Illicit Substance Use Among a Sample of Subsidized Housing Residents: Concordance, Longitudinal Trends, and Quality of Life

DISSERTATION

Presented to the School of Public Health

Behavioral and Community Health
University of North Texas Health Science Center
Fort Worth, TX 76107

In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy
by

Alexis Rendon, MPH, CPH
PhD Candidate

April 2019

Table of Contents

Acknowledgments

Special thanks to Dr. Scott Walters, Dr. Emily Spence-Almaguer, Dr. Eun-Young Mun, Dr. Melvin Livingston, and Dr. Sumihiro Suzuki for their contributions to this dissertation. In addition, recognition of the entire m.chat team, whose hard work made this dissertation possible.

List of Tables

List of Figures

Abstract

This three-paper model dissertation investigates issues related to self-reported substance use. Self-report is a less invasive and expensive method of collecting substance use behavior when compared to a toxicological test, but the self-report method has been shown to be unreliable in some populations. We found that self-report missed some use captured by a saliva toxicological test administered to a subsidized housing population enrolled in a technology-assisted health coaching program. Concordance was highest among marijuana users and increased over time. Higher rates of concordance were found when the recall window was expanded from a restricted biological recall window to match the toxicological test to the full 90 day window of the Timeline Follow-Back. Participants who reported using substances more frequently reported having more problems related to their substance use. We also found that both substance use problems and the frequency of consumption of a combined Other category of substances, including cocaine, amphetamine, methamphetamine, opiates, prescription pills, or phencyclidine were predictive of lower quality of life. This dissertation validates previous literature indicating that self-report is a fair to moderately good measure of actual substance use behavior in vulnerable populations that may intentionally or unintentionally misreport their substance use. Programs limited to self-reported measures may consider widening their recall windows to increase accuracy of self-report.

Chapter 1: Introduction

1.1. Study Rationale

Substance use disorders pose a significant barrier to maintaining stable housing (Willenbring, Whelan, Dahlquist, & O'Neal, 1990) and are associated with increased health care costs (Kessell, Bhatia, Bamberger, & Kushel, 2006), especially among those with low-income or comorbid health conditions. Housing assistance programs for low-income households, often referred to as subsidized housing, have been created to aid rent burdened households. Because of the well-established relationship between homelessness and substance use problems (Stein, Dixon, & Nyamathi, 2008; Torchalla, Strehlau, Li, & Krausz, 2011), many substance users are in need of both shelter and substance use treatment. There are a variety of different types of housing programs considered to be subsidized housing. Housing assistance can range from cost-sharing with the tenant to fully subsidized rent. Programs can be temporary or long-term and can also include supportive, case-management services. Subsidized housing programs also vary in their requirement of sobriety from psychoactive substances. Housing programs that require continued abstinence from alcohol or psychoactive substances can be a barrier for individuals with substance use disorders to maintain housing (Morgenstern, Hogue, Dasaro, Kuerbis, & Dauber, 2008; Palepu, Marshall, Lai, Wood, & Kerr, 2010; Watson, Wagner, & Rivers, 2013; Willenbring, Whelan, Dahlquist, & O'Neal, 1990). Housing programs that require substance use abstinence operate on the Treatment First Model, in which on-going substance use is addressed first before the need for "shelter" or housing assistance. In contrast, the true Housing First Model operates on the idea that homeless substance users should be provided housing first, before any health behavior modifications regarding substance use can be made. This is based on the idea that the combination of housing and supportive services provides the highest probability of success in both reducing health care costs and decreasing substance use (Kirst, Zerger, Misir, Hwang, & Stergiopoulos, 2015). This concept has been supported by findings of reduced alcohol use (Kirst et al., 2015) and a 53% decrease in public service and medical costs (Larimer et al., 2009) compared to the treatment first approach.

Although the effect of supportive housing programs on health care expenditures has been well established, the effect of supportive housing on reducing substance use among tenants remains under debate. Lack of stable, affordable housing has been cited as a barrier to substance use treatment and abstinence (Zerger, 2002). Despite predictions of greater success when compared to traditional Treatment First Model, studies of supportive housing programs have found higher rates of illicit substance use (Edens, Mares, Tsai, & Rosenheck, 2011). Substance users had poorer mental health outcomes and subjective quality of life scores when compared to their non-using counterparts across

multiple measures of mental health status (Edens et al., 2011), including the Study 12-Item Short Form (SF-12), three subscales of the Brief Symptom Inventory (BSI), and observed psychotic behavior rating scale. Several studies have reported declines in alcohol use, but not in illicit substance use after entering supportive housing (Kirst et al., 2015; Larimer et al., 2009; Padgett, 2006). In contrast, another study reported that supportive housing residents increased substance treatment services by 22% after being housed, which resulted in decreased substance use (Mondello & House, 2007).

A major limitation of these studies that attempt to determine the effect of supportive housing services on substance use behavior is that they tend to rely exclusively on self-report as a measure of psychoactive substance use (Kirst et al., 2015; Padgett, 2006). Self-report may be a biased measure of true substance use in populations with housing insecurity due to concerns about the accuracy of recall and the influence of social desirability (Padgett, 2006). This bias is of particular concern in populations with co-occurring mental illness, substance use, and homelessness. Kirst et al. (2015) used self-report measures to ascertain alcohol and illicit substance use to compare supportive housing treatment to treatment as usual (TAU) with no supportive services. This study found that supportive housing clients reduced a greater number of alcohol problems over time compared to TAU, but found no treatment effect on substance use. Kirst et al. list the method of data collection exclusively as self-report may underestimate the prevalence of illicit substance use in this population. Underestimated substance use could lead to inaccurate conclusions about the effects of types of subsidized housing programs on substance use behavior.

Many studies have conducted tests on the validity of self-reported substance use and found evidence that certain groups may be more likely to underreport substance use, compared to biological tests (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). For example, Schumacher et al. (1995) found an average 30-day concordance rate of 68% among 131 homeless crack cocaine users, compared to urinalysis. This finding was relatively low when compared to the agreement between self-report and biological test for other populations. A meta-analysis found that the percent agreement between self-report and biological measures of substance users ranged from 87.3%–90.9% for marijuana and 79.3%–84.1% for cocaine (Hjorthøj et al., 2012). Among studies in which substance users had psychiatric co-morbidities, the percent agreement ranged from 80.4%–83.8% (Hjorthøj et al., 2012). Estimates of substance use may be underestimated by up to 34.9% depending on the substance and population (Hjorthøj et al., 2012; Schumacher et al., 1995). If self-report is an inaccurate measure of the substance use behavior of individuals on subsidized housing, conclusions

made about the progress of reducing or eliminating substance use in this population may be invalid. Before conclusions are made about the effects of supportive housing programs on substance use, measuring substance use via self-report must be validated as a reliable measure of actual substance use in low-income housing residents.

In regards to substance use research, the concept of concordance refers to the agreement between self-reported drug use and toxicological confirmation (Figure 1). Comparison of self-reported substance use and toxicological confirmation results in the formation of four groups: concordant users (self-report and toxicology positive), concordant non-users (self-report and toxicology negative), and two types of discordant users. Discordant users are either underreporting (self-report negative, toxicology positive) or over reporting (self-report positive, toxicology negative). There are various types of toxicology media, include sweat, blood, urine, saliva, and hair. Each has different windows of detection and varies in degree of invasiveness (Dolan, Rouen, & Kimber, 2004). Compared to urinalysis, saliva sample testing has low invasiveness and similar levels of sensitivity (Bennett, Davies, & Thomas, 2003).

	Toxicology Positive	**Toxicology Negative**
Self-Report Positive	Concordant User	Discordant User (False Positive)
Self-Report Negative	Discordant User (False Negative)	Concordant Non-user

Figure 1. Two-by-Two Table of Substance Use Concordance

Verbal recall of substance use is common in research studies assessing drug use behavior (Darke, 1998; Hjorthøj et al., 2012; Schumacher et al., 1995). Various approaches aimed at increasing the reliability of the self-report of substance users has led to the development of methods like the Timeline Follow-back (TLFB), which utilizes a visual calendar to enhance recall of substance use (L. Sobell & Sobell, 1992). Originally developed to measure alcohol consumption, the TLFB has since been widely used in cross sectional and prospective studies of drug use (Hjorthøj et al., 2012). While the TLFB is generally concordant with biological measures, some populations may be more accurate in their reporting (Harrison, 1997; Hjorthøj et al., 2012; Napper et al., 2010; Rosay, Najaka, & Herz, 2007; Secades-Villa & Fernandez-Hermida, 2003).

While self-reported measures of substance use are easier to administer compared to other toxicological tests like hair, blood, saliva, or urine samples, they have not been sufficiently validated in populations of individuals on subsidized housing programs. Our previous study, Rendon et al. (2017), found that the concordance of self-reported substance use reported via TLFB ranged between 43.8% -

69.7%, depending on the substance type. Further study is needed to validate self-report via TLFB in subsidized housing residents. Greater understanding about the substance use behaviors of subsidized housing residents is needed to further develop effective supportive services for tenants- including substances users who accurately report their substance use and those who misreport their substance use behaviors.

There are a number of reasons substance users may intentionally or unintentionally misreport their substance use. Some of the reasons for underreporting may be lack of standardized methodology, social desirability or stigma of reporting illicit substance use, inability to recall substance use further back in time, and fear of legal repercussions (Napper et al., 2010). In order to accurately report substance use, individuals must first be aware of their behavior. This memory process has been previously described by Matt et al. as an individual's effort in retrieving information incorporated with context, timing proximity, retrieval cues (2003). Those with mental illnesses may have a more difficult time recalling substance use because of the increased effort needed to retrieve or remember this behavior. Next, individuals go through an estimation process when retrieving information about substance use behavior. This "editing" occurs when it is impossible or undesirable to report on all instances of a behavior (Matt et al., 2003). For instance, asking someone how many times they used in the last 90 days requires both memory and estimation. Individuals often extrapolate a smaller time period (i.e. number of times used in a week) to a larger time period to estimate a response. It is possible that some individuals inaccurately estimate their true behavior without the intent to deceive. In contrast, individuals may intentionally increase or decrease their estimate depending on their judgment of social desirability. A well-studied phenomenon, social desirability occurs when an individual edits their response in order to "obtain approval by responding in a culturally appropriate and acceptable manner" (Crowne & Marlowe, 1960). The effects of social desirability on self-reported substance use may influence an individual one of two ways. First, individuals may exaggerate their willingness to change resulting in amplification of a reported problem (Zemore, 2012). Social desirability may also influence individuals to minimize their substance use. Stigma associated with substance use may influence individuals to diminish their substance use behavior. Individuals with criminal justice backgrounds may be especially wary of reporting substance use behavior, even when given assurances of confidentiality. Often, individuals may not share illicit substance use behavior due to stigma or fear of negative consequences such as losing a resource like housing or limitations in medical treatment or medications (Clark, Zyambo, Li, & Cropsey, 2016). Both concordant and discordant reporters of substance use go through the same process of recall, which involves memory and estimation. Estimates

can be miscalculated without ill intent. They may also be inflated or diminished by the effects social desirability on a reported behavior. If services like subsidized housing are contingent upon sobriety, individuals who relapse or continue to use psychoactive substances may have greater motivation to misrepresent their substance use.

While substance use concordance is relatively well-studied as a cross-sectional phenomenon in substance use literature, fewer studies have followed how concordance changes over time. Findings regarding how concordance changes with time are mixed- concordance has been shown to both increase and decrease over time. One study conducted among adolescent minority substance users found that concordance increases over time (Dillon, Turner, Robbins, & Szapocznik, 2005). A longitudinal study of 500 individuals on probationers found that concordance decreased as time went on by comparing self-report collected via Addiction Severity Index for past 30 day use to urine drug screen assays (Clark et al., 2016). This study found a time by substance interaction, in which more stigmatized substances like cocaine and opiates were denied at higher rates than less stigmatized substances like marijuana and alcohol. Additionally, those who were at risk for higher levels of stigma were more likely to misreport their substance use (males, African Americans, younger age) (Clark et al., 2016). Although Clark et al. add valuable information to the topic, some gaps remain unfilled. First, this study focused on probationers and not housing residents. While it is possible that there is some overlap, this study should be replicated in subsidized housing residents. Second, Clark et al., use a urine test. This toxicological method is often used as a gold standard, but is less invasive than a saliva sample. While toxicological tests should capture the same information, the effects of the test themselves may influence participation in the study. Many probationers are required to give urine samples as terms of their probation. Saliva sampling is a less invasive way to capture a biological measure of substance use in a population that may be resistant to giving urine samples. Finally, Clark et al. did not use the TLFB method for collecting substance use among their participants. While similar in design by comparing a self-report of substance use behavior to a biological measure, we propose to compare a different method of self-report and a different biological measure in a different population. Few studies on the topic of concordance over time make it difficult to form a consensus on the topic. It is currently unknown if individuals receiving subsidized housing become more concordant as they progress through a program (Schumacher et al., 1995). More studies are needed in various types of populations to add further information to the debate about how concordance changes over time.

While the effects of time on concordance are unknown among subsidized housing residents, there are rationales to support either outcomes- concordance increasing or decreasing with time. For instance, the argument that people become more accurate in their self-report of substance use over time is supported by the idea that they may be more comfortable disclosing information that they previously edited because it was considered socially undesirable. The argument that people become less honest about substance use over time is supported by the idea that substance users are influenced by perceived social desirability among those delivering the intervention (Napper et al., 2010). In this case, participants may minimize their substance use because of their relationship with their interventionist. Using the same logic, it is possible that social desirability may influence a substance user to exaggerate substance use frequency or intensity or even solicit misrepresentation about substance using behaviors from a non-user. Studied extensively in adolescents, social desirability has not been adequately studied in adult substance users with co-morbid mental conditions. Additionally, participants may be affected by questions posed to them about their substance use behavior. This phenomenon is known as assessment reactivity and has been well-established in the literature in regards to self-reported drinking behavior (McCambridge & Kypri, 2011; Moyer, Finney, Swearingen, & Vergun, 2002; LC Sobell & Sobell, 1981; Walters, Vader, Harris, & Jouriles, 2009). Assessment reactivity affects the individual's ability to recall behavior. Questions asked to participants can serve as their own retrieval cue in the memory process of recalling behavior. Information unintentionally given during assessments may influence an individual's future estimation of that behavior. For instance, asking someone how many standard servings of alcohol they have consumed in the last 90 days while defining that a standard serving is 1.5 oz. of liquor, 5 oz. of wine, or 12 oz. of beer may inadvertently influence the individual's future consumption of alcohol. This increased or heightened awareness could lead to actual decrease in the amount of alcohol or frequency of illicit substances consumed. It could also make individuals better reporters of their own use. The impact of knowing if concordance increases or decreases over time could have an impact on how substance use data is collected for longitudinal interventions. For example, if concordance decreases over time, it may be useful to implement biological testing into an exclusively self-reported program to increase accuracy of true substance use at follow-up. In contrast, it may be sufficient to use self-reported substance use only at follow-up if concordance increases with time.

In addition to studying the validity of self-reported substance use behavior, it is important to study other behavioral factors associated with substance use. One known correlated concept is perceived quality of life. Individuals who use substances have poorer quality of life outcomes compared to those who do not use substances (Havassy & Arns, 1998; McKenna et al., 1996). Quality of life is a

useful measure in health services research with chronic disorders where the goal is improving the patient living condition, rather than complete absence of symptoms (Wood-Dauphinee, 1999). For substance users with chronic physical or mental health conditions, quality of life is an important measure of the patient's subjective view of their overall well-being (Pasareanu, Opsal, Vederhus, Kristensen, & Clausen, 2015). Determining the rate of concordance and the validity of TLFB in a sample of subsidized housing residents will be more useful with the addition of information about related behaviors like quality of life. Information about perceived quality of life may lend more information to generate further hypotheses about the possible change or lack of change of the concordance rate. For instance, if quality of life scores improve and concordance rate increases, we may recommend further study into the relationship between quality of life and concordance to determine the specific mechanism of action.

1.2. Research Aims

This study will use data collected by the 1115 Medicaid Waiver Program: Mobile Community Health Assistance for Tenants (m.chat). The data used for analyses were collected from 644 adult low-income housing residents who were participating in a health coaching program. The m.chat program was a health coaching program for tenants of subsidized housing. The sample is unique because it enrolled participants with mental health conditions, often an exclusion criterion in other substance use studies. Additionally, the m.chat program was not a substance use treatment study. Substance use was not required for eligibility, but it was measured two ways at baseline and each 6-month follow-up. This study utilized both a subjective self-report and an objective biological measure of recent substance use behavior in this population with high prevalence of psychoactive substance use. This allows for the comparison of self-report to the oral fluid substance test to determine the accuracy of the subjective measure. The following specific aims will be addressed:

Aim 1.

Determine the rates of concordance and discordance among individuals who use psychoactive substances at baseline and demographic differences between concordant and discordant users. Specifically:

Do rates of concordance differ by substance type (marijuana, amphetamine/methamphetamine/PCP, and cocaine) among residents of subsidized housing programs? If so, should self-report be continued to be used as an exclusive measure of substance use behavior in this population?

Many studies measuring substance use have exclusively relied on self-reported information (Kirst et al., 2015; Padgett, 2006). Despite evidence that people who belong to particularly vulnerable populations tend to underreport psychoactive substance use, compared to biological tests (Digiusto et al., 1996; Goldfinger et al., 1996; Harrison, 1997; Napper et al., 2010; Schumacher et al., 1995; Sloan et al., 2004) self-report remains primarily the exclusive measure of substance use based on cost. While comparisons between self-report and biological measures have been conducted in other populations, the self-report and biological measure from a sample of subsidized housing residents have yet to be compared. Data from the m.chat population will be used to assess the rates of concordance and discordance for three substance groups: marijuana, amphetamine/methamphetamine/PCP, and cocaine.

Do demographic variables or other self-reported behaviors predict underreporting discordance at baseline? If so, how can this information be used to advise other researchers or interventionists?

Predictive variables could be used by other researchers or interventionists in estimating the rate of misreporting substance use in demographically similar populations where biological measures are not feasible.

Aim 2.

Examine if the rate of concordance between self-reported substance use and nonuse, and oral fluid test differ across time (from baseline to 6, and 12 months) for marijuana, cocaine, and amphetamine/methamphetamine/PCP.

While it is known that self-report can be unreliable in other populations, we will use m.chat data to determine if the rate of concordance changes at follow-up compared to baseline. Determination of the effect of the concordance rate of illicit substance use over time is necessary before studying what causes the change in concordance. Based on our findings, we can make future recommendations on how to approach the underlying mechanism of change regarding concordance. For instance, if we find that concordance increases over time, we can recommend future studies investigate both the effects of positive social desirability and assessment reactivity to be able to determine what causes the increase of concordance over time. In contrast, if we find the opposite to be true, that concordance decreases over time, we may recommend that future studies investigate negative social desirability, fear of negative repercussions, and assessment reactivity. By determining if the concordance rate changes at follow-up, this study will be able to generate hypotheses regarding why the concordance rate changed in this sample. Additionally, knowing how the concordance rate changes over time may have practical

implications for subsidized housing programs. Future programs may consider the consequences of exclusively relying on self-report for longitudinal substance use treatment programs and interventions.

Aim 3.

Examine if self-reported substance use frequency predicts perceived quality of life. Does increased consumption of different substances lead to more severe substance use problems? Are substance use problems predictive of quality of life?

Similar to Manuscript 2, Manuscript 3 will also use baseline and follow-up data from the m.chat program. Substances users defined as a positive self-report while controlling for age, race, and gender. Outcomes of interest in Manuscript 3 include quality of life measured by the Quality of Life Satisfaction Questionnaire (Q-LES-Q). This variable is associated with substance use and was the most convenient outcomes to assess alongside substance use.

1.3. Methods

Study Design: m.chat Program

The Mobile Community Health Assistance for Tenants (m.chat) project began recruitment for the intervention in 2014 and finished intervention delivery in December of 2017. Funded by the Medicaid 1115 Waiver program in Texas, this program combined in-person health coaching and technology to address six different health domains: substance use, diet, exercise, medication compliance, social support, and recreation/leisure (Walters & Spence-Almaguer, 2015). The m.chat program eligibility requirements included PSH residency (November 2014 – October 2016) or any subsidized housing residency (November 2016 – December 2017). The program was expanded in November 2016 to serve residents of other housing programs. PSH residents were continued to be enrolled in the study after the eligibility expansion. Participants were required to be at least 18 years old and either Medicaid enrolled or low-income and uninsured at the time of screening. Participants were recruited from Tarrant County from housing lists provided by housing agencies, via case manager referrals, and word of mouth referrals. In addition, participants had to self-report at least one of the following mental health problems: prescribed medication for psychological or emotional problems, experienced hallucinations in the past year, currently receiving a pension for a psychiatric disability, or scoring ≥9 on the 9-item Patient Health Questionnaire (PHQ-9) depression screener (Kroenke, Spitzer, & Williams, 2001). Exclusion criteria included (1) residing in other types of housing not considered PSH (e.g., Transitional Housing or homeless shelter) before November 2016, (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/or (4) limited autonomy or decision-making capabilities (e.g., substantially neurologically or cognitively impaired). Participants were given assurances of confidentiality. Informed consent was obtained from each study participant. The project was approved by the Institutional Review Board (IRB) at the University of North Texas Health Science Center (UNTHSC).

Assessments were completed at baseline and 6-month intervals. At each assessment, participants were asked about their substance use in the past 90 days via the Timeline Follow-back (TLFB) method and were given an oral fluid test. The baseline, 6-month, and 12-month follow-ups will be used to categorize individuals by substance use status: non-user (both TLFB negative and oral fluid test negative), concordant user (both TLFB positive and oral fluid test positive), or discordant user (TLFB negative and oral fluid test positive and TLFB positive and oral fluid test negative).

To our knowledge, only our previous study Rendon et al. (2017) has estimated the prevalence of substance use in a supportive housing or a similar low-income adult population with mental health disorders using a biological measure to validate self-report. While the TLFB method has been validated in other populations, this method has still been shown to have a high rate of discordance in adults with mental illnesses (Stasiewicz et al., 2008). Instead of using a toxicological measure, the TLFB or other self-reported methods are often used as the primary measure of substance use in this population. This poses a problem when TLFB is used as the exclusive measure in this population. Measures used in populations for which they have not been validated can lead to poor data collection and incorrect conclusions.

Due to the voluntary nature of the health coaching program, results may not generalize to housing residents who are not motivated to make healthy behavior changes. Findings from this research may help those wanting more accurate information regarding substance use in low-income housing populations to more effectively address differential needs for both substance users and non-using clients.

ACCEPTED FOR PUBLICATION, PENDING FINAL PDF

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, MPH[1]

Eun-Young Mun, Ph.D.[1]

Emily Spence-Almaguer, Ph.D.[1]

Scott Walters, Ph.D.[1]

[1]University of North Texas Health Science Center

School of Public Health

Department of Health Behavior and Health Systems

Fort Worth, TX

Correspondence and requests to: Scott Walters, PhD, School of Public Health, The University of North Texas Health Science Center, 3500 Camp Bowie Blvd, Fort Worth, TX 76107; Email: scott.walters@unthsc.edu.

Acknowledgement
This study was supported by CMS DSRIP 138980111.2.6 and in part by R01 AA019511.
Funding: 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.

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 (L. 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 et al., 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 et al., 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 et al., 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 et al., 2012).

Studies examining the agreement between self-report and toxicological confirmation show that some groups tend to underreport substance use (Digiusto et al., 1996; Goldfinger et al., 1996; Harrison, 1997; Napper et al., 2010; Polcin, 2016; Schumacher et al., 1995; Sloan et al., 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., 2004a; 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 et al., 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 (Morgenstern et al., 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., 2004b, 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 et al., 2004; Clark et al., 2016; Dillon et al., 2005; Rohsenow, Martin, Eaton, & Monti, 2007). For instance, Babor et al. (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 et al., 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 (L. 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 hours (3 days) for cocaine, PCP, amphetamine, and methamphetamine and up to 120 hours (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 hours) 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 $1,000 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.

Table 1. Participant Demographics

		Baseline *N* = 644		Follow-up 1 *N* = 516		Follow-up 2 *N* = 361	
Demographics		*n*	%	*n*	%	*n*	%
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				
		Test+		Test-			Test+		Test-			Test+		Test-	
Marijuana	**TLFB+**	**46**	***44.2%***	**58**	***55.8%***	104	**TLFB+**	**30**	***57.7%***	**22**	***42.3%***	52	**TLFB+**	**31**	***96.9%***	**1**	***3.1%***	32
	TLFB-	**36**	***6.7%***	**500**	***93.3%***	536	**TLFB-**	**26**	***5.7%***	**430**	***94.3%***	456	**TLFB-**	**25**	***7.7%***	**300**	***92.3%***	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)** 0.50 (0.38, 0.62)						**Cohen's kappa (95% CI)** 0.67 (0.55, 0.78)					
		Test+		**Test-**				**Test+**		**Test-**				**Test+**		**Test-**		
Cocaine	**TLFB+**	**41**	***87.2%***	**6**	***12.8%***	47	**TLFB+**	**6**	***46.2%***	**7**	***53.8%***	13	**TLFB+**	**5**	***55.6%***	**4**	**44.4%**	9
	TLFB-	**87**	***14.8%***	**501**	***85.2%***	588	**TLFB-**	**41**	***8.3%***	**454**	***91.7%***	495	**TLFB-**	**35**	***10.1%***	**313**	***89.9%***	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 (0.03, 0.30)						**Cohen's kappa (95% CI)** 0.17 (0.02 0.32)					
		Test+		**Test-**				**Test+**		**Test-**				**Test+**		**Test-**		
PCP, Amphetamine, Methamphetamine	**TLFB+**	**17**	***81.0%***	**4**	***19.0%***	21	**TLFB+**	**1**	***14.3%***	**6**	**85.7%**	7	**TLFB+**	**7**	***77.8%***	**2**	***22.2%***	9
	TLFB-	**32**	***5.2%***	**582**	***94.8%***	614	**TLFB-**	**73**	***14.8%***	**420**	**85.2%**	493	**TLFB-**	**21**	***6.0%***	**327**	***94.0%***	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)** -0.00 (-0.05, 0.04)						**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

		Substance Group (OR, 95% CI)					
		Marijuana		Cocaine		Amphetamine, Methamphetamine, PCP	
Time Point	**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		**Chi-Square**	***p*-value**	**Chi-Square**	***p*-value**	**Chi-Square**	***p*-value**
***df* = 2**		18.94	<.01	5.12	0.08	22.66	<.01

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 et al., 2004; Godley et al., 2002; Stasiewicz et al., 2008). The agreement for cocaine was higher than previous studies which ranged between 70% - 87% (Carroll et al., 2004a; 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 hours 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 hours (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.

Quality of Life and Substance Use Frequency Among a Sample of Subsidized Housing Residents

Alexis Rendon, MPH[1]

Eun-Young Mun, Ph.D.[1]

Emily Spence-Almaguer, Ph.D.[1]

Scott Walters, Ph.D.[1]

[1]University of North Texas Health Science Center

School of Public Health

Department of Health Behavior and Health Systems

Fort Worth, TX

Correspondence and requests to: Scott Walters, PhD, School of Public Health, The University of North Texas Health Science Center, 3500 Camp Bowie Blvd, Fort Worth, TX 76107; Email: scott.walters@unthsc.edu.

1. Introduction

Quality of life (QoL) is a useful measure in health services research where the goal is improving the patient living condition, rather than complete absence of symptoms (Wood-Dauphinee, 1999). Defined by the Food and Drug Administration as a "multi-domain concept that represents the patient's perception of the effect of illness and treatment on physical, psychological, and social aspects of life" (U.S. Department of Health and Human Services: Food and Drug Administration, 2009), QoL can be measured as an overall subjective evaluation of multiple categories like physical health, social interactions, and satisfaction with economic and living situations (Pasareanu et al., 2015). The subjective nature of QoL offers a unique assessment by the individual experiencing the behavior. This subjective evaluation is not well-captured by other measures besides QoL (Tiffany, Friedman, Greenfield, Hasin, & Jackson, 2012). There have been many instruments designed to study QoL. A meta-analysis discussed four historically different conceptions of quality of life due to variance in definition and the multidimensional nature of the concept of quality of life: objective QoL, subjective experiences, incorporated objective and subjective domains, and health-related QoL. The Quality of Life Enjoyment and Satisfaction Questionnaire Short Form (Q-LES-Q-SF) is derived from Endicott's original instrument designed to capture subjective experience regarding enjoyment and satisfaction of daily functioning (Endicott, Nee, Harrison, & Blumenthal, 1993). This abbreviated version is the most frequently used QoL instrument in both psychopharmacology and clinical trials (Stevanovic, 2011).

QoL is an important concept in substance use research due to the subjective nature of the concept of QoL and because QoL can function as a measure of the impact of the substance use on actual functioning across multiple domains (Tiffany et al., 2012). Substance use behavior is defined as a substance use disorder (SUD) via the Diagnostic and Statistical Manual of Mental Disorders fifth edition (DSM-5) when at least two symptoms are met regarding: 1) increased consumption over amount or period intended, 2) unsuccessful quit attempts, 3) resources like time spent on obtaining, using, and recovering from using a substance, 4) the presence of cravings, 5) failure to fulfill work, school, or home obligations, 6) continued use despite social or interpersonal problems caused by substance use, 7) given up or reduced social, occupational, or recreational activities, 8) continued use despite physically hazardous situations, 9) continued use despite persistent or recurrent physical or psychological problem caused or exacerbated by substance use, 10) tolerance, and 11) symptoms of withdrawal (American Psychiatric Association, 2013).

The limited research on QoL in non-treatment seeking, illicit substance using populations tends to find an inverse relationship between alcohol consumption and QoL. In a study of 1,336 primary care, non-treatment seeking patients, a higher frequency of alcohol consumption was associated with lower QoL compared to individuals who drank less amount or less frequently (Donovan, Mattson, Cisler, Longabaugh, & Zweben, 2005; Volk, Cantor, Steinbauer, & Cass, 1997). A recent meta-analysis of 14 studies regarding recreational marijuana and QoL found that a threshold effect may exist in which frequent use of marijuana resulted in lower QOL scores (Goldenberg, IsHak, & Danovitch, 2017). This study calls for further investigation into the actual threshold value at which frequency or dosage of marijuana begins to impair QoL. Two of the studies included in this meta-analysis found that smoking marijuana more than once per week resulted in lower QoL scores compared to non-users in Switzerland (Dey, Gmel, & Mohler-Kuo, 2013; Dey, Gmel, Studer, & Mohler-Kuo, 2014). Lower QoL may be indicative of greater treatment readiness. One of the main reasons substance users might seek treatment is the negative consequences associated with substance use behavior (Laudet, Becker, & White, 2009). If QoL is primarily measured in only treatment-seeking substance users with lower QoL scores experiencing negative consequences, this may lead to an underestimation of the QoL scores among the general population of users.

Individuals with SUDs often experience problems that are disruptive to daily functioning outside of the immediate realm of substance use (Tiffany et al., 2012). Because these consequences can cause substantial distress to the user, it is the impact of these consequences caused by problems that underlie the main societal concern about addiction. SUDs affect QoL via consequences related to substance use that impact life functioning. Lower perceived QoL may be influenced by consequences associated with substance use either directly, like criminal justice interactions or indirectly, like comorbid health problems. For example, QoL may be indirectly affected by a cough that occurred as a result of smoking marijuana. This consequence may indirectly affect QoL if the cough prevents the individual from doing their normal functioning tasks. A consequence with a more direct effect on QoL could be incarceration due to substance use. This consequence would directly impact an individual's ability to carry out their normal daily functions and impede their life enjoyment. The most effective treatments for substance use are targeted at reducing or ameliorating negative consequences associated with substance use behavior (Tiffany et al., 2012). This type of harm-reduction approach addresses clinically and socially relevant consequences of substance use. Therefore measuring substance use problems and consequences is important for determining treatment program efficacy.

Similar to the ASI, the Inventory of Drug Use Consequences (InDUC) was designed as an instrument to measure the severity of these types of direct and indirect adverse consequences of poly-substance use on five scales: Impulse Control (harmful risk-taking), Social Responsibility (recognition of a lack of fulfillment of personal responsibilities like school or work), Physical (harm to physical well-being or appearance), Interpersonal (harm to relationships with others), and Intrapersonal (disappointment in self) (Tonigan & Miller, 2002). This instrument has shown good test-retest reliability, internal validity, and sensitivity to change (Gillaspy & Campbell, 2006; Tonigan & Miller, 2002), even in a population with co-occurring substance use and mental illness (Bennett, Nidecker, Strong Kinnaman, Li, & Bellack, 2009). Derived from a similar instrument specific to alcohol, the Drinker Inventory of Consequences (DrInC) (Miller, Tonigan, & Longabaugh, 1995), the InDUC has been validated in treatment-seeking populations. Increased frequency of substance use is not always indicative of increased consequences of substance use behavior (Tiffany et al., 2012). The InDUC is sensitive to changes in frequency of use. For example, over a 90-day period a reduction of 40% in the frequency of substance use led to a 33% decline in substance-related consequences as measured by the InDUC (Tonigan & Miller, 2002). While lower QoL may be explainable at least in part by the consequences associated with substance use, it is unknown if QoL is affected differently by different kinds of substances. Traditional scoring of the InDUC asks users to think about consequences that resulted from any type of substance or alcohol use. However, it is possible that some substances may be associated with more severe negative consequences. For instance, even a single day of opiate use might be associated with severe consequences, while a single day of marijuana use or alcohol use might not. Further, illicit substance use is most often quantified as days of use, while alcohol consumption is quantified as standard servings per day (1.5 oz of liquor, 5 oz of wine, or 12 oz of regular beer) or number of heavy drinking days (>=5 drinks per day for men; >=4 drinks per day for women) (National Institute on Alcohol Abuse and Alcoholism, n.d.). Other illicit substance use behavior is more difficult to quantify outside of days of use as there are no standards of dosage. Even given similar doses, different methods of consumption, like smoking, snorting, ingesting, or injecting may lead to different levels of intoxication depending on how a substance was consumed.

Use of more than a single type of substance is common as using a single substance is a risk factor for developing dependence to additional substances (Jones, Logan, Gladden, & Bohm, 2015). In addition to increased risk of dependence, consequences of substance use are more severe among poly-substance users. It has been established in the literature that SUD persistence rates are highest among poly-substance users (Evans, Grella, Washington, & Upchurch, 2017; Moss, Chen, & Yi, 2014; Nosyk et al., 2014). Poly-substance users are more vulnerable to risks associated with substance use, especially

when mixing substances during the same period of use. Poly-substance use is a risk factor for both overdose and death (Jones et al., 2015). In the United States, most overdose deaths involve use of multiple substances (Jones, Mack, & Paulozzi, 2013; Warner, Chen, & Makuc, 2009).

Although it has been established that substance use behavior is correlated with lower perceived QOL via comorbid health conditions and psychosocial outcomes (De Maeyer, Vanderplasschen, & Broekaert, 2010; Havassy & Arns, 1998; Tiffany, Friedman, Greenfield, Hasin, & Jackson, 2012), QoL remains understudied as an endpoint in the substance abuse field (Laudet et al., 2009; Muller, Skurtveit, & Clausen, 2016; Smith & Larson, 2003). This is likely due to a lack of consensus for adoption of a standardized measure of QoL, underdeveloped arguments regarding the importance of including of QoL as a measure in substance use research, and a lack of a universal definition for QoL (Tiffany et al., 2012). Unfortunately, this often results in QoL measurement included on a study-by-study basis. When QoL is used, it is often used as a measure among users who have sought out treatment (De Maeyer et al., 2010).

The importance of utilizing measures beyond the scope of substance use behavior has been previously established in the literature (Maisto & Cooper, 1980; Pattison, Sobell, & Sobell, 1977; Tiffany et al., 2012). The Addiction Severity Index (ASI) (McLellan, Luborsky, Woody, & O'Brien, 1980) was created in response to these recommendations and has since been widely incorporated into treatment assessments. In additional international studies among treatment seekers, substance use severity as measured by the Addiction Severity Index (ASI) demonstrated that illicit substance users with more severe ASI scores had poorer QoL scores (Campêlo et al., 2017; Marini, Schnornberger, Brandalise, Bergozza, & Heldt, 2013). Despite recommendations to incorporate important concepts related to QoL into substance abuse research, this does not yet occur universally (Tiffany et al., 2012).

The present study aimed to determine if the frequency of different types of substance use were associated with QoL scores among substance users enrolled in a health coaching program. We defined use as the number of days of self-reported illicit substance use in the last 90 days, and alcohol use as the number of heavy drinking days in the last 90 days. Frequency of use for different drug types was used to predict QoL, while controlling for substance use consequences. While other studies have examined severity of use via instruments like the ASI, to our knowledge our study is the first to examine QoL separately by substance group type, while controlling for consequences of use.

2. Methods

This study used data collected by the 1115 Medicaid Waiver Program: Mobile Community Health Assistance for Tenants (m.chat). The Mobile Community Health Assistance for Tenants (m.chat) project recruited patients from November, 2014 to December, 2017. Funded by the Medicaid 1115 Waiver program in Texas, this program combined in-person health coaching and technology to address six different health domains: substance use, diet, exercise, medication compliance, social support, and recreation/leisure (Walters, Spence-Almaguer, Hill & Abraham, 2015). The m.chat program eligibility requirements included permanent supportive housing (PSH) residency (November 2014 – October 2016) or any subsidized housing residency (November 2016 – December 2017). Participants were required to be at least 18 years old and either Medicaid enrolled or low-income and uninsured at the time of screening. Participants were recruited from Tarrant County from housing lists provided by housing agencies, via case manager referrals, and word of mouth referrals. In addition, participants had to self-report at least one of the following mental health problems: prescribed medication for a psychological or emotional problem, experienced hallucinations in the past year, currently receiving a pension for a psychiatric disability, or scoring ≥9 on the 9-item Patient Health Questionnaire (PHQ-9) depression screener (Kroenke et al., 2001). Exclusion criteria included (1) residing in other types of housing not considered PSH (e.g., Transitional Housing or homeless shelter) before November 2016, (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/or (4) limited autonomy or decision-making capabilities (e.g., substantially neurologically or cognitively impaired). Participants were given assurances of confidentiality. Informed consent was obtained from each study participant. The project was approved by the Institutional Review Board (IRB) at the University of North Texas Health Science Center (UNTHSC).

Structured psychosocial assessments were completed at baseline, 6-month follow-up, and 12-month follow-up. 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 Timeline Follow-Back (TLFB) method (L. Sobell & Sobell, 1992). Substance users were determined at each assessment (baseline, 6-month, and 12-month follow-up via self-report as a user of any of the following substance use categories: (1) alcohol binge drinking (≥5 standard drinks for males or ≥4 for females), (2) any use of marijuana, (3) or any use of an "Other" drug, including cocaine, amphetamine, methamphetamine, opiates, prescription pills, or phencyclidine. Opiates were combined into the Other

substance category due to low prevalence of self-reported use. Additionally, poly-substance use days were recorded when one or more substance category was reported to be used within the same day.

Data were collected from 1,507 assessments from adults (*n* = 643 at baseline; 43.9% male; 32.7% White, 56.9% Black; 90.2% age 36+) residing in subsidized housing who were participating in m.chat. The sample is unique because it included participants with mental health conditions, often an exclusion criterion in other substance use studies. Additionally, the m.chat program was not a substance use treatment study, nor was substance use required for program eligibility. Substance use was captured via a 90-day self-reported TLFB, which has been shown to be a valid and reliable measure of substance use. Additionally, in a previous study we found good levels of agreement between self-report and toxicological confirmation of recent use of marijuana, cocaine, amphetamine, methamphetamine, and phencyclidine (Rendon et al., 2017). Frequency was captured as a proportion of substance use within the last 90 days prior to the interview. Frequency was treated as a continuous discrete variable taking values between 0 and 1 for analysis. For analysis of consumption levels, we classified participants into four categories: those who reported using 0 days (none) (frequency = 0), 1-30 days (low consumption) (frequency = 0.01 – 0.33), 31-60 days (moderate consumption) (frequency = 0.34 – 0.67), and 61-90 days (heavy consumption) (frequency = 0.68 – 1.00). Results were still presented via these frequency subgroups. Due to low counts in some of these subgroups, the results should be interpreted with caution. Aggregate data for sub-groups is displayed in Table 2, but due to the small n, could have been skewed by a single individual.

To measure substance problems, we used a modified 15-item version of the InDUC, which has been shown to have good internal consistency and to be highly correlated with the longer version (Bennett et al., 2009; Kiluk, Dreifuss, Weiss, Morgenstern, & Carroll, 2013) . The primary outcome in this analysis was quality of life measured by the Quality of Life Satisfaction and Enjoyment Questionnaire Short Form (Q-LES-Q-SF). The reliability, validity, and factor structure of the Q-LES-Q has been established in a variety of populations, including community samples and samples with mental health problems (Riendeau et al., 2018). Raw scores range from 14-70, with higher scores indicating a higher quality of life satisfaction. Sum scores were converted to percentage maximum possible score for analyses with scores ranging from 0% - 100%. Percentage maximum possible scores were adjusted to account for missing on purpose values. The sum of the raw InDUC score was used as a covariate for analyses. Other covariates included age, gender, and race. We used a hierarchical two-level model to fit the data predicting QOL in SAS version 9.4 (SAS Institute Inc, Cary, NC). The first level was repeated time

(baseline, 6 months, and 12 months) which was nested within the second level of person-level data for the conditional model. We relied on the second model because it allowed us to capture both within subject variation and between subject variability in predicting QoL.

3. Results

This sample was primarily composed of adults aged 36 and older (90.2%) with a median age of 52.2 years at baseline (Table 1). Self-reported users of any use of illicit substances or alcohol binge consumption made up 39.64%, 44.92%, and 35.27% of the sample at baseline, 6-month follow-up, and 12-month follow-up, respectively. Among participants with any use reported in the three categories of interest (binge drinking, marijuana, or Other substance category via cocaine, amphetamine, methamphetamine, opiates, prescription pills, or phencyclidine) binge drinking and marijuana occurred at the highest proportion of days across all three time points (Table 2). For all substances, most users reported using less than 30 out of the last 90 days (frequency = 0.01 – 0.33) (Table 2). Binge drinking frequency remained relatively stable across time, overall slightly decreasing from 21.77% of participants at baseline to 17.52% at 12 month follow-up (Table 2). Marijuana was the most commonly reported substance used at baseline (27.37%), 6 months (25.59%), and 12 months (26.89%). In addition to being the most commonly reported used substance, marijuana also had the highest proportion of users who used at least 60 days of the past 90 days (frequency = 0.67-1.00) at baseline (4.82%), 6 months (4.30%), and 12 months (5.14%) compared to other substances. Substances in the combined Other category had lower rates of use overall compared to binge drinking and marijuana. Both Other substance category use and poly-substance use decreased over time (Table 2).

Table 4. Participant Demographics

Demographics		Baseline *N* = 643		Follow-up 1 *N* = 512		Follow-up 2 *N* = 331	
		n	%	*n*	%	*n*	%
Age	**Young Adult (18 - 35 years)**	63	9.80%	42	8.20%	19	5.74%
	Middle Adult (36 - 55 years)	364	56.61%	277	54.10%	182	54.98%
	Older Adult (56+ years)	215	33.44%	193	37.70%	130	39.27%
Sex	**Male**	282	43.86%	222	43.36%	141	42.60%
	Female	361	56.14%	290	56.64%	190	57.40%
Race	**Black/African American**	366	56.92%	284	55.47%	178	53.78%
	White	210	32.66%	174	33.84%	127	38.37%
	Hispanic	37	5.75%	33	6.45%	15	4.53%
	Other/Multi-racial	30	4.67%	31	6.05%	11	3.32%

Table 5. Self-Reported Substance Use Frequency within the Last 90 Days

		Baseline					
	Frequency	**n**	**%**	**lnDUC**	**Std. Dev.**	**QOL**	**Std. Dev.**
Binge Drinking	**0**	503	78.23%	3.34	7.79	55.71%	0.16
	0.01 - 0.33	95	14.77%	9.93	10.80	53.31%	0.15
	0.34 - 0.67	26	4.04%	16.54	13.89	52.47%	0.16
	0.68 - 1.00	19	2.95%	17.00	13.85	49.99%	0.16
Marijuana	**0**	467	72.63%	3.60	8.23	55.55%	0.16
	0.01 - 0.33	117	18.20%	9.09	11.06	52.81%	0.16
	0.34 - 0.67	28	4.35%	11.93	12.83	55.36%	0.16
	0.68 - 1.00	31	4.82%	9.61	12.24	55.77%	0.14
Other Substance Category	**0**	483	75.12%	2.69	7.04	55.79%	0.16
	0.01 - 0.33	137	21.31%	11.50	11.34	54.07%	0.15
	0.34 - 0.67	13	2.02%	21.77	14.24	44.09%	0.16
	0.68 - 1.00	9	1.40%	24.00	7.60	45.63%	0.21
Poly-Substance Use	**0**	534	83.05%	3.54	7.84	55.78%	0.16
	0.01 - 0.33	87	13.53%	11.61	11.75	51.81%	0.15
	0.34 - 0.67	10	1.56%	28.90	13.84	42.86%	0.12
	0.68 - 1.00	11	1.71%	15.73	10.93	56.82%	0.13
		6 Months					
	Frequency	**n**	**%**	**lnDUC**	**Std. Dev.**	**QOL**	**Std. Dev.**
Binge Drinking	**0**	413	80.66%	2.75	6.62	60.57%	0.16
	0.01 - 0.33	78	15.23%	8.23	9.64	58.75%	0.15
	0.34 - 0.67	11	2.15%	7.18	8.33	64.94%	0.13

	0.68 - 1.00	10	1.95%	18.90	13.96	54.11%	0.14
Marijuana	**0**	381	74.41%	3.06	7.37	61.03%	0.16
	0.01 - 0.33	97	18.95%	6.59	8.59	58.08%	0.16
	0.34 - 0.67	12	2.34%	7.75	8.99	59.52%	0.12
	0.68 - 1.00	22	4.30%	7.27	9.80	56.98%	0.13
Other Substance Category	**0**	412	80.47%	2.30	5.79	61.10%	0.16
	0.01 - 0.33	88	17.19%	9.72	9.86	57.50%	0.14
	0.34 - 0.67	6	1.17%	20.00	12.38	57.74%	0.18
	0.68 - 1.00	5	0.98%	23.40	17.31	43.93%	0.07
Poly-Substance Use	**0**	460	89.84%	3.00	6.87	60.55%	0.16
	0.01 - 0.33	46	8.98%	12.37	9.90	58.09%	0.15
	0.34 - 0.67	4	0.78%	12.25	10.72	60.27%	0.14
	0.68 - 1.00	2	0.39%	30.5	14.85	43.75%	0.09

12 Months

	Frequency	**n**	**%**	**InDUC**	**Std. Dev.**	**QOL**	**Std. Dev.**
Binge Drinking	**0**	273	82.48%	2.69	6.62	60.51%	0.16
	0.01 - 0.33	40	12.08%	6.40	7.46	62.32%	0.15
	0.34 - 0.67	8	2.42%	22.50	12.28	49.78%	0.16
	0.68 - 1.00	10	3.02%	15.30	11.03	54.11%	0.18
Marijuana	**0**	242	73.11%	2.71	6.95	61.98%	0.16
	0.01 - 0.33	61	18.43%	7.43	9.74	57.13%	0.15
	0.34 - 0.67	11	3.32%	7.00	9.06	48.21%	0.17
	0.68 - 1.00	17	5.14%	8.12	9.05	55.15%	0.19
Other Substance Category	**0**	272	82.18%	2.47	6.41	61.26%	0.16
	0.01 - 0.33	53	16.01%	10.72	10.42	56.60%	0.16
	0.34 - 0.67	4	1.21%	15.25	10.69	46.43%	0.20
	0.68 - 1.00	2	0.60%	11.00	14.14	50.90%	0.16
Poly-Substance Use	**0**	299	90.33%	3.05	7.15	60.89%	0.16
	0.01 - 0.33	25	7.55%	12.80	10.01	55.50%	0.19
	0.34 - 0.67	4	1.21%	10.25	6.99	54.91%	0.13
	0.68 - 1.00	3	0.91%	16.33	13.81	46.43%	0.07

Among users, days of use were positively associated with InDUC scores in every category (Table 2). The increase in InDUC score was less pronounced for marijuana days, compared to the increase in InDUC scores for increasing frequency of binge alcohol use, and consumption of substances in the Other category and poly-substance use. While a difference in InDUC scores between substance types was noted, we did not identify a differential pattern of QoL scores by substance use type. Overall, substance users had lower QoL scores compared to non-users (Table 2). Across time, the average QoL score was

55.06% (std 0.62) for the sample and increased over time. This metric represents the percentage maximum possible score on the Q-LES-Q-SF instrument, meaning that the average QoL score for the same was 55.06% of the maximum score. At baseline, average QoL scores for the sample ranged from 42.86% to 55.79%, with frequent binge drinkers, users of Other category substances, and poly-substance users representing the lowest overall QoL scores (Table 2). At 6 months, average QoL scores ranged from 43.75% to 61.10%, with frequent users of Other category substances and poly-substance users showing the lowest QoL scores (Table 2). At 12 months, average QoL scores ranged from 46.43% to 62.32%, with moderate frequency use of substances from the Other category and heavy poly-substance use representing the lowest QoL scores.

We found that InDUC score, use of substances in the Other category (including cocaine, amphetamines, methamphetamines, opiates, prescription pills, and phencyclidine), and race were statistically significant predictors of QoL while controlling for age and gender (Table 3). Poly-substance users had similar QoL scores compared to single substance users (Table 2) across time and frequency of consumption.

Table 6. Predicting QoL Among Users and Non-Users Using Hierarchical Linear Modeling

Unconditional Model		
Effect	**F Value**	**Pr > F**
Time	**3.41**	**0.07**
Time*Time	**18.83**	**<0.01**
Conditional Model		
Effect	**F Value**	**Pr > F**
Time	3.18	0.07
Time*Time	16.97	<0.01
InDUC	22.03	<0.01
Binge Drinking	0.41	0.52
Marijuana	0.05	0.83
Other Substance Category	3.92	0.05
Poly-Substance Use	1.52	0.22
Age	1.5	0.22
Race	3.85	0.01
Gender	0.08	0.78

Table 7. Model Covariance Parameter Estimates

Unconditional Model Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error
Intercept	**CN**	0.01659	0.001166
Residual		0.008325	0.000404

Conditional Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error
Intercept	**CN**	31.6930	2.5766
Residual		23.5868	1.1758

4. Discussion

This study examined QoL among non-users and people who reported different types of substance use at baseline, 6-months, and 12-months among a sample of adults residing in subsidized housing. This study is unique because the program was not a substance use intervention and substance use was not required for program eligibility.

Our findings that overall QoL scores were higher among non-users fit within existing literature (De Maeyer, Vanderplasschen, & Broekaert, 2010; Havassy & Arns, 1998; Tiffany, Friedman, Greenfield, Hasin, & Jackson, 2012). We found that InDUC scores, race, and days of use of substances in the Other category (cocaine, amphetamine, methamphetamine, opiates, prescription pills, or phencyclidine) were significant predictors of QoL. These findings support our hypothesis that quality of life among substance users is closely related to substance use problems measured by the five InDUC concepts: impulse control, social responsibility, and physical, interpersonal, and intrapersonal consequences. These findings align with previous literature regarding a similar instrument, the ASI, where higher scores predicted lower QoL (Campêlo et al., 2017; Marini et al., 2013). ASI's usefulness in measuring substance use severity in populations with co-occurring substance use and mental illness has been well documented (Hodgins & el-Guebaly, 1992). The present study adds information about the shorter InDUC instrument, which has been more recently introduced into populations with co-occurring substance use and mental health problems (Bennett et al., 2009). We found similar results to other studies that used the ASI to measure substance use problems using the InDUC instrument. In the current study, higher InDUC scores indicative of more substance related problems correlating with lower QoL.

In this study, we found that substance use behavior varied over time. A higher proportion of the sample self-reported substance use at the 6-month follow-up than at baseline or 12-month follow-up. The reasons for this variation is unknown for this study, but may be due to loss to follow-up, change in reporting behavior, or change in substance use. We did not find a statistically significant difference in loss to follow-up between users and non-users. It is possible that substance users may have felt more comfortable self-reporting substance use after completing some time in the program. Fear of negative consequences of reporting illicit behaviors may have caused some participants to misrepresent substance use at baseline. It is also possible that some participants increased their substance use behavior at 6 months and stopped using at 12 months. Studies that recruit and follow-up at set intervals may be susceptible to season or calendar effects on substance use behavior due to variation in activities during different seasons. We did not find seasonal effects of substance use patterns as enrollment and follow-ups occurred continuously for a 30 month period. Among participants who self-reported binge alcohol, marijuana, or Other illicit substances, we found that marijuana and binge alcohol had the highest frequency of reported use within the past 90 days across all time points.

We found that InDUC scores generally increased as frequency of use increased across all substance classes. This finding regarding alcohol binge consumption supports previous findings where a higher frequency of binge alcohol consumption was associated with lower QoL (Donovan et al., 2005; Volk et al., 1997). In addition, InDUC scores tended to increase as substance use days increased, but at a less severe rate for marijuana compared to binge alcohol and the Other drug category. For example, the InDUC score for marijuana users at baseline ranged from an average of 9.09 for low consumption users (frequency = 0.01 – 0.33) to 11.93 for moderate consumption users (frequency = 0.34 – 0.67). At baseline, the change in average InDUC score for the same frequency of use for alcohol binge drinking was 9.93 to 16.54, Other substances were 11.50 to 21.77, and poly-substance use was 11.61 to 28.90 (Table 2). Because opiates were included in this Other substance category, we are unable to make conclusions about the effects of the frequency of opiates separate from other substances in this category. Combining multiple substances into a single Other substance category was necessary for analyses due to low frequency of these behaviors. Ideally, cocaine and opiates would be considered for analysis as their own categories because different drug classes undoubtedly result in different kinds of problems. By lumping substances into a combined category, we are unable to differentiate problems between the substances types included in this category.

Our findings suggest that the kinds of problems measured by the InDUC may not be affected by increasing frequency of marijuana use. For example, the interpersonal domain may not be affected by

the frequency of marijuana use. Marijuana users who consume at higher rates may not be as affected by feeling disappointed in themselves for increased frequency compared to other substances. Social stigma may also play into severity of self-assessed substance use problems. Marijuana may be considered a softer drug compared to other illicit substances. This reduced social stigma may relax individual's perception of risks and consequences associated with marijuana consumption compared to other more stigmatized substances like cocaine or opiates. This may explain, at least in part, why InDUC scores increased at lower rates compared to other substances as frequency of consumption increased. Previous studies have suggested the existence of a threshold effect of consuming marijuana more than once per week (frequency > 0.14) on decreased QoL scores (Dey et al., 2013, 2014). While our findings demonstrate an increase in InDUC score between users and non-users at baseline, 6 months, and 12 months, the highest InDUC scores among marijuana users were not always among the most frequent users. For example, at baseline the highest average InDUC score was among the moderate frequency group (frequency = 0.34 - 0.67) at 11.93 (n = 28). This was higher than the mean score for the high frequency consumption group (>0.67) at 9.61. A similar pattern was demonstrated at 6 months and 12 months. We did not find the difference in range between low, moderate, and heavy consumption for marijuana as we did for other substances. Any marijuana use (low, moderate, and heavy consumption) had similar InDUC scores at baseline (9.09 – 11.93), 6 months (6.59 - 7.75), and 12 months (7.00 – 8.12) when compared to no marijuana use at baseline (3.60), 6 months, (3.06), and 12 months (2.71). Our findings align with the threshold affect proposed by Dey et al. (2013, 2014) in which at least low frequency of use results in a difference between users and non-users of marijuana. In our study we identified this difference at InDUC assessment of substance use problems. While controlling for these problems, we did not find a significant effect of frequency of marijuana consumption on QoL as did Dey et al. (2013, 2014).

The present study determined that the frequency of substance use of Other category substances including cocaine, amphetamine, methamphetamine, opiates, prescription pills, or phencyclidine was significantly associated with QoL scores among substance users. This study was unique in its classification of substance use frequency, using the number of days of self-reported substance use within the last 90 days. While other studies have examined severity of use via instruments like the ASI, to our knowledge our study is the first to examine QoL by self-reported frequency of use by substance group type while controlling for measured consequences of use via InDUC.

4.1 Limitations

Overall, the sample had low QoL scores. Thus, the results may not generalize to other populations. Frequency of use was categorized by self-report only. We have extensively studied the misreporting rate in this population. Actual substance use frequency may be higher than what was captured via self-report. In addition, due to low rates of opiate use we were unable to analyze opiate use frequency as its own category. Opiates were included in the analysis but were categorized as part of the Other substance category, which also included cocaine, amphetamine, methamphetamine, prescription pills, and phencyclidine.

Investigation of subgroups of substance use frequency led to small representation in some of the subgroups (Table 2). Thus, we recommend using caution when interpreting the InDUC and quality of life scores for these small subgroups.

5. Conclusion

Among people with mental health conditions residing in subsidized housing population, QoL was significantly predicted by InDUC score, race, and frequency of use of Other (cocaine, amphetamine, methamphetamine, opiates, prescription pills, or phencyclidine) substances. Substance use problems increased as frequency of consumption increased for alcohol binge drinking and drugs in the Other category. InDUC scores were higher for users of any frequency of marijuana compared to non-users, but InDUC scores were less severe compared to substances from Other categories. Future studies should consider examining both frequency and severity (e.g. dosage and consumption method) of use for different substance types to determine if a threshold effect exists on InDUC and QoL scores. In addition, the different types of problems should be studied further to determine if any one of the five domains of the InDUC instrument contribute more to subjective quality of life. Determination of a substance use threshold could be clinically relevant to settings in which an immediate goal is reducing use below a certain level as a way to avoid substance use consequences and improve quality of life before addressing the ultimate treatment goal of abstinence.

Chapter 5: Discussion

5.1. Purpose of the Study

The purpose of this three-paper model dissertation was to examine three topics related to substance use behavior in a study of subsidized housing residents enrolled in the m.chat health coaching program. This program was unique as it provided health coaching on patient-selected goals and action items within six domains: substance use, medication adherence, diet, exercise, social support, and recreation and leisure. While this program offered resources for participants interested in reducing or eliminating substance use, substance use (or sobriety) was not a criterion for participation. All three papers involved factors associated with people's self-reported substance use. The first area of research was focused on the validity of self-reported illicit substance use at baseline. Manuscript 1 compared self-reported illicit substance use of three categories of psychoactive, illicit substances (marijuana, cocaine, and a combined category of amphetamine, methamphetamine, and phencyclidine) to determine the validity of the Timeline Follow-Back (TLFB) method using saliva toxicological test as a gold standard. Manuscript 2 focused on the longitudinal trends of agreement between TLFB and saliva toxicological test in the sample across baseline, 6 month, and 12 month follow-up. Similar to Manuscript 2, Manuscript 3 used baseline and follow-up data from the m.chat program. Manuscript 3 focused on predicting subjective quality of life (QoL) measured via the Quality of Life Enjoyment and Satisfaction Questionnaire Short Form (Q-LES-Q-SF) among self-reported users of marijuana, binge alcohol, and a combined Other category of cocaine, amphetamine, methamphetamine, opiates, prescription pills, or phencyclidine. By framing these topics as three separate research questions, this dissertation provided information on ways to collect and predict substance use, and identified correlates of substance use in this high-risk population.

This final chapter discusses the major findings of the three manuscripts. We conclude with a brief summary, discussion of study limitations, and areas for future research. Via three manuscripts, we accomplished the following aims in addressing our following research questions:

Aim 1.

Determine the rates of concordance among participants who use psychoactive substances at baseline and demographic differences between concordant and discordant users. Specifically: Do rates of concordance differ by substance type group (marijuana, cocaine, or amphetamine, methamphetamine, and phencyclidine) among residents of subsidized housing programs? If so, should self-report be continued to be used as an exclusive measure of substance use behavior in this

population because of its relative ease of collection? Do demographic variables or other self-reported behaviors predict discordance at baseline? If so, how can this information be used to advise other researchers or interventionists?

Aim 2.

Examine if the rate of agreement between self-reported substance use and nonuse, and saliva toxicological test differ across time (from baseline to 6 months, and 12 months) for three substance groups: marijuana, cocaine, and combined amphetamine, methamphetamine, and phencyclidine. Determine if self-report is a reliable measure of substance use compared to saliva toxicological test in this sample. After determining how the rate of agreement between self-report and saliva toxicological confirmation changes longitudinally, generate hypotheses regarding why the agreement rate changed over time.

Aim 3.

Examine if frequency of self-reported days of substance use, including binge drinking, marijuana, Other substances (including cocaine, amphetamine, methamphetamine, opiates, prescription pills, or phencyclidine), and poly-substance use within the same day predict QoL while controlling for substance use problems. Determine if substance use problems measured via the InDUC predict QoL in the sample. Frequency was defined as the number of days of self-reported substance use within the last 90 days.

5.2. Key Findings

While accomplishing our aims for each manuscript, we had a number of key findings. In Manuscript 1, we discovered that relying exclusively on self-report at baseline to capture substance use captured less than half of actual users, depending on substance type. This finding that self-report tends to underestimate substance use behavior was consistent with previous studies of other vulnerable populations (Aiemagno et al., 1996; Digiusto et al., 1996; Hjorthøj et al., 2012; Schumacher et al., 1995) and is of concern in a field where substance use behavior is almost always collected exclusively via self-report (Kirst et al., 2015; Padgett, 2006). Additionally, we found that age, race, medical insurance status, and alcohol use were significant predictors of discordance between self-reported substance use and saliva toxicological confirmation. The odds of misreporting were greater among older participants, non-White participants, those without any medical insurance, and those who did not report drinking any alcohol in the past 90 days. We discuss some of these findings further in areas for future study.

In Manuscript 1, our results were consistent with a meta-analysis that determined the validity of TLFB (Hjorthøj et al., 2012), except that we were able to better detect non-users based on our higher

levels of specificity. We also found an increase in the positive predictive value (PPV) across substance use types when we increased the window for recall for comparison to the saliva toxicological confirmation. The toxicological test captures use within the past 72 hours for amphetamine, methamphetamine, phencyclidine, and cocaine and use within the past 30 days for marijuana (Redwood Toxicology Laboratory, 2015). For example, the concordance for cocaine was 43.8% at baseline when self-report via TLFB was limited to recall within the past 72 hours. When we allowed an increased window of recall of 90 days, the concordance for recent cocaine use increased to 60.4%. This suggests a greater willingness among participants to acknowledge substance use that occurring further back in time. This pattern of increased concordance between toxicological confirmatory tests and self-reported use that is further back in time has also been demonstrated in other populations (Harrison, 1997; Hjorthøj et al., 2012). Interestingly, we also found that a significant number of marijuana users self-reported use, but did not test positive on the saliva toxicological confirmation. We speculate that this may be attributable to either a lower rate of sensitivity in the saliva toxicological test or a bias in the sample towards over-reporting of marijuana use.

The effects of time on agreement between self-reported substance use and saliva toxicological confirmation was previously unknown. In Manuscript 2, we found that agreement between self-reported substance use via TLFB and saliva toxicological confirmation was generally high across time. Agreement for marijuana (85% - 93%) fell within agreement reported in other studies (86% - 98%) (Babor et al., 2004; Godley et al., 2002; Stasiewicz et al., 2008) and improved over time. Agreement for cocaine (85% - 91%) was higher compared to other studies (70% - 87%) (Carroll et al., 2004a; Elkashef et al., 2005; McDowell et al., 2005; Pettinati et al., 2008a; Pettinati et al., 2008; Somoza et al., 2008b; Stasiewicz et al., 2008) and decreased over time. For the combined amphetamine, methamphetamine, and phencyclidine group, agreement (84% - 94%) over time was inconclusive, likely due to the few cases in critical cells.

In Manuscript 3, we found that among any participants that self-reported using binge alcohol, marijuana, or the Other category of illicit substances (cocaine, amphetamine, methamphetamine, opiates, prescription pills, or phencyclidine, marijuana and binge alcohol had the highest frequency of reported use within the past 90 days at baseline, 6 months, and 12 months. This study uniquely analyzed substance use frequency as a proportion of days used within the last 90 days. Our finding that overall QoL scores were higher among non-users compared to substance users fits within existing literature (De Maeyer, Vanderplasschen, & Broekaert, 2010; Havassy & Arns, 1998; Tiffany, Friedman, Greenfield, Hasin, & Jackson, 2012). In addition, InDUC scores, race, and days of use of Other categorized

substances were significant predictors of QoL. Previous studies using a similar tool, the Addiction Severity Index (ASI), found that more substance-related problems was predictive of lower QoL (Campêlo et al., 2017; Marini et al., 2013). Our study adds to the literature about the shorter InDUC instrument. To our knowledge our study is the first to examine QoL by frequency of self-reported substance use while controlling for measured consequences of use via InDUC. We found that InDUC scores generally increased as frequency of use increased for all substances, but at a lower rate for marijuana. This suggests that greater days of marijuana use are associated with fewer problems, compared to binge drinking and the Other substance use category.

5.3. Practice Implications

These results have a number of implications that may be useful in research or clinical settings among similar populations where substance use information is collected. First, self-reported substance use may be more useful as a proxy of recent use than current use when toxicological testing is not possible. We found that taking into account individuals' self-reported historical substance use increased the agreement between self-report and toxicological confirmation for marijuana, cocaine, amphetamine, methamphetamine, and phencyclidine. Our findings are supported by previous literature in which increased concordance was noted between self-reported substance use reported further back in time (Harrison, 1997; Hjorthøj et al., 2012). We suggest that future researchers using the TLFB to collect data on self-reported substance use should consider using the full 90-day TLFB window as a proxy measure of recent use of these substances. Widening the self-report recall windows may be a better proxy of self-reported substance use in similarly vulnerable populations with a high prevalence of mental illness.

Due to the possibly of low sensitivity for marijuana use on the saliva toxicological test, studies that exclusively rely on a saliva toxicological test may underestimate true rates of use of marijuana. Based on our findings, we recommend using self-report instead of (or in addition to) the saliva toxicological test to capture marijuana use.

Our findings suggest that saliva toxicological tests withstand community field assessments in which samples were likely exposed to higher or lower than ideal temperatures during transport back to the office and shipment to the laboratory leading to possible sample degradation. Saliva toxicological tests were administered by a research assistant during the psychosocial interviews at baseline, 6 months, and 12 months. Research participants held the swab under their tongue to collect adequate saliva sample. A blue color change indicated that enough sample had been collected. The participant

placed the swab into the bluffer-filled tube and closed the top of the sample. The research assistant placed a tamper evident seal on the tube and labeled the tube with the participant's study identification number. The sealed tube stayed in possession of the research assistant until the interview was completed. If conducted offsite, the tube was transported back to the research office. Samples were held for 1-2 weeks until an adequate number of samples had been received to warrant shipping costs to the laboratory. Validation of the collection instrument and the testing had been conducted previously in vitro with set storage temperature at 12 hours before analysis (Quintela & Crouch, 2006). It was important to determine if this device and method of toxicological testing in a real-life field setting was valid. Without the validity of non-clinic based study sample collection, the application of this device and method would be limited. We found generally high levels of agreement between the saliva toxicological test and self-reported substance use. This suggests that saliva toxicological testing withstands field assessments and is a viable option in community field settings when capturing substance use via toxicological testing is warranted. We found that greater substance use problems measured via the InDUC were significantly related to lower QoL among substance users. This finding fits with current literature which suggests that consequences related to substance use behavior cause substantial distress to the user and thus societal concern (Tiffany et al., 2012). These consequences may directly or indirectly impact the user's ability to carry out normal daily functions and impede their life enjoyment. Lower perceived QoL may be influenced by the perceived severity of consequences associated with substance use either directly, like criminal justice interactions or indirectly, like comorbid health problems. Determination of a substance use threshold could be clinically relevant to settings in which a harm-reduction approach is adopted with the immediate goal of reducing substance use below a certain level as a way to avoid substance use consequences and improve QoL (Tiffany et al., 2012).

5.4. Limitations

There were a number of limitations for these three manuscripts. Of note, the sample changed from Manuscript 1 (n = 334) to Manuscript 2 and 3 (n = 643) because the study was still ongoing when Manuscript 1 appeared. Manuscript 1 was published before the parent program completion. As a result, analyses for Manuscript 1 were conducted on a subset of participants used for Manuscripts 2 and 3.

Both Manuscript 1 and 2 were affected by the sensitivity of the saliva toxicological test. We expect the effect of this limitation to be minimal on our misreporting estimates as saliva toxicological testing has been accepted as a sensitive measure of recent cocaine, amphetamine, methamphetamine, and marijuana use (Allen, 2011; Dams, Choo, Lambert, Jones, & Huestis, 2007). We found this method to

be advantageous due to low potential of adulteration, minimal possibility of environmental contamination, low analytical costs, and low refusal rates (Gjerde, Øiestad, & Christophersen, 2011) compared to more invasive methods of toxicological testing via urine or blood sampling.

All of the manuscripts were affected by recall bias related to self-reporting behavior. The studies were unable to determine if misreporting was intentional or if some users were unable to accurately recall their substance use behavior. 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. We discuss this concept further in areas for future research.

We were limited in our ability to test the agreement between self-report and saliva toxicological confirmation for alcohol because the parent program did not use saliva toxicological test for alcohol confirmation. Therefore, Manuscripts 1 and 2 were unable to compare self-reported alcohol consumption to a saliva toxicological confirmatory test or make any conclusions about the agreement for alcohol consumption because the toxicological test used by the parent program did not test for this. We felt it was important to include binge alcohol use in our analysis for Manuscript 3. Due to this limitation, we exclusively used self-reported substance use in Manuscript 3 whereas Manuscripts 1 and 2 used both self-reported substance use and saliva toxicological confirmation.

Another limitation of these studies is that all patients were given assurances of confidentiality, which undoubtedly influenced the accuracy of their self-reporting. When such assurances are not given, or resources are contingent upon an individual's self-reported use, it seems reasonable that this would result in even higher rates of misreporting.

We experienced limitations due to sample size. In Manuscript 2, we were unable to examine individual-level factors due to the small N for positive-positive and discordant cases at 6 month and 12 month follow-ups. Instead, we looked at population level patterns of agreement. Similarly, we had few cases of amphetamine, methamphetamine, and phencyclidine use. We would need to examine larger samples to improve estimate precision for these substances throughout all time points. In Manuscript 3, the low frequency of use led us to combine opiate use into the combined Other category cocaine, amphetamine, methamphetamine, prescription pills, and phencyclidine. Ideally, our sample would have been large enough to consider opiate usage independently. As a combined substance category, we found that frequency of use within the past 90 days of this Other substance category was a significant predictor of QoL. Due to the combined nature of the category, we are unable to make conclusions about any of the substance types independently.

Finally, 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) as the parent program was a voluntary health coaching program. One of the endpoints of the parent program was to improve QoL among participants. This program was designed to help participants make health-related changes. Positive changes in health-related outcomes would have likely led to improved daily functioning and improved life enjoyment, the mechanism by which we believe QoL is affected by substance-related problems. Improvements in either functioning or life enjoyment would have been reflected as an increase in QoL over time, as was seen in Manuscript 3. Manuscript 1 findings are most directly applicable to PSH residents who use substances; however, they may also have implications in similar populations of low-income clients with mental health disorders. In Manuscript 3, the overall sample of both users and non-users had low QoL scores. This may be due to the inclusion criterion of a mental health problem defined as either 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 et al., 2001). The results of Manuscript 3 may not generalize to other substance-using populations without mental health problems.

5.5. Future research

Our findings from the three manuscripts generated additional questions that could be considered for future research. First, we found that individuals with medical insurance were less likely to misreport substance use. If these findings are replicable, future research should be conducted to determine the mechanism behind the association between a lack of medical insurance and substance use misreporting. For instance, it might be that individuals without insurance feared negative repercussions of reporting substance use. Some participants might have believed that admitting to substance use might somehow affect their eligibility for medical insurance or other need-based program assistance in the future. On the other hand, participants with insurance were more likely to report their substance use concordantly and did not seem to fear losing eligibility for need-based assistance. Similarly, individuals who reported alcohol consumption in the past 90 days were more likely to accurately self-report their substance use, compared to those who did not report alcohol consumption. It is possible that people who felt comfortable disclosing alcohol use also felt less stigmatized about reporting use of other illicit substances. Because this study did not include a biological test for alcohol, it is not possible to determine if participants who misreported substance use also misreported alcohol use.

Further research is needed to determine if actual alcohol use, or reported alcohol use is the better predictor of substance use reporting.

From Manuscript 2, longitudinal trends of amphetamine, methamphetamine, and phencyclidine need to be studied in a larger sample due to the relatively low frequency of these substance use behaviors. Additionally, further study is needed to establish why concordance changes over time. We found that agreement was generally high, but increased over time for marijuana and decreased for cocaine. In Manuscript 2, we were unable to test any mechanisms to explain why concordance changes over time. There are a number of reasons substance users may intentionally or unintentionally misreport their substance use. Some of the reasons for underreporting may be lack of standardized methodology, social desirability or stigma of reporting illicit substance use, inability to recall substance use further back in time, and fear of legal repercussions (Napper et al., 2010). We can speculate an explanation as to why concordance would change differently for different substances over time, but further testing is needed to explore the mechanism. We hypothesize that there may be two co-occurring mechanisms at play: recall bias and assessment reactivity.

In regard to recall bias, participants may feel more comfortable disclosing use of less stigmatized substances like marijuana over time as they become more comfortable in trusting a program. In contrast, cocaine agreement may have decreased due to perceived social desirability among those delivering the intervention (Napper et al., 2010). Participation in the parent program meant establishing a one-on-one relationship with a health coach. For participants who's self-report and saliva toxicological test agreed at baseline, some participants may have continued to use but felt like they should report no use to perceived social desirability constraints. They may have felt like reporting accurate use at follow-up may have let their health coach down or to seem like they were not making progress in the program. In this case, participants may minimize their substance use because of their relationship with their interventionist. Using the same logic, it is possible that social desirability may influence a substance user to exaggerate substance use frequency or intensity or even solicit misrepresentation about substance using behaviors from a non-user because they want to adhere to their perception of expected behavioral norms. Studied extensively in adolescents, social desirability has not been adequately studied in adult substance users with comorbid mental conditions. Mental health problems may have a substantial impact on unintentional recall bias by affecting an individual's ability to accurately recall past substance use.

Additionally, participants may make actual changes in substance use as a result of questions posed to them about their substance use behavior. This phenomenon is known as assessment reactivity

and has been well-established in the literature in regards to self-reported drinking behavior (McCambridge & Kypri, 2011; Moyer et al., 2002; LC Sobell & Sobell, 1981; Walters et al., 2009). Assessment reactivity may also affect the individual's ability to recall behavior, for instance serving as retrieval cues for behavior. Information unintentionally given during assessments may influence an individual's future estimation of that behavior. For instance, asking someone how many standard servings of alcohol they have consumed in the last 90 days while defining that a standard serving is 1.5 oz. of liquor, 5 oz. of wine, or 12 oz. of beer may inadvertently influence the individual's future consumption of alcohol. This increased or heightened awareness could lead to actual decrease in the amount of alcohol or frequency of illicit substances consumed. It could also make individuals better reporters of their own use.

In Manuscript 3, substance use problems were less severe for individuals who used more days of marijuana compared to the same number of days of binge drinking or Other substances. Future studies should consider examining frequency and severity of use (e.g. dosage and consumption method) of use for different substance types to determine if a threshold effect exists on InDUC and QoL scores. Additionally, it may be of interest to look at different types of problems. The InDUC measures problems in five scales: Impulse Control (harmful risk-taking), Social Responsibility (recognition of a lack of fulfillment of personal responsibilities like school or work), Physical (harm to physical well-being or appearance), Interpersonal (harm to relationships with others), and Intrapersonal (disappointment in self) (Tonigan & Miller, 2002). Problems in one scale may disproportionally affect an individual's subjective QoL for different types of substances. Lower perceived QoL may be influenced by consequences associated with substance use either directly or indirectly. For example, QoL may be indirectly affected by a cough that occurred as a result of smoking marijuana. This consequence may indirectly affect QoL if the cough impedes the individual's ability to complete normal daily functions. A consequence with a more direct effect on QoL could be incarceration due to substance use. This consequence would directly impact an individual's ability to carry out their normal daily functions and impede their life enjoyment. In addition, the different types of problems should be studied further to determine if any one of the five domains of the InDUC instrument contribute more to subjective quality of life.

5.6. Conclusion

While analyzing substance use behavior among a sample of subsidized housing residents, we made a number of significant findings in regards to how substance use data is collected. In Manuscript 1,

we found that exclusively relying on self-report method for substance use could result in an underestimation of true substance use. In contrast, exclusively using saliva toxicological testing for substances would have resulted in an underestimate of use for marijuana. We found that participants who misreported substance use were more likely to be older, non-White, have no medical insurance, and not report any alcohol use. When biochemical verification is not possible, increasing the recall window may help increase the accuracy of self-reported substance use. In Manuscript 2, we found that the 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 that the agreement of marijuana increased over time and decreased over time for cocaine. In Manuscript 3, we found that QoL was significantly predicted by InDUC score, race, and frequency of use of other (cocaine, amphetamine, methamphetamine, opiates, prescription pills, or phencyclidine) substances. Overall, QoL was lower for users of any substance compared to non-users.

In sum, five key lessons may be useful when planning and evaluating an intervention in a similarly vulnerable population:

1) Depending on the substance category, self-report may not adequately capture recent substance use. We found that agreement decreased over time for cocaine. Exclusively relying on self-report may result in underestimation of actual use.

2) In our study, using a wider time window of 90 days resulted in higher concordance with the saliva toxicological test compared to the corresponding recall window limited by the saliva toxicological test (3 days for cocaine, methamphetamine, amphetamine, and phencyclidine and 30 days for marijuana), indicating that distant recall was a better measure for current substance use behaviors in this population.

3) Predictors of misreporting can be used to estimate discordance among similar populations when biological verification is not feasible. Participants were more likely to misreport if they were older, non-White, and uninsured. These predictive variables can be used by other research teams to estimate misreporting in a demographically-similar population.

4) With regard to marijuana use, exclusively relying on an oral fluid biological test may result in an underestimation of actual substance use. Due to high agreement at baseline and increasing agreement over time, exclusive use of self-report may suffice to accurately capture marijuana use.

5) Quality of life was lower for users of any substance compared to non-users. In addition, quality of life was significantly predicted by substance use problems measured via the InDUC

instrument, race, and frequency of the Other category (cocaine, amphetamine, methamphetamine, opiates, prescription pills, or phencyclidine) of substances.

References

Aiemagno, S. a, Cochran, D., Feucht, T. E., Stephens, R. C., Butts, J. M., & Wolfe, S. a. (1996). Assessing substance abuse treatment needs among the homeless: a telephone-based interactive voice response system. *American Journal of Public Health*, *86*(11), 1626–1628. http://doi.org/10.2105/AJPH.86.11.1626

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders*. *American Journal of Psychiatry* (5th ed.). http://doi.org/10.1176/appi.books.9780890425596.744053

Babor, T. F. & 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. https://doi.org/10.1037/0022-006X.72.3.455.

Bennett, G. A., 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. http://doi.org/10.1016/j.drugalcdep.2003.08.001

Bennett, M. E., Nidecker, M., Strong Kinnaman, J. E., Li, L., & Bellack, A. S. (2009). Examination of the inventory of drug use consequences with individuals with serious and persistent mental illness and co-occurring substance use disorders. *The American Journal of Drug and Alcohol Abuse*, *35*(5), 385–390. http://doi.org/10.1080/00952990903177228

Campêlo, S. R., Barbosa, M. A., Dias, D. R., Caixeta, C. C., Leles, C. R., & Porto, C. C. (2017). Association between severity of illicit drug dependence and quality of life in a psychosocial care center in BRAZIL: Cross-sectional study. *Health and Quality of Life Outcomes*. http://doi.org/10.1186/s12955-017-0795-5

Carroll, K. M., Ball, S. A., Nich, C., Martino, S., Frankforter, T. L., Farentinos, C., … Woody, G. E. (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. http://doi.org/10.1016/j.drugalcdep.2005.08.002

Carroll, K. M., Fenton, L. R., Ball, S. A., Nich, C., Frankforter, T. L., Shi, J., & Rounsaville, B. J. (2004). Efficacy of Disulfiram and Cognitive Behavior Therapy in Cocaine-DependentOutpatients. *Archives of General Psychiatry*, *61*(3), 264. http://doi.org/10.1001/archpsyc.61.3.264

Clark, C. B., Zyambo, C. M., Li, Y., & Cropsey, K. L. (2016). The impact of non-concordant self-report of substance use in clinical trials research. *Addictive Behaviors*, *58*, 74–79. http://doi.org/10.1016/j.addbeh.2016.02.023

Cone, E. J. (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. http://doi.org/10.1016/j.forsciint.2011.01.046

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*(4), 349–354. http://doi.org/10.1037/h0047358

Darke, S. (1998). Self-report among injecting drug users: A review. *Drug and Alcohol Dependence*, *51*(3), 253–269. http://doi.org/10.1016/S0376-8716(98)00028-3

De Maeyer, J., Vanderplasschen, W., & Broekaert, E. (2010). Quality of life among opiate-dependent individuals: A review of the literature. *International Journal of Drug Policy*. http://doi.org/10.1016/j.drugpo.2010.01.010

Dey, M., Gmel, G., & Mohler-Kuo, M. (2013). Body mass index and health-related quality of life among young Swiss men. *BMC Public Health*. http://doi.org/10.1186/1471-2458-13-1028

Dey, M., Gmel, G., Studer, J., & Mohler-Kuo, M. (2014). Health-risk behaviors and quality of life among young men. *Quality of Life Research*. http://doi.org/10.1007/s11136-013-0524-4

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.

Dillon, F. R., Turner, C. W., Robbins, M. S., & 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. https://doi.org/10.1037/0893-164X.19.4.404.

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. https://doi.org/10.1080/09595230410001704208.

Donovan, D., Mattson, M. E., Cisler, R. A., Longabaugh, R., & Zweben, A. (2005). Quality of Life as an Outcome Measure in Alcoholism Treatment Research. *Journal of Studies on Alcohol*, *15*, 119–139.

Edens, E. L., Mares, A. S., Tsai, J., & Rosenheck, R. a. (2011). Does active substance use at housing entry impair outcomes in supported housing for chronically homeless persons? *Psychiatric Services (Washington, D.C.)*, *62*(2), 171–178. http://doi.org/10.1176/appi.ps.62.2.171

Elkashef, A., Holmes, T. H., Bloch, D. A., Shoptaw, S., Kampman, K., Reid, M. S., … 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. http://doi.org/10.1111/j.1360-0443.2005.00986.x

Endicott, J., Nee, J., Harrison, W., & Blumenthal, R. (1993). Quality of Life Enjoyment and Satisfaction

Questionnaire: a new measure. *Psychopharmacology Bulletin*, *29*(2), 321–326. http://doi.org/10.1111/j.1365-2850.2011.01735.x

Evans, E. A., Grella, C. E., Washington, D. L., & Upchurch, D. M. (2017). Gender and race/ethnic differences in the persistence of alcohol, drug, and poly-substance use disorders. *Drug and Alcohol Dependence*, *174*, 128–136. http://doi.org/10.1016/j.drugalcdep.2017.01.021

Gillaspy, J. A., & Campbell, T. C. (2006). Reliability and validity of scores from the inventory of drug use consequences. *Journal of Addictions and Offender Counseling*, *27*(1), 17–27. http://doi.org/10.1002/j.2161-1874.2006.tb00015.x

Godley, M. D., Godley, S. H., Dennis, M. L., Funk, R., & Passetti, L. L. (2002). Preliminary outcomes from the assertive continuing care experiment for adolescents discharged from residential treatment. *Journal of Substance Abuse Treatment*, *23*(1), 21–32. http://doi.org/10.1016/S0740-5472(02)00230-1

Goldenberg, M., IsHak, W. W., & Danovitch, I. (2017). Quality of life and recreational cannabis use. *American Journal on Addictions*. http://doi.org/10.1111/ajad.12486

Goldfinger, S. M., Schutt, R. K., Seidman, L. J., Turner, W. M., Penk, W. E., & Tolomiczenko, G. S. (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. http://dx.doi.org/10.1097/00005053-199611000-00003.

Goldstein, A. S., Gavojdea, S., & Zogg, D. F. (1994). Oral Collection Device and Method for Immunoassay. United States.

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. http://doi.org/10.1016/j.schres.2009.07.020

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.

Havassy, B. E., & Arns, P. G. (1998). Relationship of Cocaine and Other Substance Dependence to Well-Being of High-Risk Psychiatric Patients. *Psychiatric Services*, *49*(7), 935–940. http://doi.org/10.1176/ps.49.7.935

Hjorthøj, C. R., Hjorthøj, A. R., & 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. http://doi.org/10.1016/j.addbeh.2011.11.025

Hodgins, D. C., & el-Guebaly, N. (1992). More data on the Addiction Severity Index. Reliability and validity

with the mentally ill substance abuser. *J Nerv Ment Dis*, *180*(3), 197–201. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1588339

Jones, C. M., Logan, J., Gladden, R. M., & Bohm, M. K. (2015). Vital Signs: Demographic and Substance Use Trends Among Heroin Users - United States, 2002-2013. *MMWR. Morbidity and Mortality Weekly Report*, *64*(26), 719–25. http://doi.org/mm6426a3 [pii]

Jones, C. M., Mack, K. A., & Paulozzi, L. J. (2013). Pharmaceutical overdose deaths, United States, 2010. *JAMA - Journal of the American Medical Association*. http://doi.org/10.1001/jama.2013.272

Kessell, E. R., Bhatia, R., Bamberger, J. D., & Kushel, M. B. (2006). Public health care utilization in a cohort of homeless adult applicants to a supportive housing program. *Journal of Urban Health*, *83*(5), 860–873. http://doi.org/10.1007/s11524-006-9083-0

Kiluk, B. D., Dreifuss, J. A., Weiss, R. D., Morgenstern, J., & Carroll, K. M. (2013). The Short Inventory of Problems - Revised (SIP-R): Psychometric properties within a large, diverse sample of substance use disorder treatment seekers. *Psychology of Addictive Behaviors*, *27*(1), 307–314. http://doi.org/10.1037/a0028445

Kiluk, B. D., Nich, C., Babuscio, T., & Carroll, K. M. (2010). Quality versus quantity: Acquisition of coping skills following computerized cognitive-behavioral therapy for substance use disorders. *Addiction*, *105*(12), 2120–2127. http://doi.org/10.1111/j.1360-0443.2010.03076.x

Kirst, M., Zerger, S., Misir, V., Hwang, S., & Stergiopoulos, V. (2015). The impact of a Housing First randomized controlled trial on substance use problems among homeless individuals with mental illness. *Drug and Alcohol Dependence*, *146*, 24–9. http://doi.org/10.1016/j.drugalcdep.2014.10.019

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613. http://doi.org/10.1046/j.1525-1497.2001.016009606.x

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174. https://doi.org/10.2307/2529310

Larimer, M. E., Malone, D. K., Garner, M. D., Atkins, D. C., Burlingham, B., Lonczak, H. S., … Marlatt, G. A. (2009). Health care and public service use and costs before and after provision of housing for chronically homeless persons with severe alcohol problems. *JAMA : The Journal of the American Medical Association*, *301*(13), 1349–1357. http://doi.org/10.1001/jama.2009.414

Laudet, A. B., Becker, J. B., & White, W. L. (2009). Don't wanna go through that madness no more: Quality of life satisfaction as predictor of sustained remission from illicit drug misuse. *Substance Use and Misuse*. http://doi.org/10.1080/10826080802714462

Maisto, S. A., & Cooper, A. M. (1980). A historical perspective on alcohol and drug treatment outcome research. In L. C. Sobell, M. B. Sobell, & E. Ward (Eds.), *Evaluating Alcohol and Drug Abuse Treatment Effectiveness* (pp. 1–14). New York: Pergamon Press.

Marini, M., Schnornberger, T. M., Brandalise, G. B., Bergozza, M., & Heldt, E. (2013). Quality of life determinants in patients of a psychosocial care center for alcohol and other drug users. *Issues in Mental Health Nursing*. http://doi.org/10.3109/01612840.2013.780118

Matt, G. E., Turingan, M. R., Dinh, Q. T., Felsch, J. A., Hovell, M. F., & Gehrman, C. (2003). Improving self-reports of drug-use: Numeric estimates as fuzzy sets. *Addiction*, *98*(9), 1239–1247. http://doi.org/10.1046/j.1360-0443.2003.00444.x

McCambridge, J., & Kypri, K. (2011). Can simply answering research questions change behaviour? Systematic review and meta analyses of brief alcohol intervention trials. *PLoS ONE*. http://doi.org/10.1371/journal.pone.0023748

McDowell, D., Nunes, E. V., Seracini, A. M., Rothenberg, J., Vosburg, S. K., Ma, G. J., & Petkova, E. (2005). Desipramine treatment of cocaine-dependent patients with depression: A placebo-controlled trial. *Drug and Alcohol Dependence*, *80*(2), 209–221. http://doi.org/10.1016/j.drugalcdep.2005.03.026

McKenna, M., Chick, J., Buxton, M., Howlett, H., Patience, D., & Ritson, B. (1996). The SECCAT survey: I. The costs and consequences of alcoholism. *Alcohol and Alcoholism*, *31*(6), 565–576. http://doi.org/10.1093/oxfordjournals.alcalc.a008192

McLellan, A. T., Luborsky, L., Woody, G. E., & O'Brien, C. P. (1980). An improved diagnostic evaluation instrument for substance abuse patients: The addiction severity index. *Journal of Nervous and Mental Disease*, *168*(1), 26–33. http://doi.org/10.1097/00005053-198001000-00006

Miller, W. R., Tonigan, J. S., & Longabaugh, R. (1995). The drinker inventory of consequences (DrInC). *The Drinker Inventory of Consequences (DrInC): An Instrument for Assessing Adverse Consequences of Alcohol Abuse*, *4*. Retrieved from http://pubs.niaaa.nih.gov/publications/drinc.pdf

Mondello, M., & House, S. (2007). Cost Analysis of Permanent Supportive Housing State of Maine - Greater Portland. *Mental Health Care*, (September), 1–22.

Morgenstern, J., Hogue, A., Dasaro, C., Kuerbis, A., & Dauber, S. (2008a). 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), 561–70. http://doi.org/10.15288/jsad.2008.69.561

Moss, H. B., Chen, C. M., & Yi, H. ye. (2014). Early adolescent patterns of alcohol, cigarettes, and marijuana polysubstance use and young adult substance use outcomes in a nationally representative sample.

Drug and Alcohol Dependence, *136*(1), 51–62. http://doi.org/10.1016/j.drugalcdep.2013.12.011

Moyer, A., Finney, J. W., Swearingen, C. E., & Vergun, P. (2002). Brief interventions for alcohol problems: A meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. *Addiction*. http://doi.org/10.1046/j.1360-0443.2002.00018.x

Muller, A. E., Skurtveit, S., & Clausen, T. (2016). Many correlates of poor quality of life among substance users entering treatment are not addiction-specific. *Health and Quality of Life Outcomes*. http://doi.org/10.1186/s12955-016-0439-1

Napper, L. E., Fisher, D. G., Johnson, M. E., & Wood, M. M. (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. http://doi.org/10.1016/j.addbeh.2009.12.006

National Institute on Alcohol Abuse and Alcoholism. (n.d.). What is a Standard Drink?

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. http://doi.org/10.1016/S0376-8716(03)00053-X

Nosyk, B., Li, L., Evans, E., Huang, D., Min, J., Kerr, T., … Hser, Y. I. (2014). Characterizing longitudinal health state transitions among heroin, cocaine, and methamphetamine users. *Drug and Alcohol Dependence*, *140*, 69–77. http://doi.org/10.1016/j.drugalcdep.2014.03.029

Padgett, D. K. (2006). Housing First Services for People Who Are Homeless With Co-Occurring Serious Mental Illness and Substance Abuse. *Research on Social Work Practice*, *16*(1), 74–83. http://doi.org/10.1177/1049731505282593

Palepu, A., Marshall, B. D. L., Lai, C., Wood, E., & Kerr, T. (2010). Addiction treatment and stable housing among a cohort of injection drug users. *PLoS ONE*, *5*(7). http://doi.org/10.1371/journal.pone.0011697

Pasareanu, A. R., Opsal, A., Vederhus, J. K., Kristensen, Ø., & Clausen, T. (2015). Quality of life improved following in-patient substance use disorder treatment. *Health and Quality of Life Outcomes*, *13*(1). http://doi.org/10.1186/s12955-015-0231-7

Pattison, E. M., Sobell, M. B., & Sobell, L. C. (1977). Emerging concepts of alcohol dependence. New York: Springer Publishing.

Pettinati, H. M., Kampman, K. M., Lynch, K. G., Suh, J. J., Dackis, C. A., Oslin, D. W., & O'Brien, C. P. (2008). Gender differences with high-dose naltrexone in patients with co-occurring cocaine and alcohol dependence. *Journal of Substance Abuse Treatment*, *34*(4), 378–390. http://doi.org/10.1016/j.jsat.2007.05.011

Pettinati, H. M., Kampman, K. M., Lynch, K. G., Xie, H., Dackis, C., Rabinowitz, A. R., & OBrien, C. P. (2008). 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. http://doi.org/10.1016/j.addbeh.2007.11.011

Polcin, D. L. (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. https://doi.org/10.1179/1573658X15Y.0000000004.

Quintela, O., & Crouch, D. J. (2006). Recovery of drugs of abuse from the immunalysis Quantisal oral fluid collection device. Journal of Analytical Toxicology, 30, 614–616. https://doi.org/10.1093/jat/30.8.614.

Redwood Toxicology Laboratory. (2015). *Laboratory Testing Reference Guide*.

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. http://doi.org/10.1016/j.addbeh.2017.02.011

Riendeau, R. P., Sullivan, J. L., Meterko, M., Stolzmann, K., Williamson, A. K., Miller, C. J., … Bauer, M. S. (2018). Factor structure of the Q-LES-Q short form in an enrolled mental health clinic population. *Quality of Life Research*, *27*(11), 2953–2964. http://doi.org/10.1007/s11136-018-1963-8

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

Rohsenow, D. J., Martin, R. A., Eaton, C. A., & Monti, P. M. (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. http://doi.org/10.15288/jsad.2007.68.641

Rosay, A. B., Najaka, S. S., & Herz, D. C. (2007). Differences in the validity of self-reported drug use across five factors: Gender, race, age, type of drug, and offense seriousness. *Journal of Quantitative Criminology*, *23*(1), 41–58. http://doi.org/10.1007/s10940-006-9018-7

Schumacher, J. E., Milby, J. B., Raczynski, J. M., 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. https://doi.org/10.1016/0740 5472(95)02009-5.

Secades-Villa, R., & Fernandez-Hermida, J. R. (2003). The validity of self-reports in a follow-up study with drug addicts. *Addict.Behav.*, *28*(0306–4603 (Print)), 1175–1182. http://doi.org/10.1016/S0306-4603(02)00219-8

Sloan, J. J., Bodapati, M. R., & Tucker, T. A. (2004). Respondent misreporting of drug use in self-reports:

Social desirability and other correlates. *Journal of Drug Issues*, *34*(2), 269–292. http://doi.org/10.1177/002204260403400202

Smith, K. W., & Larson, M. J. (2003). Quality of life assessments by adult substance abusers receiving publicly funded treatment in Massachusetts. *American Journal of Drug and Alcohol Abuse*. http://doi.org/10.1081/ADA-120020517

Sobell, L., & Sobell, M. (1981). Frequent follow-up as data gathering and continued care with alcoholics. *Substance Use & Misuse*, *16*(6), 1077–1086. http://doi.org/10.3109/10826088109038914

Sobell, L. C., & Sobell, M. B. (1992). Timeline follow-back. In: Litten, R.Z., Allen, J.P. (eds) Measuring Alcohol Consumption. Humana Press, Totowa, NJ, USA. https://doi.org/10.1007/978-1-4612-0357-5_3.

Somoza, E., Somoza, P., Lewis, D., Li, S. H., 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. http://doi.org/10.1016/j.drugalcdep.2007.09.007

Stasiewicz, P. R., Vincent, P. C., Bradizza, C. M., Connors, G. J., Maisto, S. A., & Mercer, N.D. (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. https://doi.org/10.1037/0893-164X.22.1.78.

Stein, J. a, Dixon, E. L., & Nyamathi, A. M. (2008). Effects of psychosocial and situational variables on substance abuse among homeless adults. *Psychology of Addictive Behaviors : Journal of the Society of Psychologists in Addictive Behaviors*, *22*(3), 410–416. http://doi.org/10.1037/0893-164X.22.3.410

Stevanovic, D. (2011). Quality of life enjoyment and satisfaction questionnaire - short form for quality of life assessments in clinical practice: A psychometric study. *Journal of Psychiatric and Mental Health Nursing*. http://doi.org/10.1111/j.1365-2850.2011.01735.x

Tiffany, S. T., Friedman, L., Greenfield, S. F., Hasin, D. S., & Jackson, R. (2012). Beyond drug use: A systematic consideration of other outcomes in evaluations of treatments for substance use disorders. *Addiction*. http://doi.org/10.1111/j.1360-0443.2011.03581.x

Tonigan, J. S., & Miller, W. R. (2002). The Inventory of Drug Use Consequences (InDUC): Test-retest stability and sensitivity to detect change. *Psychology of Addictive Behaviors*. http://doi.org/10.1037/0893-164X.16.2.165

Torchalla, I., Strehlau, V., Li, K., & Krausz, M. (2011). Substance use and predictors of substance

dependence in homeless women. *Drug and Alcohol Dependence*, *118*(2–3), 173–179. http://doi.org/10.1016/j.drugalcdep.2011.03.016

U.S. Department of Health and Human Services: Food and Drug Administration. (2009). Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. *Clinical/Medical Federal Register*, (December), 1–39. http://doi.org/10.1111/j.1524-4733.2009.00609.x

Volk, R. J., Cantor, S. B., Steinbauer, J. R., & Cass, A. R. (1997). Alcohol use disorders, consumption patterns, and health-related quality of life of primary care patients. *Alcoholism: Clinical and Experimental Research*. http://doi.org/10.1111/j.1530-0277.1997.tb03855.x

von Eye, A., & Mun, E. Y. (2003). Characteristics of measures for 2 × 2 tables. Understanding Statistics, 2(4), 243-266. https://doi.org/10.1207/S15328031US0204_02.

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

Walters, S. T., & Spence-Almaguer, E. E. (2015). Integrating health coaching and technology with vulnerable clients. *Social Work Today*, 6–7.

Walters, S. T., Vader, A. M., Harris, T. R., & Jouriles, E. N. (2009). Reactivity to alcohol assessment measures: An experimental test. *Addiction*, *104*(8), 1305–1310. http://doi.org/10.1111/j.1360-0443.2009.02632.x

Warner, M., Chen, L. H., & Makuc, D. M. (2009). Increase in fatal poisonings involving opioid analgesics in the United States, 1999-2006. *NCHS Data Brief*, (22), 1–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19796521

Watson, D. P., Wagner, D. E., & Rivers, M. (2013). Understanding the critical ingredients for facilitating consumer change in housing first programming: A case study approach. *Journal of Behavioral Health Services and Research*, *40*(2), 169–179. http://doi.org/10.1007/s11414-012-9312-0

Willenbring, M. L., Whelan, J. A., Dahlquist, J. S., & O'Neal, M. E. (1990). Community treatment of the chronic public inebriate I: Implementation. Alcoholism Treatment Quarterly, 7(1), 79–97. https://doi.org/10.1300/J020V07N01_07.

Wood-Dauphinee, S. (1999). Assessing quality of life in clinical research: from where have we come and where are we going? *Journal of Clinical Epidemiology*, *52*(4), 355–363. http://doi.org/10.1016/S0895-4356(98)00179-6

Zemore, S. E. (2012). The effect of social desirability on reported motivation, substance use severity, and treatment attendance. *Journal of Substance Abuse Treatment*, *42*(4), 400–412.

http://doi.org/10.1016/j.jsat.2011.09.013

Zerger, S. (2002). Substance abuse treatment: What works for homeless people? A Review of the literature. National Health Care for the Homeless Council. Retrieved from https://www.nhchc.org/wp-content/uploads/2012/02/SubstanceAbuseTreatmentLitReview.pdf.

Addictive Behaviors 70 (2017) 90–96

Contents lists available at ScienceDirect

Addictive Behaviors

journal homepage: www.elsevier.com/locate/addictbeh

What's the agreement between self-reported and biochemical verification of drug use? A look at permanent supportive housing residents

Alexis Rendon [a,*], Melvin Livingston [b], Sumihiro Suzuki [b], Whitney Hill [a], Scott Walters [a]

[a] University of North Texas Health Science Center, School of Public Health, Department of Behavioral and Community Health, Fort Worth, TX, United States
[b] University of North Texas Health Science Center, School of Public Health, Department of Biostatistics and Epidemiology, Fort Worth, TX, United States

HIGHLIGHTS

- Timeline Follow-back self-reported drug use is compared to an oral fluid test.
- Self-report may not adequately capture recent drug use in similar populations.
- Older, non-White, and uninsured participants were more likely to misreport use.
- With marijuana, relying on an oral fluid test may underestimate actual drug use.

ARTICLE INFO

Article history:
Received 3 September 2016
Received in revised form 6 January 2017
Accepted 8 February 2017
Available online 10 February 2017

Keywords:
Substance abuse
Self-report
Timeline Follow-back
Chronic homelessness
Supportive housing

ABSTRACT

Self-reported substance use is commonly used as an outcome measure in treatment research. We evaluated the validity of self-reported drug use in a sample of 334 adults with mental health problems who were residing in supportive housing programs. The primary analysis was the calculation of the positive predictive values (PPVs) of self-report compared to an oral fluid test taken at the same time. A sensitivity analysis compared the positive predictive values of two self-reported drug use histories: biological testing window (ranging between the past 96 h to 30 days depending on drug type) or the full past 90-day comparison window (maximum length recorded during interview). A multivariable logistic regression was used to predict discordance between self-report and the drug test for users. Self-reported drug use and oral fluid drug tests were compared to determine the positive predictive value for amphetamines/methamphetamines/PCP (47.1% agreement), cocaine (43.8% agreement), and marijuana (69.7% agreement) drug tests. Participants who misreported their drug use were more likely to be older, non-White, have no medical insurance, and not report any alcohol use. In general, amphetamine/methamphetamine/PCP and cocaine use was adequately captured by the biological test, while marijuana use was best captured by a combination of self-report and biological data. Using the full past 90 day comparison window resulted in higher concordance with the oral fluid drug test, indicating that self-reported drug use in the past 90 days may be a proxy for drug use within the biological testing window. Self-report has some disadvantages when used as the sole measure of drug use in this population.

1. Introduction

The most severe type of homelessness is chronic homelessness (CH), defined as individuals who are homeless for at least a year within the last three years or who have had four separate, distinct, and sustained stays of homelessness in the past year (The Substance Abuse and Mental Health Services Administration [SAMHSA], 2011; Tsai, Lapidos, Rosenheck, & Harpaz-Rotem, 2013; United States Department of Housing and Development, 2007). Although CH individuals make up about 25% of the homeless population, they account for a disproportionate share of health and social services costs (Burt & Aron, 2001; Caton, Wilkins, & Anderson, 2007; Larimer et al., 2009). Two common features of CH individuals are mental health problems and substance use. For instance, the prevalence of lifetime mental illness in the CH population is 74%–83% (Edens, Mares, & Rosenheck, 2011), compared to lifetime rates of 4.2% in the general population (The Substance Abuse and Mental Health Services Administration [SAMHSA], 2010). Similarly, rates of

* Corresponding author at: School of Public Health, The University of North Texas Health Science Center, 3500 Camp Bowie Blvd Suite 708, Fort Worth, TX 76107, United States.
E-mail address: alexis.rendon@live.unthsc.edu (A. Rendon).

http://dx.doi.org/10.1016/j.addbeh.2017.02.011
0306-4603/

lifetime substance use disorders among CH individuals are approximately 68%–73% (Edens, Mares, & Rosenheck, 2011) compared to lifetime rates of 9.4% in the general population (SAMHSA, 2011). Those with co-occurring mental health and substance use disorders often suffer from more severe non-compliant behaviors compared to those with mental health difficulties or substance use alone (Drake & Wallach, 1989).

Permanent supportive housing (PSH) combines housing and supportive case management to meet the needs of CH individuals (Larimer et al., 2009; Polcin, 2016). There is evidence that PSH reduces overall healthcare costs (Larimer et al., 2009); however, the effect of PSH on substance use remains under debate (Edens, Mares, Tsai, & Rosenheck, 2011; Kirst, Zerger, Misir, Hwang, & Stergiopoulos, 2015; Polcin, 2016). Several studies have reported declines in alcohol use, but not in illicit drug use after entering PSH (Kirst et al., 2015; Larimer et al., 2009; Padgett, 2006). While these studies found no change in rates of illicit substance use, another study reported that PSH residents increased drug treatment services by 22% after being housed, which resulted in decreased drug use (Mondello & House, 2007). One limitation of these studies is that they tend to rely exclusively on self-report as a measure of substance use (Kirst et al., 2015; Padgett, 2006).

Verbal recall of substance use is common in research studies assessing drug use behavior (Darke, 1998; Hjorthøj, Hjorthøj, & Nordentoft, 2012; Schumacher et al., 1995), and likewise assessments of the efficacy of supportive housing programs tend to rely on this method (Napper, Fisher, Johnson, & Wood, 2010; Larimer et al., 2009). The most common method, Timeline Follow-back (TLFB), is a self-report instrument that utilizes a visual calendar to enhance recall of substance use (Sobell & Sobell, 1992). Originally developed to measure alcohol consumption, the TLFB has since been widely used in cross sectional and prospective studies of drug use (Hjorthøj et al., 2012). While researchers have used the TLFB method for recall of up to the past 12 months, a 90-day TLFB is common in substance abuse treatment studies (Dennis, Funk, Godley, Godley, & Waldron, 2004; Sobell, Brown, Leo, & Sobell, 1996).

While the TLFB is generally concordant with biological measures, some populations may be more accurate in their reporting (Harrison, 1997; Hjorthøj et al., 2012; Napper et al., 2010; Rosay, Najaka, & Herz, 2007; Secades-Villa & Fernandez-Hermida, 2003). For example, in a study of people being discharged from drug treatment, self-report was an accurate measure of amphetamine use within the past 48 h, with 95% agreement when compared to drug urine tests (Napper et al., 2010). However, in other studies, discordance has been as high as 34.9% among users, depending on the substance type and the reporting population (Hjorthøj et al., 2012; Schumacher et al., 1995). For example, Schumacher et al. (1995) found an average 30-day concordance rate of 68% among 131 homeless crack cocaine users, compared to urinalysis. A meta-analysis found that the percent agreement between self-report and biological measures ranged from 87.3%–90.9% for marijuana and 79.3%–84.1% for cocaine (Hjorthøj et al., 2012). Among studies in which substance users had psychiatric co-morbidities, the percent agreement ranged from 80.4%–83.8% (Hjorthøj et al., 2012). There are several reasons studies might report differences in the accuracy of self-report. The lack of standardized methodology, social desirability or stigma of reporting illicit drug use, inability to recall drug use further back in time, and fear of legal repercussions may lead individuals to misreport their substance use (Napper et al., 2010).

Despite the acknowledged tendency of self-report to underestimate actual drug use, self-report is still the primary measure of drug use in studies of homeless substance users (Napper et al., 2010). A study conducted during 2004–2008 estimated the prevalence of past 30 day illicit drug use among 756 CH research participants as 36%–39% at baseline (Edens, Mares, & Rosenheck, 2011). However, because these data were collected exclusively by self-report, it is possible that this underestimates true drug use in this population. The validity of self-reported drug use in both CH and supportive housing populations has not been adequately studied (Polcin, 2016). To our knowledge, no other study has estimated the prevalence of substance use in a supportive housing or a similar low-income population with mental health disorders using a biological measure to validate self-report.

Demographic factors have sometimes been associated with misreporting, but the overall patterns are unclear. For instance, there is disagreement in the literature about whether age is a predictor of misreporting drug use (Katz, Webb, Gartin, & Marshall, 1997; McElrath, Dunham, & Cromwell, 1995; McNagny & Parker, 1992; Rosay et al., 2007). The relationship between race and reporting drug use has also been a point of disagreement (Rosay et al., 2007). Studies of other vulnerable populations have found that Blacks are less likely to have a concordant self-report and urinalysis (White et al., 2014). Race was not a significant contributor to a study of self-reporting drug use validity among arrestees (Sloan, Odapati, & Ucker, 2004), but was in another study of arrestees (McElrath et al., 1995). This demonstrates that the relationship between race and misreported drug use may be population-specific. However, evidence suggests that sex is not a predictor of misreporting (Sloan et al., 2004). Finally, it is unclear whether insurance status is a predictor of misreporting. However, compared to those with private insurance, uninsured individuals have increased odds of alcohol and substance abuse disorders and also experience barriers to accessing substance abuse treatment services (Wu, Kouzis, & Schlenger, 2003), and thus there is reason to believe that insurance status might be associated with misreporting.

While it is known that supportive housing individuals incur large healthcare related costs, knowing who is more likely to misreport drug use can be helpful to researchers who design and evaluate programs for similar populations. For instance, without knowing the validity of self-reported drug use, researchers will not be able to accurately measure drug use or make valid conclusions about the efficacy of interventions. This study aimed to determine the validity of self-reported drug use compared to a biological drug test and assess predictors of misreporting in a group of people residing in supportive housing. The overall goal was to provide further information for others who are seeking to obtain accurate measures of substance abuse in vulnerable populations.

2. Methods

2.1. Participants

Participants were adults (18 years and older), residing in PSH in Fort Worth, TX, who were interested in participating in a voluntary health coaching program. To be eligible, participants must have been Medicaid-enrolled or low-income uninsured (Medicaid eligible), and self-reported one of the following mental health problems: prescribed medication for psychological or emotional problems, experiencing hallucinations, receiving a pension for a psychiatric disability, or scoring > 9 on the 9-item Patient Health Questionnaire (PHQ-9) depression screener. Exclusion criteria included (1) residing in other types of housing not considered PSH (e.g., Transitional Housing or homeless shelter), (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/or (4) limited autonomy or decision-making capabilities (e.g., substantially neurologically or cognitively impaired). Convenience sampling of six local housing agencies resulted in 463 people who were screened for eligibility. The final sample consisted of 334 participants who met the inclusion criteria (Fig. 1). (Among the 399 PHS residents who were screened for eligibility, approximately 83.7% met the other inclusion criteria.) The project was approved by the Institutional Review Board of the University of North Texas Health Science Center, and participants were given assurances of confidentiality. Informed consent was obtained from each study participant.

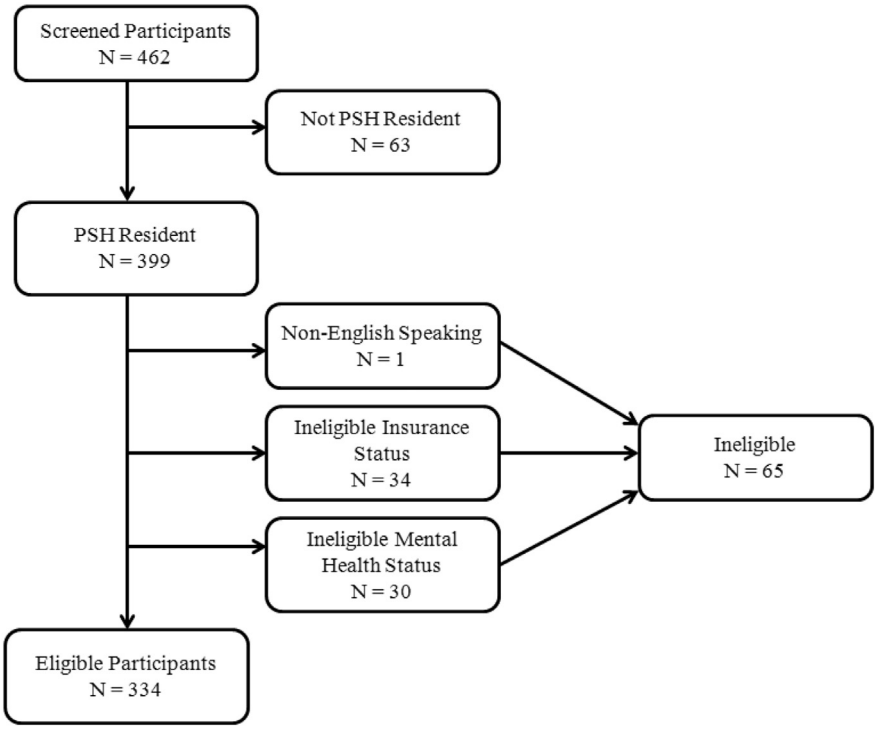

Fig. 1. Study recruitment and enrollment.

2.2. Instruments

Self-report and oral fluid drug test data were gathered during baseline in-person interviews conducted between November 2014 and December 2015. The Timeline Follow-back (TLFB) method was used to collect self-reported data for the 90 days prior to the baseline interview. This assessment was administered by trained research assistants who used anchor dates (e.g., birthdays, holidays, anniversaries, appointments, life events) to help participants recall and establish patterns of drug and alcohol use during the past 90 days. For illicit drug use, participants indicated on which days they had used any substance. For alcohol use, the number of standard drinks was also recorded for each day indicated by the participant. This method of self-report has been widely used to estimate substance use (Hjorthøj et al., 2012; Sobell et al., 1996).

2.3. Procedure

An oral fluid drug test, also administered by the research assistant, was used to test for the presence of opiates, oxycodone, PCP, cocaine, marijuana, benzodiazepine, barbiturate, amphetamine, and methamphetamine.[1] The testing device included a swab attached to a plastic handle; participants placed the swab under their tongues until a small window at the end of the handle turned blue. This color change indicated that an adequate oral fluid sample had been collected for testing. At this point, the participant placed the swab into a buffer fluid-filled tube and the lid was snapped on by the participant. The sample was labeled with a bar code that referenced the participant's unique identification number, date and time of oral fluid sample collection, and name and contact information of the research assistant administering the test. Enzyme Immunoassay (EIA) was used to determine if samples were positive or negative. Positive samples were confirmed with gas chromatography/mass spectrometry.

2.4. Data analysis

Due to the low prevalence of use, the stimulant categories were combined into a single category. Responses to the self-report TLFB were compared to results from the oral fluid drug test. These results were compared for two separate time windows. The first window (Definition 1) compared the self-report to the suggested drug detection window (3 days for amphetamine/methamphetamine/PCP and cocaine or 30 days for marijuana). This method has been widely used in other studies comparing the validity of self-reported drug use to a biological standard (Hjorthøj et al., 2012). The second window (Definition 2) compared the drug test with the full 90-day self-report (Table 2). Participants who tested negative and reported no use were categorized as non-users. Participants whose drug test matched their self-report were categorized as concordant users (Fig. 2). Participants who did not report using drugs, but whose test was positive were categorized as discordant users (Fig. 2). Finally, those who reported using drug use, but whose test was not positive for drugs were categorized as discordant users (Fig. 2). Medical records were not obtained in this study. Therefore, due to the lack of verification, drug classes that corresponded to self-reported medical prescriptions were not used in the analyses (opiate ($n = 10$), barbiturate ($n = 1$), benzodiazepine ($n = 1$), and methadone ($n = 2$)).

Positive predictive values (PPVs), negative predictive values (NPVs), sensitivity, and specificity were determined for each drug category for both the recommended biological testing windows and the full 90 day TLFB window. The PPV were used to calculate concordance and discordance rates by substance type. The number of self-reported drug users was divided by the number of users determined by the oral fluid test for each time window (biological testing window and 90 day window). A comparison of the PPVs between the biological window and full 90-day window was conducted to test the robustness of the results. This was conducted as a sensitivity analysis to determine how many users

[1] The Quantisal™ Oral Fluid Collection Device used for oral fluid sample collection.

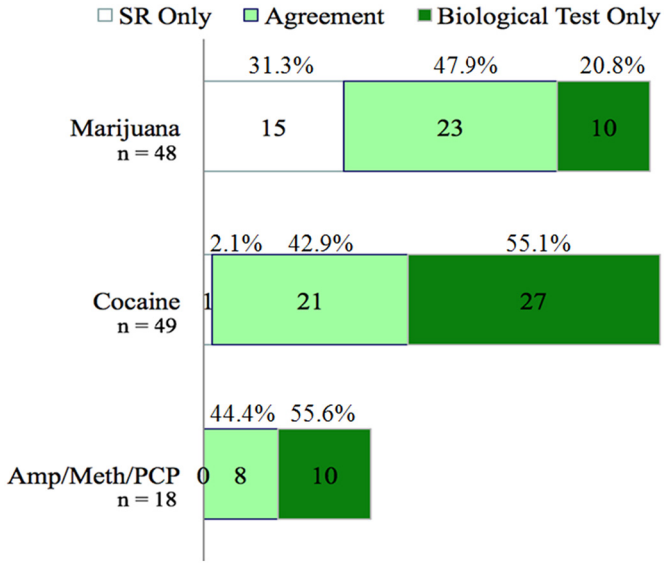

Fig. 2. Determination of user by 90-day self-report (SR) only, biological test only, or a combination of both.

who were discordant in the biological testing window became concordant once the reporting window was expanded to the full 90 days (Table 2).

3. Results

Substances measured in this study included amphetamines, methamphetamines, PCP, cocaine, and marijuana. The primary analysis evaluated the validity of self-reported drug use compared to the oral fluid drug test. Secondary analyses evaluated predictors of discordance between self-report and the oral fluid test. Univariate and multivariable logistic regression were used to predict discordance between self-reported drug use and the biological fluid test. Odds ratios were used to interpret the relationship between predictor variables and discordance. Predictive variables for this multivariable logistic regression analysis included race, gender, age, PHQ-9 depression severity score, alcohol use, time in a PSH program, and quality of life. Quality of life was measured by the Quality of Life Enjoyment and Satisfaction Questionnaire (QLESQ) (Endicott, Nee, Harrison, & Blumenthal, 1993). Analyses were completed using SAS statistical software, version 9.3 (SAS Institute Inc., Cary NC). Low representation of races other than White and Black/African American (n = 17 [5.1%]) led to all other races being combined into a non-White category (n = 193) (Table 1).

Table 1
Participant Characteristics (N = 334).

	N (%) M (SD)
Male	159 (47.6%)
Race	
White	141 (42.2%)
Non-White	193 (57.8%)
Medical insurance status	
Any (Medicaid, Medicare)	200 (59.9%)
Uninsured	134 (40.1%)
Self-reported alcohol use	
Any alcohol (past 90 days)	161 (48.2%)
No alcohol (past 90 days)	173 (51.8%)
Age (years)	51.3 (9.0)
Time in PSH (months)	25.1 (25.8)

Note: M = mean, SD = standard deviation, PSH = permanent supportive housing.

Substance users who self-reported use and tested positive on the biological drug test were classified as concordant users (Fig. 2). Overall, more marijuana users were concordant at 47.9% (n = 23) compared to 44.4% of amphetamine/methamphetamine/PCP users (n = 8) and 42.9% of cocaine users (n = 21) (Fig. 2).

3.1. Participant characteristics

The prevalence of drug use among participants by either a self-report or a positive drug test was 43.1% (n = 144). Among drug users who either self-reported or tested positive on the oral fluid test, 28.5% were positive for multiple substances (n = 41). A total of 48.2% of participants self-reported any alcohol use in the past 90 days (Table 1). Among participants who used alcohol, 47.2% reported drinking 5 or more drinks for men or 4 or more drinks for women on at least one day in the past 90 days.

3.2. Positive predictive value

Marijuana had the highest PPV at 69.7%, followed by amphetamine/methamphetamine/PCP at 47.1%, and cocaine at 43.8%. Widening the windows to the recall period of the past 90 days resulted in higher PPVs for all three drugs types: 84.8% for marijuana 70.6% for amphetamine/methamphetamine/PCP, and 60.4% for cocaine.

3.3. Logistic regression

Results from the logistic regression showed that age, race, medical insurance status, and self-reported alcohol use were significant predictors of discordance (Table 3). Gender was not a significant predictor of discordance (OR: 1.113, 95% CI: [0.503, 2.466]). The gender difference between participant and interviewer was not statistically significant in a univariate analysis and was thus excluded from further analysis. Univariate regression of the relationship between participant and interviewer race showed that this difference was not statistically significant and was not included in further analysis. The odds of misreporting drug use increased as age increased by one year (OR: 1.068, 95% CI: [1.014, 1.125]). Whites were less likely to misreport drug use compared to non-Whites (OR: 0.413, 95% CI: [0.178, 0.957]). Those with any medical insurance were less likely to misreport drug use compared to those without insurance (OR: 0.413, 95% CI: [0.181, 0.946]). Individuals who admitted to using alcohol in the past 90 days were less likely to misreport illicit drug use (OR: 0.406, 95% CI: [0.180, 0.915]).

4. Discussion

This study investigated the validity of self-reported drug use in a population of formerly homeless individuals residing in supportive housing. The results of this study are most directly applicable to PSH residents who use substances, however they may also have implications in similar populations of low-income clients with mental health disorders. These results are generally consistent with results from other studies of vulnerable populations (Aiemagno et al., 1996; Digiusto, Seres, Bibby, & Batey, 1996; Hjorthøj et al., 2012; Schumacher et al., 1995), where self-report tends to underestimate drug use. Age, race, medical insurance status, and alcohol use were significant predictors of discordance between self-reported drug use and the biochemical test. People with medical insurance and those who reported any alcohol use were less likely to misreport drug use. Further study of these associations may be useful in predicting misreporting in similar populations.

People who reported alcohol consumption in the past 90 days were more likely to accurately self-report their drug use, compared to those who did not report alcohol consumption. As the biological test did not capture alcohol use, the nature of this relationship remains untested. It is possible that people who feel comfortable disclosing alcohol use also feel less stigmatized about reporting drug use. Because this study

did not include a biological test for alcohol, it is not possible to determine if participants who misreported drug use also misreported alcohol use. Further research is needed to determine if actual alcohol use, or reported alcohol use is the better predictor of drug use reporting.

Compared to Whites, non-Whites were more likely to misreport drug use. However, a univariate analysis showed that there was not a significant relationship between interviewer and interviewee race. The relationship between interviewer and interviewee gender was also nonsignificant, which is consistent with previous findings examining gender as a predictor for misreporting drug use (Edens, Mares, & Rosenheck, 2011; Sloan et al., 2004). Participants who were uninsured were more likely to misreport drug use compared to those with insurance. Further study is needed to explore why those without insurance were more likely to misreport drug use.

This study found that relying exclusively on self-report to capture drug use among formerly homeless individuals captured less than half of actual users, depending on drug type (Table 2). These results are broadly consistent with a meta-analysis by Hjorthøj et al. (2012), which reported similar levels of sensitivity (0.60–0.88) to this study (0.44–0.7), but lower levels of specificity (0.42–1.0) compared to this study (0.95–1.0) (Hjorthøj et al., 2012). This means that this study was able to detect drug use as well as previous studies but better able to detect non-users compared to previous studies. The range in these estimates represents the difference between the two definitions of concordance: 90 day window for recall compared to biological window (3 days for amphetamines/methamphetamines/PCP and cocaine; 30 days for marijuana). For example, when self-report and the biological test were compared for cocaine for the biological window, the concordance rate was 43.8%. When the self-report recall window was widened from 3 days to 90 days, the concordance rate increased to 60.4%. In fact, all PPVs were higher using the 90-day window as compared to the recommended oral fluid drug test window. This increase in PPVs across all three substance types may reflect a greater willingness to acknowledge drug use that was further back in time. Self-reported past drug use may be a better proxy of recent use than asking about current use. Because people are not always reporting substance use when it occurs, prior substance use information may form a reasonable proxy of recent (non-reported) use in similar populations. The 90-day window did a better job of capturing more recent use that the biological window of the TLFB. This suggests that it may be advantageous for future researchers to use data from the full 90-day TLFB window as a proxy measure of recent drug use. This pattern of increased concordance between biochemical results and self-reported use that is further back in time has also been shown in other populations (Harrison, 1997; Hjorthøj et al., 2012). These findings could be used to better estimate true drug use rates as well as treatment effects in a field where self-report is used almost exclusively (Magura & Kang, 1996; Rosay et al., 2007).

Interestingly, a significant number of marijuana users self-reported use, but did not test positive on the biological drug test (n = 15, 31.3%). This may be attributable to a lower sensitivity rate in the oral fluid biological test for marijuana or to a bias in the sample towards reporting marijuana use. It is beyond the scope of this study to determine the cause of the mismatch between the self-report and the biological test for marijuana, however, it seems possible that relying exclusively on a biological test for marijuana may underestimate true rates of use.

Table 3
Odds ratio estimates for discordance predictors among drug users.

Variable	OR	95% CI
Age	1.068	(1.014, 1.125)
Gender	1.113	(0.503, 2.466)
Race	0.413	(0.178, 0.957)
Insurance	0.413	(0.181, 0.946)
Alcohol	0.406	(0.180, 0.915)

Note: N = 115. Reference categories: non-Whites, males, uninsured, and non-drinkers. OR = odds ratio.

These findings suggest that the accuracy of self-report can be improved by increasing the length of time for recall. The study methods limited the analysis to the past 90 days, which demonstrated higher concordance rates with the biochemical measure (Table 2). When biological testing is not feasible, widening self-report windows to 90 days may be a better proxy for measuring current use in this population. Widening the self-report recall windows may be a better proxy of self-reported drug use in similarly vulnerable populations with a high prevalence of mental illness.

The concordance rate between self-reported and biochemical verification of drug use was generally lower in our study compared to other studies. The findings of the full 90 day window for marijuana concordance reached 84.8%, which is in line with previous estimates, but other concordance rates fell short of estimates obtained in other studies (Hjorthøj et al., 2012; Napper et al., 2010; White et al., 2014). There may be additional factors in our population that would decrease the accuracy of self-report when compared to studies of people with co-morbid psychiatric conditions, but not a history of chronic homelessness (Hjorthøj et al., 2012).

4.1. Limitations

Our study had several limitations. First, the sensitivity of the oral fluid drug testing method may result in an underestimation of actual drug use, which would affect the misreporting estimate. However, oral fluid testing has been accepted as a sensitive measure of recent cocaine, amphetamine, methamphetamine, and marijuana usage (Allen, 2011; Dams, Choo, Lambert, Jones, & Huestis, 2007), and the expected impact on these estimates should be minimal. This method is also advantageous due to the low potential of adulteration, minimal possibility of

Table 2
Statistics of self-report against biochemical test for biological and 90 day windows (N = 334).

		Definition 1: biological testing window[a]	Definition 2: 90 day window
Marijuana	PPV (95% CI)	0.697 (0.511, 0.838)	0.848 (0.673, 0.943)
	NPV (95% CI)	0.967 (0.940, 0.984)	0.951 (0.921, 0.972)
	Sensitivity (95% CI)	0.697 (0.513, 0.844)	0.651 (0.491, 0.790)
	Specificity (95% CI)	0.951 (0.921, 0.972)	0.983 (0.961, 0.995)
Cocaine	PPV (95% CI)	0.438 (0.298, 0.587)	0.604 (0.453, 0.739)
	NPV (95% CI)	0.915 (0.879, 0.943)	0.939 (0.906, 0.963)
	Sensitivity (95% CI)	0.438 (0.295, 0.588)	0.604 (0.453, 0.742)
	Specificity (95% CI)	0.997 (0.981, 0.999)	0.997 (0.981, 0.999)
Amphetamine/methamphetamine/PCP	PPV (95% CI)	0.471 (0.239, 0.715)	0.706 (0.440, 0.886)
	NPV (95% CI)	0.970 (0.945, 0.986)	0.982 (0.961, 0.993)
	Sensitivity (95% CI)	0.444 (0.215, 0.692)	0.667 (0.410, 0.867)
	Specificity (95% CI)	1.000 (0.989, 1.000)	1.000 (0.989, 1.000)

Note: PPV = positive predictive value, NPV = negative predictive value.
[a] 72 h for amphetamine/methamphetamine/PCP and cocaine; 30 days for marijuana.

environmental contamination, low analytical costs, and low refusal rates (Gjerde, Øiestad, & Christophersen, 2011). Second, this study was unable to determine if misreporting was intentional or if some users were unable to recall their drug usage. The primary objective of this research was to determine the rate of agreement between self-report and biological verification, regardless of the reason for misreporting. Third, a biological test was not used to test for alcohol use. Therefore, this study was unable to compare self-reported alcohol use with a biological measure as with the other substances. Fourth, while participants were aware that they would be giving an oral fluid sample, they were not explicitly told that their self-report and biological drug test would be compared for concordance. Finally, participants in this study were given assurances of confidentiality, which undoubtedly influenced the accuracy of their self-reporting. When such assurances are not given, or resources are contingent upon an individual's self-reported use, it seems reasonable that this would result in even higher rates of misreporting.

5. Conclusions

Self-report has some disadvantages when used as the sole measure of drug use among supportive housing residents. If this study had relied exclusively on the self-report method, that would have resulted in an underestimation of true drug use. For instance, this study found that only 44% of biochemically-confirmed cocaine users self-reported their use. However, had this study exclusively relied on oral fluid testing for marijuana, it would have failed to capture an additional 15 users (31.3%) (Fig. 2). Participants who misreported drug use were more likely to be older, non-White, have no medical insurance, and not report any alcohol use. When biochemical verification is not possible, increasing the recall window may help increase the accuracy of self-reported drug use.

In sum, three key lessons may be useful when planning and evaluating an intervention in a similarly vulnerable population.

1) Depending on the drug category, self-report may not adequately capture recent drug use. In our study, using a wider time window of 90 days resulted in higher concordance with the oral fluid drug test, indicating that distant recall was a better measure for current drug use behaviors in this population.
2) Predictors of misreporting can be used to estimate discordance among similar populations when biological verification is not feasible. Participants were more likely to misreport if they were older, non-White, and uninsured. While we may not know the reasons for misreporting, these predictive variables can be used by other research teams to estimate misreporting in a demographically similar population.
3) With regard to marijuana use, exclusively relying on an oral fluid biological test for this population may result in an underestimation of actual drug use.

Author's disclosure

Funding for this study was provided through a Medicaid 1115 Waiver to the State of Texas. Centers for Medicare and Medicaid Services (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.

Authors Rendon, Suzuki, and Walters designed the study. Author Rendon conducted literature searches and provided summaries of previous research studies and conducted the statistical analysis. Author Rendon wrote the first draft of the manuscript and all authors contributed to and have approved the final manuscript.

All authors declare that they have no conflicts of interest.

References

Aiemagno, S. A., Cochran, D., Feucht, T. E., Stephens, R. C., Butts, J. M., & Wolfe, S. A. (1996). Assessing substance abuse treatment needs among the homeless: A telephone-based interactive voice response system. *American Journal of Public Health*, *86*(11), 1626–1628. http://dx.doi.org/10.2105/AJPH.86.11.1626.

Allen, K. R. (2011). Screening for drugs of abuse: Which matrix, oral fluid or urine? *Annals of Clinical Biochemistry*, *48*(6), 531–541. http://dx.doi.org/10.1258/acb.2011.011116.

Burt, M., & Aron, L. Y. (2001). *Helping America's homeless*. Washington, DC: Urban Institute.

Caton, C. L. M., Wilkins, C., & Anderson, J. (2007). People who experience long-term homelessness: Characteristics and interventions. *Toward understanding homelessness: The 2007 National Symposium on homelessness research* (pp. 1–44).

Dams, R., Choo, R. E., Lambert, W. E., Jones, H., & Huestis, M. A. (2007). Oral fluid as an alternative matrix to monitor opiate and cocaine use in substance-abuse treatment patients. *Drug and Alcohol Dependence*, *87*(2–3), 258–267. http://dx.doi.org/10.1016/j.drugalcdep.2006.08.020.

Darke, S. (1998). Self-report among injecting drug users: A review. *Drug and Alcohol Dependence*, *51*(3), 253–269. http://dx.doi.org/10.1016/S0376-8716(98)00028-3.

Dennis, M. L., Funk, R., Godley, S. H., Godley, M. D., & Waldron, H. (2004). Cross-validation of the alcohol and cannabis use measures in the Global Appraisal of Individual Needs (GAIN) and Timeline Followback (TLFB; form 90) among adolescents in substance abuse treatment. *Addiction*, *99*(Suppl. 2), 120–128. http://dx.doi.org/10.1111/j.1360-0443.2004.00859.x.

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.

Drake, R. E., & Wallach, M. A. (1989). Substance abuse among the chronic mentally ill. *Hospital & Community Psychiatry*, *40*(10), 1041–1046 Retrieved from http://www.ncbi.nlm.nih.gov/htbin-post/Entrez/query?db=m&form=6&dopt=r&uid=2807205

Edens, E. L., Mares, A. S., & Rosenheck, R. A. (2011). Chronically homeless women report high rates of substance use problems equivalent to chronically homeless men. *Women's Health Issues*, *21*(5), 383–389. http://dx.doi.org/10.1016/j.whi.2011.03.004.

Edens, E. L., Mares, A. S., Tsai, J., & Rosenheck, R. A. (2011). Does active substance use at housing entry impair outcomes in supported housing for chronically homeless persons? *Psychiatric Services*, *62*(2), 171–178. http://dx.doi.org/10.1176/appi.ps.62.2.171.

Endicott, J., Nee, J., Harrison, W., & Blumenthal, R. (1993). Quality of life enjoyment and satisfaction questionnaire: A new measure. *Psychopharmacology Bulletin*, *29*(2), 321–326. http://dx.doi.org/10.1111/j.1365-2850.2011.01735.x.

Gjerde, H., Øiestad, E. L., & Christophersen, A. S. (2011). Using biological samples in epidemiological research on drugs of abuse. *Norsk Epidemiologi*, *21*(1), 5–14 Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84855764470&partnerID=40&md5=a6fee292da929397c2a05d211c691d1a

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.

Hjorthøj, C. R., Hjorthøj, A. R., & 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. http://dx.doi.org/10.1016/j.addbeh.2011.11.025.

Katz, C. M., Webb, V. J., Gartin, P. R., & Marshall, C. E. (1997). The validity of self-reported marijuana and cocaine use. *Journal of Criminal Justice*, *25*(1), 31–41. http://dx.doi.org/10.1016/S0047-2352(96)00049-9.

Kirst, M., Zerger, S., Misir, V., Hwang, S., & Stergiopoulos, V. (2015). The impact of a housing first randomized controlled trial on substance use problems among homeless individuals with mental illness. *Drug and Alcohol Dependence*, *146*, 24–29. http://dx.doi.org/10.1016/j.drugalcdep.2014.10.019.

Larimer, M. E., Malone, D. K., Garner, M. D., Atkins, D. C., Burlingham, B., Lonczak, H. S., ... Marlatt, G. A. (2009). Health care and public service use and costs before and after provision of housing for chronically homeless persons with severe alcohol problems. *JAMA: The Journal of the American Medical Association*, *301*(13), 1349–1357. http://dx.doi.org/10.1001/jama.2009.414.

Magura, S., & Kang, S. Y. (1996). Validity of self-reported drug use in high risk populations: A meta-analytical review. *Substance Use & Misuse*, *31*(9), 1131–1153. http://dx.doi.org/10.3109/10826089609063969.

McElrath, K., Dunham, R., & Cromwell, P. (1995). Validity of self-reported cocaine and opiate use among arrestees in five cities. *Journal of Criminal Justice*, *23*(6), 531–540. http://dx.doi.org/10.1016/0047-2352(95)00044-5.

McNagny, S. E., & Parker, R. M. (1992). High prevalence of recent cocaine use and the unreliability of patient self-report in an inner-city walk-in clinic. *Journal of the American Medical Association*, *267*(8), 1106–1108. http://dx.doi.org/10.1001/jama.1992.03480080076030.

Mondello, M., & House, S. (2007). Cost analysis of permanent supportive housing state of Maine - Greater Portland. *Mental Health Care*, 1–22 (September).

Napper, L. E., Fisher, D. G., Johnson, M. E., & Wood, M. M. (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. http://dx.doi.org/10.1016/j.addbeh.2009.12.006.

Padgett, D. K. (2006). Housing first services for people who are homeless with co-occurring serious mental illness and substance abuse. *Research on Social Work Practice*, *16*(1), 74–83. http://dx.doi.org/10.1177/1049731505282593.

Polcin, D. L. (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). http://dx.doi.org/10.1105/tpc.105.037978.ln.

Rosay, A. B., Najaka, S. S., & Herz, D. C. (2007). Differences in the validity of self-reported drug use across five factors: Gender, race, age, type of drug, and offense seriousness. *Journal of Quantitative Criminology*, *23*(1), 41–58. http://dx.doi.org/10.1007/s10940-006-9018-7.

Schumacher, J. E., Milby, J. B., Raczynski, J. M., 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 Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8583516

Secades-Villa, R., & Fernandez-Hermida, J. R. (2003). The validity of self-reports in a follow-up study with drug addicts. *Addictive Behaviors*, *28*, 1175–1182 (0306–4603 (Print)) http://dx.doi.org/10.1016/S0306-4603(02)00219-8.

Sloan, J. J. I., Odapati, M. A. R. B., & Ucker, T. A. A. T. (2004). Respondent misreporting of drug use in self-reports: Social desirability and other correlates. *Journal of Drug Issues*, 269–292.

Sobell, L., & Sobell, M. (1992). Timeline follow-back: A technique for assessing self-reported alcohol consumption. *Measuring alcohol consumption psychosocial and biochemical methods* (pp. 41–72). http://dx.doi.org/10.1007/978-1-4612-0357-5_3.

Sobell, L. C., Brown, J., Leo, G. I., & Sobell, M. B. (1996). The reliability of the alcohol Timeline Followback when administered by telephone and by computer. *Drug and Alcohol Dependence*, *42*(1), 49–54. http://dx.doi.org/10.1016/0376-8716(96)01263-X.

The Substance Abuse and Mental Health Services Administration (2010i). *Results from the 2010 National Survey on Drug Use and Health: mental health findings.*

The Substance Abuse and Mental Health Services Administration (2011i). Current statistics on the prevalence and characteristics of people experiencing homelessness in the united states, (July), 1–23. Retrieved from www.samhsa.gov

Tsai, J., Lapidos, A., Rosenheck, R. a., & Harpaz-Rotem, I. (2013). Longitudinal association of therapeutic alliance and clinical outcomes in supported housing for chronically homeless adults. *Community Mental Health Journal*, *49*(4), 438–443. http://dx.doi.org/10.1007/s10597-012-9518-x.

United States Department of Housing and Development (2007). *Defining chronic homelessness: A technical guide for HUD programs (September) , 1–16.*

White, D., Rosenberg, E. S., Cooper, H. L. F., Del Rio, C., Sanchez, T. H., Salazar, L. F., & Sullivan, P. S. (2014). Racial differences in the validity of self-reported drug use among men who have sex with men in Atlanta, GA. *Drug and Alcohol Dependence*, *138*(1), 146–153. http://dx.doi.org/10.1016/j.drugalcdep.2014.02.025.

Wu, L. T., Kouzis, A. C., & Schlenger, W. E. (2003). Substance use, dependence, and service utilization among the US uninsured nonelderly population. *American Journal of Public Health*, *93*(12), 2079–2085. http://dx.doi.org/10.2105/AJPH.93.12.2079.