabinol (THC) (50 ng/ml) and oxycodone (100 ng/ml). Urine samples were assigned a positive or negative value for each of the following five groups: opioids (morphine/opiates, methadone, and oxycodone), cocaine, cannabis, benzodiazepines, and amphetamines (methamphetamine and amphetamine).

Based on these values, five measures of concordance of self-report with urine samples were computed: Cohen's kappa (κ), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). In this case we used the urine toxicology result as the “gold standard”; thus “true positive” was defined as having a positive urine toxicology screen result. κ is a statistical measure of inter-rater