9-3-2013

Validity of self-reported drug use during pregnancy

Mahek Garg

Mahek Garg

Candidate

College of Pharmacy

Department

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

Dr. Ludmila Bakhireva , Chairperson

Dr. Matthew Borrego

Dr. Ajna Hamidovic

VALIDITY OF SELF-REPORTED DRUG USE DURING PREGNANCY

by

Mahek Garg

Bachelor of Pharmacy

Thesis

Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

PHARMACEUTICAL SCIENCES

The University of New Mexico
Albuquerque, New Mexico

July, 2013

VALIDITY OF SELF-REPORTED DRUG USE DURING PREGNANCY

BY

MAHEK GARG

B.Pharm., Indraprastha University, India, 2010

M.S., Pharmaceutical Sciences, University of New Mexico, 2013

ABSTRACT

Introduction: Epidemiology studies often rely on maternal self-reports for drug use information, however, the degree of drug use under-reporting among pregnant women is largely unknown. The purpose of this study is to assess the accuracy of self-reports for methadone, buprenorphine, opioids (prescription opioids and heroin), marijuana, benzodiazepines, amphetamines/methamphetamines, and cocaine/crack-cocaine in a population of pregnant women.

Methods: Analysis was based on 102 pregnant women enrolled in the 'Biomarkers in Pregnancy Study' (BIPS) cohort at the University of New Mexico. Women attending the UNM Milagro clinic, designated to pregnant women with the current or past history of substance abuse, were enrolled during one of the first prenatal care visits and followed up to term. Self-reported information about drug use was compared with the results of the urine drug screens conducted during the third trimester. Simple kappa and prevalence-and-bias-adjusted kappa coefficients were calculated as measures of agreement. Sensitivity and specificity of self-reports for each drug class were also estimated using

urine toxicology screening as the gold standard. In addition, logistic regression was conducted to evaluate the effect of number of toxicology screens on agreement.

Results: The mean maternal age of the sample was 26.4 ± 4.9 years and included a large proportion of ethnic minority (78% Hispanics/Latina) and socially disadvantaged (51% < less than high school education and 95% Medicaid-insured) pregnant women. On average, these patients had 4.8 ± 3.0 urine drug screens in the third trimester. For methadone-maintenance therapy, there was a perfect agreement between self-reports and urine screens (k and PABAK =1.0, 100% sensitivity and specificity). Simple kappa coefficients for other classes of drugs revealed varied levels of agreement, however, PABAK coefficients indicated moderate to almost perfect agreement for other classes of drugs. Sensitivity of self-reports was low for all classes of drugs, with marijuana and opioids more acceptable than other classes of drugs. The specificity of self-report was high for classes of drugs. Logistic regression revealed no association between number of toxicology screens and agreement.

Discussion: These results indicated that sensitivity of self-reports for all classes of drugs was low with opioids and marijuana more acceptable than other drugs.

Table of Contents

FIGURES

TABLES

CHAPTER 1: INTRODUCTION

1.1. Background

According to the 2011 annual National Survey on Drug Use and Health conducted by the Substance Abuse and Mental Health Services Administration[1], approximately 22.5 million Americans reported using an illicit drug or psychotherapeutic medications such as pain reliever, tranquilizer or stimulant in the month prior to the survey. This number has increased by 8.3 percent from the previous estimates of the year 2002. Results of the survey indicated that marijuana was the most commonly abused illicit drug with an estimated 18.1 million current users in the year 2011. The survey also reported that prescription drug abuse, that is the non-medical use of prescription drugs, was the second most prevalent illicit drug use category and approximately 2.7% of the US population reported taking prescription drugs non-medically. The most commonly abused medications included pain reliever, tranquilizers, stimulants, and sedatives. In addition, high rates of other risky behaviors were also reported among those who abused drugs[1].

Implications of the illicit drug use epidemic and its consequences are significant. It is a serious public health concern that presents tremendous burden on both individuals and society, estimable in terms of morbidity, mortality, untoward health and medical consequences, societal and economic costs associated with the addiction and its consequences. Alcohol and other substances of abuse including marijuana, LSD, heroin, cocaine, tobacco and prescription drugs account for nearly 590,000 deaths in the US every year[2]. Apart from high mortality rates, drug use and addiction is associated with high morbidity causing approximately 40 million illnesses and injuries each year[2]. The

economic burden of substance use is estimated to be greater than $600 billion every year, which includes health care costs, lost productivity and crimes[3,4]. Societal costs measurable in terms of social costs to families and communities include family violence, initiation of substance abuse among children of drug addict parents, crimes, passive exposure, accidents, and divorce.[2]

Substance abuse among pregnant women is most alarming and poses complex management issues for the healthcare system. Data from the National Survey on Drug Use and Health shows that the prevalence of drug use among pregnant women aged 15-44 years was 5.0% averaged for the years 2010 and 2011[5]. The use of illicit drugs among younger pregnant women was substantially higher with 20.9 percent drug users among pregnant women aged 15 to 17, 8.2 percent among pregnant women aged 18 to 25, and 2.2 percent among pregnant women aged 26 to 44 combined for the years 2010 and 2011.

Drug use during pregnancy has serious complications for both the mothers and the exposed fetuses. Various long-term prospective and retrospective studies have shown high risks of morbidity rates in substance abusing pregnant women and their offsprings[6] [7]. In brief, teratogenic agents such as opioids and cocaine are associated with an increased risk of miscarriages, pre-mature births, congenital birth defects, various long-term and short-term health, behavioral and cognitive problems[7,8]. These agents exert their effects directly by either passing through the placenta and interfering with the fetal development or indirectly due to poor and irresponsible maternal behavior and associated postnatal environment.[7-10] The gravity of drug use during pregnancy and its associated consequences make it imperative for researchers to accurately identify all drug use during

pregnancy by assessing the validity of frequently employed drug use methods and ensuring that the methods employed for the same are of high validity.

Methods such as self-report and biochemical analysis of bodily fluids are available to assess maternal consumption of substances during pregnancy[11]. Each method has its own advantages and disadvantages. Based on the availability of resources and research design, one of the two methods can be employed to assess drug use during pregnancy.

Self-report represents a popular and commonly used method of data collection in social and health services research. Self-reports are easy to administer, not only in face-to-face to interviews but also by mail, telephone, or as self-administered questionnaires. Compared to biochemical analysis of bodily fluids and parts, self-reported data is less expensive to collect and at times, it may be the only mode to gather information about subjective research questions, such as perceptions and behaviors of the study population. Self-reports can be used for not only providing information about drug use but also frequency and time-period of drug use, if disclosed by the respondents.

However, information collected by self-reports can be highly unreliable if special emphasis is not given to the factors that can influence the response such as wording of the interviews, expectations of the interviewer, anonymity, use of audio-visual aids in the data collection etc.[12] Moreover, information collected by self-reports are highly susceptible to various validity issues. The concept of validity relates to the question "Are we measuring what we intend to measure" and one of the important type of validity is *criterion validity*, which in this study, refers to the extent to which the subjective self-reported data are verified by agreement with another indicator of the same phenomenon

believed to be of higher validity[13]. One of the factors that can undermine the validity of self-reported drug use during pregnancy is *social desirability bias*, which refers to the unwillingness of the respondents to respond accurately for behaviors that might be disapproved by the society. According to the social desirability theory, the more highly stigmatized and negatively sanctioned a behavior, the stronger the tendency to deny having engaged in that. This outlook indicates that unreliable reports, either underreported or over reported may occur as a result of the perceived acceptability of the correct response [13].

Due to advancements in drug screening techniques, various measures including analysis of bodily fluids such as urine, saliva, meconium (first stool of an infant) and hair assays are available for assessing the validity of self-reported drug use [11]. Precision and accuracy of these drug screening techniques make them ideal comparators for establishing the validity of self-reports by comparing the sensitivity and specificity of self-reports with any of the available above-mentioned objective methods. As legitimate and reasonable concerns regarding the validity of self-reported maternal drug use during pregnancy exist, it is incumbent upon researchers to address these issues. Therefore, the objective of this study is to assess the validity of self-reported drug use during pregnancy using urine toxicology screen results as the criterion.

1.2. Significance

Drug use during pregnancy is associated with various developmental and neurobehavioral dysfunctions among prenatally exposed infants which makes it essential to accurately assess the prevalence of drug use during pregnancy[7]. Timely identification of pregnant

women with substance abuse behavior and exposed infants would help in the delivery of timely clinical and behavioral interventions. In the absence of good quality evidence on the validity of most commonly employed methods in assessing the drug abuse during pregnancy, the interventions might not reach the exposed women and infants. This study would enable the researchers in establishing the validity of self-reported information and would answer if relying on self-reported information for risky behaviors like drug use during pregnancy is adequate or not.

Reliable self-reports can be very helpful in assessing the effectiveness and compliance of various substance abuse treatment and rehabilitative programs as they are easy to obtain and relatively less expensive than the laboratory measures, provided their validity has been established.

Very few studies in the past have compared the validity of self-reported data against laboratory tests but methodological issues such as time-frames, number of biological samples exist. Moreover, there is a dearth of comprehensive data specifically looking at self-reported drug use during pregnancy among high-risk population of pregnant women. Our study aims at comparing the validity of self-reported drug use using urine toxicology screens as the criterion, stratifying results by different drug classes and in a population of pregnant women with current or past history of substance abuse enrolled in a substance abuse treatment clinic. Findings of this study would contribute to the limited knowledge that exists in the literature about the agreement of self-reported drug use during pregnancy with an objective measure.

1.3. Specific Aims and Hypothesis

SPECIFIC AIM I: To estimate the agreement between patient self-reports and urine toxicology screens for the following classes of drugs: marijuana, methadone, prescription opioids and heroin, amphetamines, cocaine, and benzodiazepines in a high-risk population of pregnant women with substance use history.

We hypothesize that the agreement between two measures would be substantial for this population as these women are on opioid maintenance therapy and receiving treatment for their substance abuse behaviors. Therefore, they would be indeed honest in reporting their consumption of drugs during pregnancy.

SPECIFIC AIM II: To compare the validity of self-reports for drug use by comparing the sensitivity of self-report for different classes of drugs.

We hypothesize that the validity of self-report would diminish with increased social undesirability of the drug used during pregnancy.

CHAPTER 2: LITERATURE REVIEW

The chapter will provide an overview of previous studies relevant to this topic and is divided into three sections as follows:

1. Negative Consequences of drug use during pregnancy: This section of the literature review begins with a brief summary on the effects of drug use during pregnancy on fetal development followed by studies that have demonstrated negative consequences of maternal drug use during pregnancy on the neonatal outcomes and later child growth and development. This section will summarize the negative consequences of only those drugs whose self-reports are evaluated for their validity in this study.

2. Validity of self-reports for smoking and alcohol use among pregnant women: This section will summarize the studies that have assessed the accuracy of self-reports for behaviors such as smoking and alcohol use among pregnant women.

3. Validity of self-reports for drug use during pregnancy compared to an objective measure: This section of the literature will summarize studies that have assessed the validity of self-reports for drug use during pregnancy. This section is then further divided according to the criterion used for validation of self-reports. It begins with studies that have evaluated the agreement between self-reports and urinalysis followed by studies using meconium, hair or umbilical cord analysis.

2.1. Negative Consequences of Drug Use during Pregnancy

2.1.1. Teratogenic effects of drugs

The teratogenic effects of drugs vary temporally and the susceptibility of the fetus to various physiologic, teratogenic, and developmental abnormalities depends on the period of exposure, dose of exposure, chronicity of exposure as well as interactional effects of polydrug use during pregnancy[14]. Different organs have different critical periods of vulnerability for malformations during the length of pregnancy, the heart is most sensitive during the third and fourth weeks of gestation, whereas external genitalia during the eighth and ninth weeks. Brain and skeleton are sensitive from the beginning of the third week to the end of pregnancy and in the post birth period. Evidence suggests that early exposure during pregnancy is associated with outcomes that are more deleterious and first 3 months of gestation are most critical for teratogenic malformations. However, changes caused by illicit or licit drugs may even occur later in pregnancy. Hence, timing of exposure during pregnancy has different effects on different organs as exposure to a drug at 24th week of gestation may not have as severe effects on heart as on the brain or the skeleton compared to exposure at early gestation[14].

2.1.2. Drug Use during Pregnancy and Infant Outcomes

COCAINE

The 'Crack baby' phenomenon led cocaine/ crack (crystals of cocaine that can be smoked) cocaine to be the most extensively studied drug of abuse for adverse neonatal outcomes. Studies have demonstrated the association of prenatal cocaine exposure with reduced intrauterine fetal growth leading to low weight, height and head circumference at birth,[6,9,15-25] increased risk of preterm births,[9,23-25] increased prevalence of small-for-gestational age (SGA) infants,[25] placental abruptions and premature rupture of membranes,[9] neurosonographic and morphological abnormalities,[25-27] increased

healthcare utilization,[28] and increased mortality.[6] Prenatal cocaine exposure has also been associated with the disruption of monoaminergic neurotransmitter systems[29] during fetal development affecting attention and behavioral regulation during early childhood. Studies have documented adverse effects of prenatal cocaine exposure on language,[30-41] executive function development such as cognition, attention, memory[42-44] and on the most crucial component inhibitory control, which refers to better emotional regulation, reasoning and ability to maintain efforts toward attainment of goals[45] while ineffective inhibitory control is associated with the development of psychopathology (e.g. ADHD) and various externalizing and internalizing difficulties.[44,46,47]

Despite the results of above mentioned studies, there is little consensus regarding the adverse effects of cocaine use on pregnancy outcomes. A systematic literature review conducted by Franks et. al.[48] reviewed physical growth, cognition, language skills, motor skills, and behavior, attention, affect, and neurophysiology outcomes in early childhood after prenatal cocaine exposure. This review concluded that there was no consistent negative correlation between prenatal cocaine exposure and physical growth, developmental test scores, or receptive or expressive language. The study found no effect of cocaine on behavior scores; however, less optimal motor scores upto the age of 7 months were reported among the exposed. Moreover, an association between prenatal cocaine exposure and decreased attentiveness, emotional expressivity, neurophysiologic and attention was suggested. Lutiger et. al.[49] evaluated reproductive effects of maternal cocaine use in a meta-analysis and concluded that very few adverse effects such as genitourinary tract malformations could be significantly associated with the cocaine use during pregnancy when polydrug users using cocaine were compared with polydrug users

without cocaine use. Comparison of cocaine alone users and no drug users, revealed a higher risk of in-utero deaths, and genitourinary tract malformations. In addition, analysis of head circumference, gestational age, birth weight, and length revealed medium effect size when cocaine users were compared with no drug users. However, when polydrug with cocaine users were compared with polydrug without cocaine users yielded small to non-existent effect size[49].

OPIOIDS

Neonatal abstinence syndrome, NAS, a generalized disorder characterized by signs and symptoms of central nervous system hyperirritability, gastrointestinal dysfunction, respiratory distress and vague autonomic symptoms including yawning, sneezing, mottling, and fever is a serious consequence of maternal opioid use during pregnancy[50]. Prenatally exposed infants become passively addicted to the in-utero exposures and undergo abstinence at birth. Infants born to opioid dependent mothers are at a risk of NAS,[51,52] longer length of stay in neonatal units, increased healthcare costs[51] along with outcomes like increased mortality, low birth weight, preterm births, antepartum hemorrhage.[51,53-59]

The use and abuse of pain relieving medications, opioid analgesics, is increasing[60,61] and their use during pregnancy pose a serious public health challenge. A recent study by Broussard et. al. evaluated the association between maternal opioid use between one month before pregnancy and first trimester with birth defects in 17,449 cases and 6,701 controls. The study found significant association between opioid use during pregnancy and birth defects like conoventricular septal defects, pulmonary valve stenosis,

atrioventricular septal defects, hypoplastic left heart syndrome, spina bifida, or gastroschisis in infants.[62] Previous studies have reported the association of opioid use during pregnancy with orofacial clefts, congenital heart defects (CHD)[63,64] . However, the study conducted by Shaw et. al., on the other hand found no association between opioids and birth defects.[65]

Opioid maintenance therapy (OMT) used in opioid-dependent pregnant women is associated with adverse neonatal outcomes when compared to drug free healthy controls;[53,59] however, studies have found significantly improved neonatal outcomes in women receiving medically supervised doses of methadone or buprenorphine compared to pregnant women abusing illicit opioids during their pregnancy.[51,52,66]

MARIJUANA

According to a review conducted by Kuczkowski[67], although marijuana is not a well-known human teratogen, recent studies suggested subtle negative effects of marijuana use during pregnancy on neurobehavioral outcomes of the exposed infants including sleep disturbances, impaired visual problem-solving, hyperactivity, inattention and increased delinquency. Low neonatal birth weight, increased complication during labor and increased proportion of preterm births were associated with maternal marijuana use during pregnancy.

Similarly, another review conducted by Minnes et. al.[8] stated that marijuana use during pregnancy is not associated with any major fetal growth or physical abnormalities, however, mild withdrawal and poor autonomic control have been observed in the exposed infants. Prenatally exposed children exhibit deficits in reading, spelling, and

higher order thinking including memory, planning, impulsivity, problem solving, and attention. However, there was no overall suppression of IQ in the exposed infants. In terms of long term behavioral and emotional consequences of prenatally exposed infants', the authors concluded that there might be an increased risk of depressive symptoms and adolescent substance abuse.[8,68]

METHAMPHETAMINES

Some studies reported that prenatal methamphetamine exposure was associated with a 3.5 times greater risk of SGA and had lower birth weight than controls.[8] Other adverse outcomes might included lower arousal from sleep, lack of energy, and withdrawal symptoms, cleft palate, cranial abnormalities, fetal growth retardation, and behavioral problems.[8,69]

BENZODIAZEPINES

According to a review conducted by McElhatton et.al.,[70] the information regarding the effect of benzodiazepines during pregnancy is limited and inconsistent. Earlier studies have indicated an increased risk of multiple malformations, including facial clefts, and cardiac malformations. However, later studies found no evidence of increased malformations among infants of benzodiazepines users and had normal postnatal development[70]. Another meta-analysis yielded an odds ratio of 1.07 for their fetal safety. The authors concluded that while benzodiazepines do not increase teratogenic risks in general but there is a two-fold increased risk of clefts with the use of benzodiazepines during pregnancy.[71]

Summary

Undeniably, substance abuse during pregnancy has serious implications for the developing fetus. Various complications like restricted intrauterine growth, congenital birth defects, preterm labor and delivery, placental rupture and abruptions, arise from their use during pregnancy. Moreover, exposed infants are at an increased vulnerability to cognitive and behavioral problems in the later stages of life. The overall adverse effects of drug use are complex, multifactorial, and not well understood. The observed outcomes could be due to a number of behavioral, social, psychological factors like postnatal environment, role of parents, their lifestyle and personality, polysubstance abuse, and nutritional status among others. The adverse outcomes associated with a specific illicit drug are difficult to disentangle from the contributing psychosocial and lifestyle factors. Polydrug exposure, home-environment, and maternal/paternal variables make it difficult to attribute untoward consequences to one exposure.

Large-scale studies, with adequate sample size and control group, in variety of environmental conditions are required in order to fully understand the influences of various substances of abuse on the developmental stages of the exposed infants. In addition, studies assessing the contributing effects of various substances are required in future.

2.2. Validity of Self-Reports for Smoking and Alcohol Use during Pregnancy

Studies assessing the accuracy of self-reports for smoking and alcohol use during pregnancy are limited. The validity of self-reported smoking status during pregnancy was compared with urinary/serum cotinine levels or carbon monoxide levels. Gilligan et. al.[72] analyzed urine samples of women attending an antenatal care clinic for cotinine levels and found that 17% of women positive for cotinine levels had misreported their smoking status as non-smokers. Shipton et. al.[73] concluded that merely relying on self-reported smoking status during pregnancy led to an underestimation of true smoking prevalence by 25% when compared to cotinine concentration in the blood samples. Similarly, Britton et.al.[74] estimated that 34.7% of women with positive urinary cotinine levels denied smoking in their self-reports. Ford et.al.[75] compared self-reported smoking status collected through postal questionnaires with blood samples and smoking status recorded in obstetric clinic retrospectively. The authors recorded underreporting, as cotinine-validated smoking prevalence was 31.3% and 27.7% whereas self-reported prevalences were 19.2% and 15.7% for first and third trimester respectively, and 18.9% from clinic records. Burstyn et. al.[76] estimated sensitivity and specificity of self-reported smoking status using urinary cotinine assays and found poor sensitivity (47%) and high specificity (95%). Gollenberg et. al.[77] assessed the validity of retrospectively reported risky maternal behaviors while trying to get pregnant. Prospective longitudinal data recorded in daily diaries, considered here as the gold standard were compared to self-reported smoking (k=0.43), caffeine (k=0.21), alcohol (k=0.20), and fish consumption (k=0.32). The study found poor to moderate validity of self-reported behaviors.

Despite substantial underreporting observed in the above-mentioned studies, few studies reported moderate agreement between self-reported smoking status and an objective

measure. Klebanoff et.al.[78] reported a high kappa coefficient of 0.83 representing substantial agreement between self-reported smoking and cotinine assays. In this study, 95% self-reported non-smokers and 87% self-reported smokers revealed their accurate smoking status, thus yielding a high correlation between the two measures. Authors of this study concluded that self-reports during pregnancy are sufficiently accurate and little would be gained using biochemical verification. This study utilized self-reported smoking status and serum samples collected 30 years ago. Secker Walker et. al.[79] examined correlations between self-reported smoking, exhaled carbon monoxide (CO) and cotinine levels at first and thirty-sixth week prenatal visits. Correlations between self-reported smoking and CO levels were 0.65 and 0.70 at the two visits whereas for cotinine/creatinine ratio, the correlations were slightly lower, 0.61 and 0.65 for the two visits, respectively indicating high correlation between self-reports and exhaled CO levels.

Rice et.al[80] examined agreement between maternal reports and medical records for a variety of perinatal behaviors. Authors reported good agreement for smoking between maternal reports and medical records (k=0.80), whereas agreement for alcohol use was poor (k=0.17). In another study conducted by Fox et.al[81], authors assessed the reliability of self-reports for smoking habits and alcohol consumption patterns during pregnancy for pregnant women participating in a randomized clinical trial of smoking cessation intervention. This study employed a test-retest design and study participants provided self-reports of their smoking and alcohol consumption prior to 18th week (15.8±3.8 weeks) and then again at 18th week of gestation. Self-reported information was compared with thiocyanate levels in their saliva samples collected during the first interview. In this

study, about half of the subjects gave identical reports of pregnancy smoking habits at test and retest and rest were minor changes. Kappa statistic for smoking status in both the groups was similar, 0.61 in the intervention and 0.56 in the control group. Similarly, for alcohol drinking patterns, kappa coefficient was 0.52 in the intervention group and 0.56 in the control group.

In a study conducted by Hessol et. al.[82], the authors interviewed 350 Latinas above 20 weeks of gestation regarding their alcohol, smoking, and medical conditions. The authors reported low kappa coefficient for self-reported alcohol use and medical records (k=0.35) whereas moderate kappa coefficient for tobacco use (k=0.79). Self-reports had lowest validity for alcohol use whereas moderate validity for tobacco use.

2.3. Validity of Self-Reports for Drug Use during Pregnancy

Limited numbers of studies have assessed the validity of self-reported illicit drug use during pregnancy using toxicology screens as the criterion. The common methods used were urinalysis, hair, meconium, or umbilical cord analysis. In the following section, studies evaluating the agreement between self-report with any of the above-mentioned objective measures are summarized:

2.3.1. Self-Reports and Urinalysis

Six studies have used urine toxicology screens to validate self-reported drug use during pregnancy and majority of them assessed agreement for cocaine use and marijuana.

Marroun et.al.[83] assessed the agreement between self-reported prenatal cannabis use and urinalysis among pregnant women enrolled in a population based birth cohort titled as the Generation R study. Self-reported maternal substance use, that is, alcohol, tobacco, and cannabis was measured using a questionnaire at the time of enrollment in the first trimester of the pregnancy. Only 35.9% of women who reported cannabis use during pregnancy had a positive urine screen. Sensitivity and specificity of self-report compared to urinalysis were 0.36 and 0.99. In this study, maternal self-report was collected at the time of enrollment (usually in the first trimester of pregnancy) and retrospectively assessed cannabis use either before pregnancy or during the last three months from the date of enrollment, whereas urine samples were collected in early, mid, and late pregnancy. No detailed description was provided how urine collection process took place and which of the three urine reports related to the self-reported cannabis use.

Christmas et. al[84]. compared the efficacy of maternal questionnaire with urine toxicology screens for the detection of substance use in 302 pregnant women presenting to a university-based obstetric clinic. Extensive questionnaires including the details of past and current maternal and paternal substance use were administered during the first prenatal visits. Urine samples collected during the study period were analyzed for a number of illicit substances like amphetamines, barbiturates, benzodiazepines, cannabinoids, benzoylecgonine, opiates, methaqualone, phencyclidine, methadone, propoxyphene, nicotine, and ethanol. Only 17 of the 41 patients (41.5%) who tested positive on the urine screens had admitted current substance use during the interview. Only 50% of self-reported current (use within last 30 days) had positive urine toxicology screens. The authors did mention that the questionnaires were administered during the

first prenatal visit. However, no information was provided about the timing of this first prenatal visit. Moreover, information on urine sample collection was ambiguous and no explicit urine collection time period was mentioned.

In another study by Horrigan[85] et. al., the authors compared three different measures a) self-report b) urine screens, and c) Substance Abuse Subtle Screening Inventory (SASSI) in order to determine which combination of measures would yield maximum sensitivity. SASSI used in this study consists of two separate screens: first logically derived 26 face valid items (scored 0 to 3) and second 52 true/false empirically derived items. The 78 items of the SASSI are divided into four clinical subscales: Face Valid Alcohol (FVA), Face Valid Other Drugs (FVOD), Obvious Attributes (OAT), and Subtle Attributes (SAT); two defensiveness subscales: Defensiveness subscale (DEF) and Supplemental Addiction Measure (SAM); and two to three supplementary subscales: Random Answer Pattern [56], Corrections subscale (COR), and Family Problems subscale (FAM), depending on the version of the questionnaire. The authors found that the 54.7% of the sample was positive for drug use by any one method. Self-reports identified 15.2% of users, whereas urinalysis identified 17.3% users and SASSRI yielding highest sensitivity, identified 43.4% users.

Bibb et. al.[86] evaluated the prevalence of illicit substance use and compared drug screening results from maternal interview, meconium, maternal and newborn urine analysis in 580 mother-newborn pairs. Maternal self-reports for the use of tetrahydrocannabinol (THC) was positive in 5.7% delivering mothers, whereas only 2.5% were positive from urine drug analysis and 1% newborns had positive meconium results. Urine screens were less sensitive in the study as samples for biochemical verification and

self-reported data were collected at the same time, which was after delivery. This might not have captured all drug users in urine analysis. For cocaine, interview, maternal urine, and meconium identified equal number of users (3.4%).

Lindsay et. al. [87] determined the accuracy of self-reported cocaine use in an urban sample of 5200 pregnant women. In this large sample of pregnant women, only 5% tested positive for cocaine use and 47% of the women with positive urine test acknowledged drug use during the interviews, thus revealing a poor correlation between self-reported cocaine use and the results of urine assays for cocaine metabolites among prenatal care seeking pregnant women.

Yonkers et. al.[88] compared self-reported marijuana or cocaine use in168 women enrolled in an integrated obstetrical/substance abuse treatment program. They found good agreement between the urine screens and self-reported use of cocaine and marijuana in this population, k= 0.74 for marijuana and k=0.70 for cocaine respectively. The good agreement between the measures could be attributed to the time frame of analysis as this study captured drug use in past month from the interview and urine collection process.

2.3.2. Self-Report and Umbilical Cord Analysis

One study conducted by Wright et. al.[89] evaluated the agreement between umbilical cord analysis and maternal self-reports for cotinine and drug levels. The commonly reported drugs used during pregnancy were methamphetamines, marijuana, and cocaine. The authors found fair agreement between maternal smoking reports and cotinine levels (kappa = 0.26 (0.07–0.5)) and poor agreement for self-reported drug use and positive

drug tests (kappa = 0.19 (−0.05–0.4)). Sensitivity of positive cord illicit drug levels was 32% and specificity was 85% compared with maternal self-report.

2.3.3. Self-Reports and Meconium Analysis

Four studies compared self-reported maternal drug use with meconium analysis of the newborns and found highly discordant results. Ostrea et. al.[90] compared the sensitivity and specificity of maternal interview, maternal hair analysis and meconium for cocaine, opiate and cannabinoid use during pregnancy. This study showed that the interviews had lowest sensitivity in detecting cocaine and opioid exposure (65% and 67%) but highest sensitivity for cannabinoid exposures (58%) when compared to combined hair and meconium analysis results. Hair analysis had a sensitivity of 100% for cocaine, 80% for opiate, and 21% for cannabinoid detection when compared with combined interview and meconium results. Meconium, on the other hand, had 87%, 77%, and 22.7% sensitivity for maternal cocaine, opiates, and cannabis use, respectively when compared with maternal interviews and hair analysis results. Sensitivity of self-report was moderate in this study as these women knew that toxicology screens would confirm their self-reported drug use.

Tassiopoulos et. al.[91] compared self-reported prenatal substance use in a cohort of 480 HIV-infected women and their children with meconium analysis. Meconium samples were available for 264 infants. Sensitivity of self-report was 80% for marijuana and 67% for cocaine in this population using meconium analysis as gold standard. For a non-random subset of mothers/infants with urine/blood tests, higher discordance between self-report and urine/blood toxicology was observed for cocaine, marijuana and opiates.

Gray et. al.[92] identified prenatal amphetamine exposure by maternal interviews and meconium analysis in 3705 participants of the Infant Development, Environment and Lifestyle (IDEAL) study. Based on the combination of maternal self-report and meconium results, 5.7% of the neonates were amphetamine exposed; maternal self-reports identified 71% of the exposed, self-report and meconium analysis identified 25.2 % whereas meconium only identified 3.8% of the exposed infants. This study found interesting results as maternal self-report was more sensitive in identifying drug exposure compared to the meconium analysis. The reasonable explanation provided by authors for these results was that the majority of this population had ceased their drug use in first or second trimester, while meconium starts forming after 20 weeks of gestation.

Lozano et. al.[93] estimated the prevalence of in-utero cannabis exposure in 974 mother-infants dyads by meconium analysis. Prenatal cannabis exposure was found in 5.3% infants whereas only 1.7% mothers self-reported their use.

2.3.4. Self-Report and Hair Analysis

Bessa et. al[94] assessed the validity of self-reported marijuana and/or cocaine use in the third trimester in pregnant adolescent women using hair analysis. Hair analysis results were positive in 6% of that study population: 4% for marijuana, 1.7 % for cocaine, and 0.3% for both. It was interesting to document that none of these patients had reported their drug use in the interviews.

Grant et. al.[95] compared 405 maternal postpartum hair samples with a structured maternal interview conducted postpartum for the detection of cocaine use during pregnancy. Cocaine or its metabolites were identified in 87% of women who reported using cocaine

at least once during pregnancy. Among women who denied cocaine use in pregnancy, 14% had a positive hair test.

Summary

There is sufficient evidence supporting the notion that accuracy of self-reports for behaviors like smoking, alcohol use, and illicit drug use during pregnancy is questionable. Obtaining accurate information about substance abuse during pregnancy poses a major challenge for healthcare providers and researchers working in this field. Self-reporting, though convenient and easy to obtain, has its own limitations. Social stigma, fear of losing the child, legal and social consequences of substance abuse during pregnancy make it difficult to gather reliable information and hence appropriate interventions may not reach the exposed infants and pregnant mothers.

In addition, accuracy of self-report may vary depending of the degree of social desirability bias associated with each substance. As evident from the studies summarized above, greater the social undesirability of a behavior, lower is the agreement between self-reports and toxicology screens. Studies that have evaluated various risky behaviors in the same population demonstrated lowest agreement for drug use followed by agreement for alcohol and then tobacco.

CHAPTER 3: METHODS

This chapter provides a detailed description of study design, methodology, and statistical analysis.

3.1. Study Design and Population

This retrospective cohort study assessed the validity of patient self-reports by comparing self-reported drug use during pregnancy provided by pregnant women against the results of urine toxicology screens. Information provided by women enrolled in the study regarding the use of marijuana, methadone, buprenorphine, prescription opioids and heroin, amphetamines, cocaine, and benzodiazepines during a structured interview conducted after delivery were compared with the results of the urine toxicology screens conducted during the third trimester of pregnancy. For the purpose of this analysis, urine toxicology screens were the criterion or gold standard, both terms used interchangeably.

Data collected from 102 pregnant women enrolled in the 'Biomarkers in Pregnancy Study' (BIPS) at the University of New Mexico was utilized for this analysis. In brief, BIPS is a prospective cohort study designed to assess the validity of several conventional and novel ethanol biomarkers for accurate confirmation of alcohol exposure during pregnancy. The study is approved by the University of New Mexico (UNM) Human Research Review Committee (HRRC).

Pregnant women seeking prenatal care at the Milagro clinic, a specialized UNM clinic for pregnant women with current or past substance use history, were enrolled for BIPS during one of their prenatal visits by a bilingual study coordinator. To be eligible for enrollment in this study, women had to give consent in either English or Spanish, be at

least 18 years old, have an ultrasound confirmed singleton pregnancy, and be less than 35 weeks gestation at enrollment. Informed consent was obtained from all the eligible and participating women. Detailed methodology of the parent study has been described previously.[96]

Since these patients were recruited from a specialized clinic, majority of them were recreational drug users. The vast majority were on opioid maintenance therapy (OMT) receiving standardized doses of either methadone or buprenorphine.

During the BIPS study, two interviews were administered to the enrolled patients to capture their general demographics, lifestyle habits, medical and reproductive history, substance use including alcohol consumption during periconceptional and pregnancy period, maternal and paternal smoking habits and illicit drug abuse. Baseline interviews were conducted at the time of enrollment (mostly in the second trimester of pregnancy) and then followed until labor and delivery. Follow-up interviews were conducted during the hospital stay after delivery. Substance use information during the period between baseline interview and delivery is collected in this second interview. Along with the interviews, maternal biological samples including urine, serum, and whole blood were collected at both baseline and follow-up visits. Repeated urine samples were collected from these women during their treatment at the clinic and were analyzed for the metabolites of various illicit and licit drugs at the TriCore Reference Laboratory (Albuquerque, New Mexico).3.2. Measurements

3.2.1. Demographics and Lifestyle Characteristics

Various sociodemographic characteristics included in this study were as follows: maternal age (continuous variable), marital status (single/never married, married/living with spouse, not married/living with partner, separated from spouse, divorced, widowed), maternal ethnicity (Hispanic/Latina), race (White, Black or African American, American Indian or Alaskan Native, Asian or Asian American Islander), education level (less than high school grad, high school grad/GED, some college/vocational/ college degree, masters/doctorate or professional degree), employment status (employed or not), maternal insurance status (no insurance, employer-based, self-purchased, Medicaid, other public insurances), place of birth (U.S. born or not), language mostly used at home (English, Spanish or other).

Variables that ascertained maternal medical and reproductive history were as follows: gestational age, pre-pregnancy weight, height, and BMI, presence of chronic conditions (hypertension, diabetes, depression, anxiety, seizure disorder, migraine, rheumatoid arthritis, thyroid, asthma/allergies, cancer, heart disease, hepatitis, liver/chronic biliary conditions, tuberculosis, pancreatic disease or other). Additional questions on maternal reproductive health were if this was a planned pregnancy (yes, not now, not any time), parity (number of live-born children), gravidity (number of pregnancies including the current one), history of miscarriage, still-born birth, terminations, ectopic pregnancy, and any complications in the current pregnancy (bleeding, high blood pressure, diabetes, other), use of medications and prenatal vitamins was also ascertained in the interviews.

Lifestyle characteristics assessed in the interviews included smoking habits, alcohol use, and illicit drug use. Women were asked to report their and their partners smoking status, if they are current smokers, were past smokers and if they quit smoking before pregnancy or after pregnancy and number of packets smoked daily.

Considering the objectives of the parent study, very comprehensive drinking information was obtained from these enrolled women. Drinking habits in the periconceptional period and during pregnancy were recorded by collecting information on binge drinking, frequency of drinking, types of drinks during pregnancy, 'high' (number of drinks it takes a person to feel high), and 'hold' (number of drinks a person can hold before passing out or falling asleep) versions of the tolerance questions., Maternal biomarkers such as serum gamma-glutamyltranspeptidase (GGT) and carbohydrate-deficient transferrin (CDT), urine ethyl glucoronide (EtG) and ethyl sulfate (EtS), whole blood phosphatidylethanol (PEth) and newborn PEth were analyzed in the samples both at baseline and follow-up to ascertain alcohol consumption. Based on maternal self-reports and biomarker results, the study sample was divided into a) Abstainers: self-reported abstainers with negative biomarkers b) Early pregnancy exposure: Self-reported exposed with negative biomarkers at delivery or self-reported controls with positive biomarker at enrollment c) Chronic or late pregnancy users: women with positive biomarkers at the delivery.

3.2.2. Self-Reported Maternal Drug use

For the purpose of this analysis, information on maternal illicit drug use in the period between enrollment and delivery is used. Structured follow-up interviews conducted

post-delivery during the hospital stay captured maternal drug use in the period between baseline and follow-up interviews. The responses from the follow-up interview formed the basis of this analysis.

Follow-up interviews explicitly inquired the use of following classes of drugs: marijuana, cocaine, crack-cocaine, heroin, methamphetamines, amphetamines, methadone, buprenorphine, ecstasy, inhalants, prescription opiates, and benzodiazepines.

Enrolled pregnant patients were also asked to report the frequency of drug consumption as one of the following categories: no use, occasional use (less than monthly), once a month, once every 2 to 3 weeks, once a week, and almost every day.

Data Modifications:

Drug class modifications:

Few drug use variables employed separately in the interview that capture information about the drugs of same classes were grouped together in the analysis. Cocaine and crack-cocaine were grouped together as cocaine; methamphetamines and amphetamines were categorized together because methamphetamines are a part of amphetamine screening. Similarly, heroin and prescription opiates were grouped together as they are tested under the same toxicology screen.

Frequency of use modifications:

Follow-up interviews explicitly captured drug use frequency among one of the following categories: no use, occasional use, once a month, once every 2 to 3

weeks, once a week, and almost every day. As we cannot distinguish between different drug use frequencies in qualitative urine screenings, self-reported drug use were dichotomized as ‘no use’ and ‘any use’. Patients reporting occasional, once a month, once every 2 to 3 weeks, once a week or almost every day were categorized together as users. Hence, enrolled patients were categorized as either non-users or users during the data analysis.

3.2.3. Urine Collection and Analysis

As mentioned previously, this study pertains to women attending the UNM Milagro clinic, a specialized clinic for women with current or past history of substance abuse. The majority of these women are on opioid maintenance therapy (OMT) and receive either methadone or buprenorphine as therapy for their addiction since regulated and properly administered OMT helps in improved outcomes, better prenatal and postnatal care. As part of the protocol at the clinic, repeated urine samples are collected from these patients during their prenatal visits and are analyzed at the TriCore Reference Laboratory for the metabolites of commonly abused drugs. The results of the urine toxicology screens for these patients are then entered in their electronic medical records at the clinic. These repeated samples help in assessing the severity of addiction by prevalent co-exposures as well as the compliance to the maintenance therapy.

'URINES DRUG of ABUSE SCREEN by MEDTOX® PROFILE-II: ER 12PANEL'[97] was used to confirm the presence of drugs or their metabolites in the urine samples of the enrolled women. MedTox Profile ER 12 Panel is a one-step colloidal metal

immunochromatographic test for rapid and qualitative detection of twelve most common classes of drugs of abuse or their metabolites in urine. Cut-off concentration for drug detection in the urine samples and time windows for the various classes of drugs are presented in Table 1. Results of these urine toxicology screens are then abstracted from the electronic medical records of these patients.

Table 1: Drug Detection Concentrations and Time Windows

Drug	**Cut-Off Concentration**	**Detection Time Window**
CANNABINOIDS	50 ng/mL	3-7 days/ 30 days for chronic users
MORPHINE	300 ng/mL	2-4 days
AMPHETAMINES	1000 ng/mL	1-5 days
METHAMPHETAMINES	1000-1500 ng/mL	3-5 days
COCAINE	300 ng/mL	2-5 days
METHADONE	300 ng/mL	3 days
BENZODIAZEPINES	300 ng/mL	7 days
OXYCODONE	100 ng/mL	1-4 days

Modifications:

Results of repeated urine toxicology screens were combined together in order to classify these patients either as drug users or non-users. A patient was classified as a non-user for

a particular drug if all of her toxicology screens from baseline to follow-up interview were negative for the presence of drug metabolites for that specific class of the drug. Similarly, users were defined by any positive urine toxicology screens for the metabolites of the specific class of drug in the specified time frame as illustrated in the Table 2.

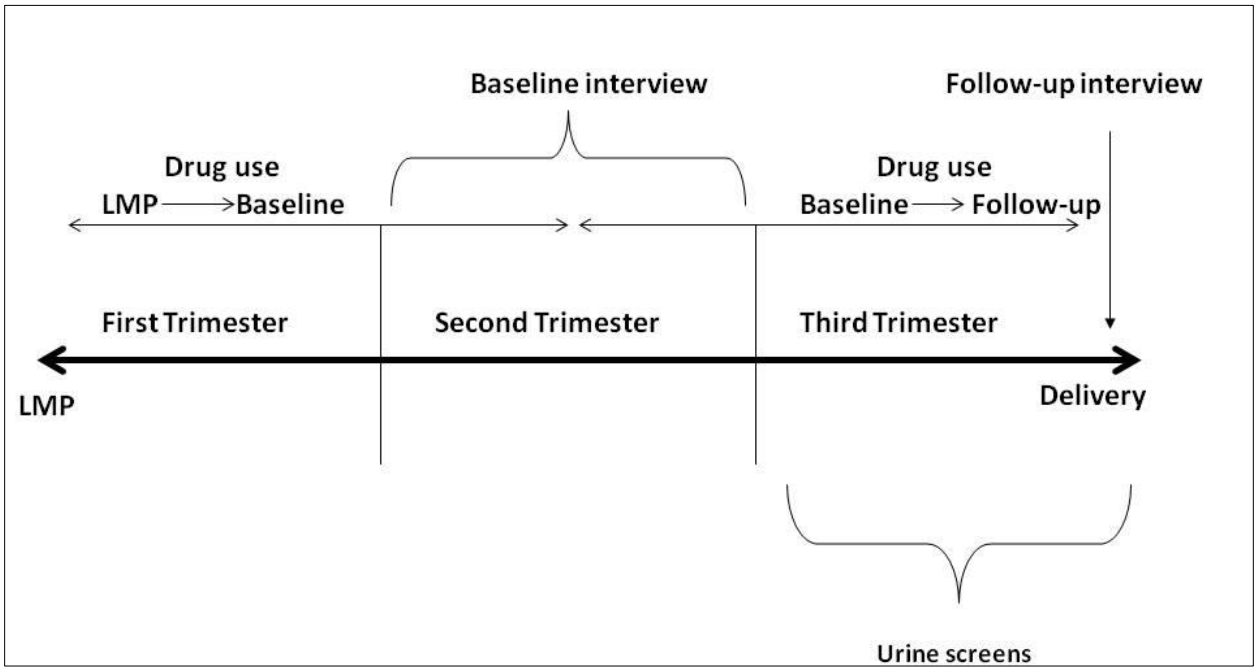
Table 2: Urine Toxicology Screen Outcomes and Interpretation

Toxicology Screen 1	**Toxicology Screen 2**	**Toxicology Screen 3**	**Final Interpretation**
Negative	Negative	Negative	Negative
Negative	Positive	Negative	Positive
Positive	Positive	Negative	Positive
Positive	Positive	Positive	Positive

3.3. Graphical Representation of the Study Frame

As depicted in the Figure 1, majority of the baseline interviews were conducted in the second trimester of the pregnancy. The baseline interview captured drug use since the last menstrual period (LMP). Follow-up interview conducted after delivery captured drug use since the baseline interview. Repeated urine screens that were conducted during the third trimester would be compared to the drug use information provided in the follow-up interview.

Figure 1: Study time frame



3.4. Data Analysis and Outcomes Measures

3.4.1. Agreement between self-reported drug use and urine toxicology screens

The primary outcome of this study, that is the level of agreement between self-reported drug use and urine toxicology screen results is evaluated by comparing self-reported drug use in the period between the baseline and follow-up interviews with the urine toxicology screens conducted in the third trimester of the pregnancy using kappa statistic.

Sensitivity and specificity of self-reports for each class of drug is calculated using urine toxicology screens as the 'gold standard'. Agreement for the purpose of this analysis is represented in Table 3.

Table 3: Representation of Outcome

SELF- REPORT	**URINE SCREENS**	**OUTCOME**
YES	YES	AGREEMENT
NO	NO	AGREEMENT
NO	YES	DISAGREEMENT
YES	NO	DISAGREEMENT

3.4.2. Prevalence of Drug Use in the Milagro Population

Prevalence of use of cannabinoids, amphetamines, benzodiazepines, methadone, buprenorphine, opioids including heroin was also estimated in the population attending the Milagro clinic. Prevalence was estimated using self-reports and positive urine toxicology screen results independently and then total prevalence was also estimated using either measure, that is, all those who either reported drug use in their interviews or were positive for drug use in the toxicology screens were considered as users for the purpose of total prevalence estimation.

3.4.3. Effect of Number of Urine Toxicology Screens on Class-Specific Agreement

Logistic regression was also performed to evaluate the effect of number of urine toxicology screens on the class-specific agreement. Agreement for the purpose of regression was defined as a binary variable and was represented as positive if both urine toxicology screens and self-reports were positive and negative if both urine toxicology screens and self-reports were negative for the same class of drug for the same patient.

3.4.4. Kappa Statistic [98]

The agreement or concordance between self-reported drug use and positive urine toxicology screens during pregnancy would be assessed using the kappa statistic. It is a common measure of precision (reliability) between different observers that takes agreement occurring merely by chance into account[98,99]. Kappa statistic applies to both objective measures like radiographs, urine toxicology screens and to subjective measurements like self-reporting. Comparing self-reported drug use with the presence of drug metabolites in urine will assess the validity of self-reported drug consumption during pregnancy.

Calculations:

Kappa coefficient is based on the difference between observed agreement and how much agreement would be expected to be present by chance alone. As shown in figure 2, a 2X2 table is used to calculate kappa coefficient.

In the figure below, cells (a) and (d) represent the number of cases when there was agreement between the two measures, that is, urine toxicology screens (gold standard) and self-report while cells (b) and (d) represent the number of times when the two measures disagree for the outcome of interest. Now it is possible that the two measures

might sometimes agree just by chance. Kappa provides a numerical rating of the degree to which this occurs and it is based on the difference between the observed and expected (by chance alone) agreement. The observed agreement is the percentage of all the cases when the different measures agree, in the example above, it can be represented by the sum of (a) and (d) divided by total cases, N.

Kappa statistic measures differences in the observed agreement from the expected agreement and its value is standardized to lie on a scale of -1 to 1.

Figure 2: Calculation of kappa statistic

		Urine Toxicology Screens		
		Positive (+)	*Negative* (-)	Total
Self-Report	*Yes*	a	b	a+b
	No	c	d	c+d
	Total	a+c	b+d	

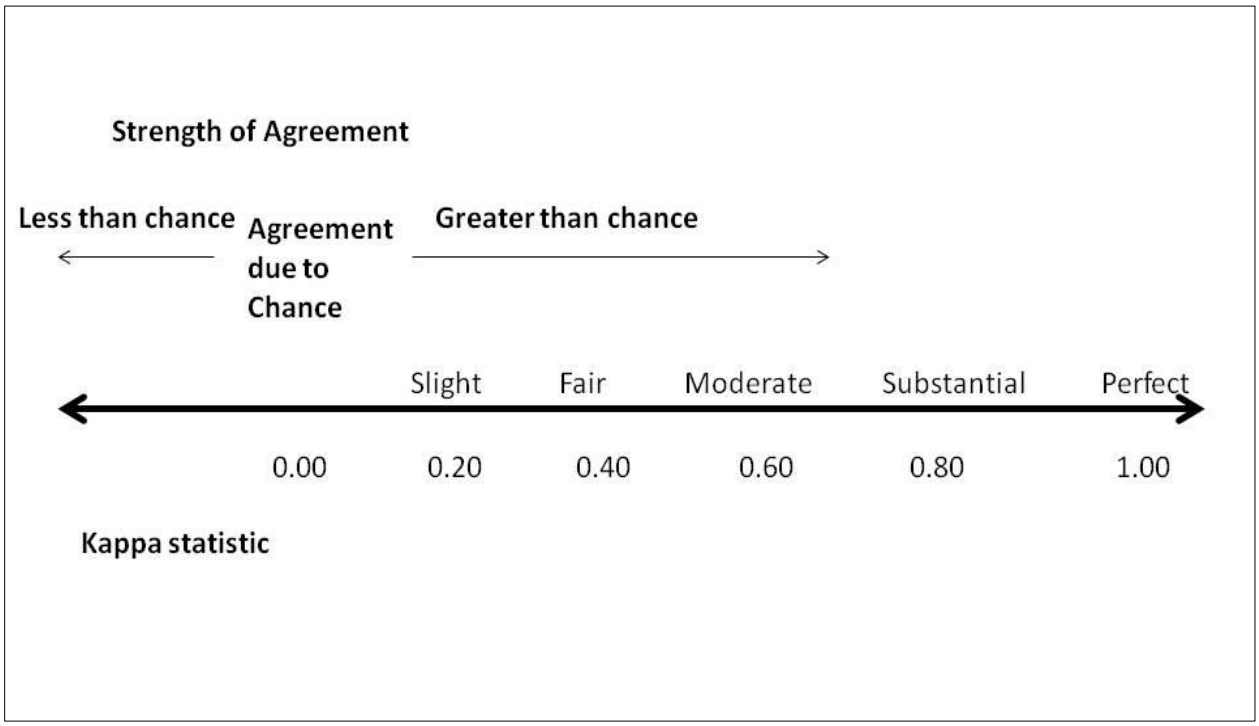
$$\text{Kappa statistic (k)} = \frac{\text{observed agreement } (P_o) - \text{expected agreement } (P_e)}{\text{1-expected agreement}}$$

where, observed agreement, $P_o = (a+d)/N$

expected agreement, $P_e = \{(a+c)(a+b) + (b+d)(c+d)\}/ N^2$

If there were no agreements between the urine toxicology screens and self-report, then (a) and (d) would be 0 and hence observed agreement would be 0. On the other hand, if there were no disagreements, cells (b) and (c) would be 0 and the observed agreement is 1.

Figure 3: Interpretation of kappa



Different authors have suggested several classifications for interpreting the value of kappa statistic. For the purpose of this analysis, we will be using the scale proposed by Landis and Koch[98]. According to it, a perfect agreement would equate to a kappa statistic of 1 whereas any value between 0.81 and 0.99 would indicate almost perfect agreement. Similarly, kappa values between 0.41-0.60 indicate moderate agreement and values between 0.61-0.80 indicate substantial agreement. Values from 0.20-0.40 indicate fair

agreement and any value less than 0.20 is slight agreement and value of 0 represents agreement due to chance and negative values, on the other hand, indicate agreement less than chance. Various kappa values and their interpretation in tabular form are presented in Table 4.

Table 4: Interpretation of Kappa statistic

Kappa value	Interpretation
1.00	Perfect agreement
0.81-0.99	Almost perfect
0.61-0.80	Substantial agreement
0.41-0.60	Moderate agreement
0.21-0.40	Fair agreement
0.0-0.20	Slight agreement
>0.0	Agreement less than chance

3.4.5. Prevalence-and-Bias-Adjusted Kappa (PABAK)

Although kappa coefficient is a widely used statistic to measure agreement between raters, it suffers from certain limitations, most importantly its dependence on prevalence. When the prevalence of a rating in the population is very high or low, the value of kappa

may indicate poor reliability even with a high-observed proportion of agreement. For example, it may signify low kappa even when the proportion of observed agreement is relatively high or vice versa. This interesting phenomenon where the observed proportion of agreement is high but the value of kappa statistic is low, is known as kappa paradox[100,101]. If this paradox is present in the data and only simple kappa coefficients are used to interpret the level of agreements, this may give misleading results. In order to overcome this limitation of the kappa coefficient, another statistic PABAK is proposed which adjusts the kappa for imbalances caused by differences in the prevalence and bias [101].

The effect of prevalence can be assessed by estimating prevalence index that is calculated as:

$$P.I. = | a-d | / N$$

If the prevalence index is high (ie, the prevalence of a positive rating is either very high or very low), chance agreement is also high and kappa is decreased accordingly[100,101].

Bias index is the extent to which raters disagree on the proportion of positive (or negative) cases and is represented as [100]:

$$B.I. = | b-c | / N$$

Kappa is higher when the bias index is high. In contrast to prevalence, the effect of bias is greater when kappa is small than when it is large. As with prevalence, the magnitude of kappa should be interpreted in light of the bias index. [100,101].

PABAK adjusts for high or low prevalence by substituting the actual values of cells a and d with their average values. Similarly, for bias, values of the cells b and c are substituted with their average values.

Although, there is as such no criterion for the use of PABAK, however, when the prevalence and/or bias index is high in the observations, then PABAK (along with simple kappa, prevalence index, bias index, observed agreement and expected agreement) should be reported.

3.4.6. Sensitivity and Specificity[102]

In addition to the measures of agreement, the validity of self-reports is also ascertained by calculating sensitivity and specificity of self-reports for the above-mentioned classes of drugs. They refer to the extent to which self-report measures what it is supposed to measure, in other words, they ensure the accuracy of self-reports using urine toxicology screens as the 'gold standard'. We would like to address that although there is no true gold standard, we will be using urine toxicology screens as the gold standard since it is the currently best available method of diagnosing and we will be validating self-reports using them. The figure below illustrates the concept of sensitivity and specificity:

Figure 4: Sensitivity and Specificity

		Urine Toxicology Screens (Gold Standard)		
		Positive (+)	*Negative* (-)	Total
Self-Report	*Yes*	True Positives	False Positives	a+b

	(a)	(b)	
No	False Negatives (c)	True Negatives (d)	c+d
Total	a+c	b+d	

Where, True postitive: the patient who has the disease and the test is positive

False positive: the patient who does not have the disease but the test is positive

True negative: the patient does not have the disease and the test is negative

False negative: the patient has the disease but the test is negative.

Sensitivity of a test is the ability of that test to correctly identify patients with the outcome of interest. In other words, it is a measure of how likely it is for a test to pick up the presence of the outcome of interest in a person who has it. It is calculated as the proportion of true positives and the sum of true positives and false negatives.

$$\text{Sensitivity} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}}$$

Specificity, on the other hand, is the ability of a test to correctly identify those patients without the outcome of interest. It is calculated as the proportion of true negatives and the sum of true negatives and false positives.

$$\text{Specificity} = \frac{\text{True negatives}}{\text{True negatives} + \text{False positives}}$$

A test with 100% sensitivity correctly identifies all patients with the outcome of interest. A test with 75% sensitivity will detect 75% of patients with the outcome of interest (true

positives) but 25% with the outcome of interest will go undetected. Similarly, a test with 100% specificity accurately identifies all patients without the outcome of interest and a test with 75% specificity correctly reports 75% of patients without the outcome of interest as test negative (true negative) but 25% patients without the outcome of interest are incorrectly identified as test positives (false positives).

Assessing sensitivity and specificity of self-report for various classes of drugs will help in demonstrating that self-report is sensitive and specific and detects differences believed to exist between groups of patients. We will be more confident that it is valid and measuring what we believe it to measuring.

3.5. Power Analysis

The power analysis for this study was conducted in PASS software (Kaysville, Utah) for comparing two independent proportions. For this analysis, alpha value that is the probability of rejecting a true null hypothesis was set to be 0.05, and 0.2 for beta which is the probability of accepting a false null hypothesis. Sample allocation was set at one, as it was assumed that the sample size in both the groups would be same. For the purpose of this power calculation, sensitivity of self-report for methadone was assumed to 99% as women enrolled in this study are on opioid maintenance therapy and it was assumed that self-reporting for methadone would be almost perfect. Power calculations are presented for different effect sizes by varying the sensitivity of self-reporting for other classes of drugs by 5% as difference in the sensitivity of self-report of methadone and other drug classes is unknown.

Table 5: Sample size calculations for different effect sizes

Power	Required sample size	Sensitivity of another drug class	Sensitivity of methadone
0.85	22	0.50	0.99
0.82	24	0.55	0.99
0.83	28	0.60	0.99
0.81	32	0.65	0.99
0.80	38	0.70	0.99
0.81	48	0.75	0.99
0.80	62	0.80	0.99
0.81	90	0.85	0.99
0.80	178	0.90	0.99

Assuming the sensitivity of self-report for methadone to be 0.99, the required sample size might vary form 22-178 depending on the detected differences between the sensitivities of self-report for methadone and another drug class. A sample size of 22-32 patients will achieve >80% power to detect a difference between the group proportion of 0.49- 0.34 using two sided Z test at alpha level 0.05.

Chapter 4: RESULTS

This chapter summarizes the results of the study. A description of the study population's socio-demographic, lifestyle characteristics along with medical and reproductive history of the study population is presented. In the subsequent section, kappa and PABAK for different classes of drugs are estimated and then in the end, sensitivity and specificity of self-report for different classes of drugs are reported.

4.1. Description of the Study Population

Table 6 and 7 present descriptive characteristics of the study population including socio-demographic, lifestyle and medical and reproductive history variables.

4.1.1. Demographic Characteristics

The maternal age of the study sample ranged from 18 to 41 years with a mean of 26.4 ± 4.9 years. Majority of the patients were enrolled during the second trimester of their pregnancy represented by a mean gestational age of 20.9 ± 7.9 weeks at enrollment. More than three-fourths of the study population was comprised of ethnic minorities with 77.5% Hispanic/Latina women. Majority of these women were White (86%) with much lower proportion of American Indian (8.9%) and Black or African American (4%) in the sample. Half of the study population was single/never married (54%), 38% were married/not- married and living with a partner, and 8% of the patients were either divorced, separated or widowed. In addition, approximately half of the enrolled population was less than high school graduate (51.5%) and only 11.8% were currently employed. Majority of these women were on either Medicaid or other public insurance (95%) with 1% reporting an employer-based or self-purchased insurance and 4%

reporting no insurance coverage. This population majorly comprised of U.S.-born women (98%) and 97% reported English as their as their primary language.

4.1.2. Lifestyle Characteristics

The lifestyle characteristics of the enrolled population are also presented in Table 6. Approximately 62% women reported being current smokers at the time of the interview and 64% reported that their partners smoked. Approximately 80% women reported smoking sometime during pregnancy and this number included both current smokers and those who quit after realizing that they were pregnancy. A total of 20.5% women reported no smoking during pregnancy, including non-smokers and those who quit before pregnancy. Alcohol consumption information is also presented in the same table.

Based on both self-reported and biomarker measures, 45% of the sample abstained from alcohol during pregnancy, 35% had early pregnancy alcohol exposure, and 20% were chronic or late pregnancy alcohol users.

4.1.3. Medical and Reproductive History

The maternal medical and reproductive history of the enrolled sample is presented in Table 7. More than a third of the study sample (37%) reported the presence of a chronic condition(s), which was defined as a medical condition that requires repeated, ongoing, or occasional treatment. The most common chronic conditions in this population included anxiety (19%), depression (16%), hepatitis (15%), migraine (5%), and asthma (4%). Other complications reported during pregnancy included bleeding (8%), high blood pressure (2%), and other complications (6%) including nausea, placenta abnormalities,

and fetal growth restriction. More than half of the study population (59%) reported history of an adverse perinatal outcomes including stillbirth (3%), termination (26%), miscarriage (46%), or ectopic pregnancy (5%). Approximately 52% of the patients had their BMI in the healthy range (18.0 to 24.9) with 138.3±30.6 pounds mean pre-pregnancy weight and 63.5 ±4.0 cm mean height. About 88% of all women reported using multivitamin and iron supplements during their pregnancy. For one fourth of the study population (25%) this was their first pregnancy and 85% patients reported their current pregnancy to be unplanned.

4.2. Prevalence of Illicit Drug Use and Opioid-Maintenance Therapy

The prevalence of various drugs of abuse were estimated using both measures (self-report and urine toxicology screens) separately and then together using either measure. The results are presented in Table 8. Methadone and buprenorphine were the most prevalent with 63.85% and 62.07% prevalence, respectively. They were followed by marijuana and cocaine with 25.3% and 24.1% of the enrolled women reporting their use, respectively. Amphetamine and benzodiazepines were the least prevalent drugs with 9.5% and 6.0% prevalence. Self-reported drug use prevalence was lower than the total prevalence captured by either measure for all classes of drugs except methadone and benzodiazepines.

In addition to prevalence, frequency of drug use stratified as no use, occasional use, once a month use, once every 2 to 3 weeks use, once a week, and almost everyday use is presented in Table 9 according to the drug classes. While majority of the study population reported no drug use for all the classes (>80%) except for methadone (35%).

Methadone and buprenorphine were used almost every day by maximum proportion of women (64% and 17.2% respectively) compared to other drug class. Inhalants and ecstasy were almost non-prevalent in this population. Cocaine, opioids, and benzodiazepines had users in all the categories of the frequency.

4.3. Effect of Number of Urine Toxicology Screens on Drug Class-Specific Agreement

The mean number of urine toxicology screens conducted during the third trimester of pregnancy in the study population was 4.8±3.0 per patient (range: 1-14).

4.4. Measures of Agreement and Validity

For measures of agreement, simple kappa coefficients with 95% confidence intervals, prevalence-adjusted-bias-adjusted kappa (PABAK) coefficients, and proportion of observed and expected agreement along with prevalence index and bias index were estimated according to the drug classes and are presented in Table 10.

Methadone: As majority of these patients were on opioid maintenance therapy receiving supervised doses of either methadone or buprenorphine for their addiction, a perfect agreement was observed between self-report and urine toxicology screens for methadone use during pregnancy (kappa and PABAK=1).

Buprenorphine: Being the alternative drug for opioid maintenance therapy, buprenorphine was the second most prevalent drug in this population. The simple kappa coefficient value for agreement between self-report and urine toxicology screens for buprenorphine use during pregnancy was 0.79 (0.57; 1.00) indicating substantial agreement. After adjusting for prevalence and bias index, the PABAK value remained

same, 0.79 signifying substantial agreement between the two measures after adjusting for prevalence. In contrast to an expected proportion of agreement equaling 0.50, the observed proportion of agreement was 0.90 with 0.91 and 0.88 proportion of positive and negative agreement respectively. The prevalence and bias index for buprenorphine were 0.14.

Cocaine: Kappa coefficient between self-report and toxicology screens for cocaine use was 0.55 (0.32; 0.79) indicating moderate agreement. The PABAK value was 0.73 indicating substantial agreement between the two measures after adjusting for prevalence and unequal distribution in the 2X2 tables. The expected and observed proportions of agreement were 0.71 and 0.87 respectively with 0.62 and 0.92 proportion of positive and negative agreement respectively. The prevalence index was -0.65 and bias index equal to 0.11. There were no differences in the PABAK values of the occasional and regular users, 0.72 and 0.74, respectively.

Marijuana: The simple kappa coefficient for marijuana was 0.61, which after adjusting for prevalence and bias changed to 0.76 signifying substantial agreement between the two measures for marijuana use during pregnancy. The expected and observed proportions of agreement were 0.69 and 0.88 respectively with 0.69 and 0.93 proportion of positive and negative agreement respectively. The prevalence index was -0.61 and bias index equal to 0.07. Moreover, there were no differences in the PABAK value of the occasional and regulars, 0.75 and 0.76 respectively.

Benzodiazepines: The kappa coefficient for agreement between self-reports and toxicology screens for benzodiazepines indicated a fair agreement (k=0.36, CI: -0.04;

0.76). However, after adjusting for low prevalence of benzodiazepines use in this population, PABAK values of 0.86 represented almost perfect agreement between the two measures. Moreover, the expected and observed proportions of agreement were 0.89 and 0.93 respectively with 0.40 and 0.96 proportion of positive and negative agreement respectively. The prevalence index was -0.88 and bias index equal of 0.

Amphetamines: Similarly, to benzodiazepines, the value of simple kappa coefficient for amphetamines represented fair agreement between the two measures (k=0.52, CI: 0.16; 0.88). However, PABAK value of 0.88 indicated an almost perfect agreement between self-reported drug use and urine toxicology screens. In addition, the expected and observed proportions of agreement were 0.87 and 0.94 respectively with 0.55 and 0.97 proportion of positive and negative agreement respectively. The prevalence index was -0.87 and bias index equal of 0.06.

4.5. Sensitivity and Specificity of Self-Reports

The sensitivity and specificity of self-report using urine toxicology screens as the gold standard are presented in Table 11. While both the sensitivity and specificity of self-report for methadone use during pregnancy was 100%, the sensitivity of self-report for buprenorphine was slightly lower than that of methadone at 83%. Sensitivity of self-report for marijuana (57.9%) and opioids including both prescription opioids and heroin (58.3%) was similar, however their specificity varied with 97% and 90%, respectively. Cocaine, benzodiazepines, and amphetamines had lowest sensitivity values of 47%, 40% and 37%, respectively, although specificity of self-report for these drugs was almost perfect (> 96%).

Additional subgroup analyses were performed by stratifying cocaine, marijuana, and opioids users by occasional and regular user. This analysis revealed that the sensitivity of self-reports decreases with increased use whereas specificity increases among frequent users. For example, sensitivity of self-report for cocaine decreased from 33.3% to 28.6% among occasional and regular users respectively whereas specificity increased from 98.4 to 100%. Similarly, the sensitivity of self-reports for opioids (47.1% to 16.7%) and marijuana (42.9% to 38.5%) decreased whereas specificity of opioids increased from 89.7 % to 98.1% whereas there was no change for marijuana after stratifying as occasional and regular users separately.

4.6. Effect of Number of Urine Screens on Agreement

Logistic regression was performed to evaluate the effect of number of urine toxicology screens on the agreement between self-reports and urine toxicology screens (agreement vs. no agreement) for each major class of drugs. Result of this analysis revealed no association between the number of urine drug screens and the level of agreement, as represented in Table 12.

Chapter 5: DISCUSSION

This chapter first briefly summarizes the key findings of the study, followed by the similarities and differences of the results with the existing literature. Limitations and strengths of the study, implications of study findings along with the recommendations for future research work then are discussed in the subsequent sections.

5.1. Summary of Key Findings

As the primary objective of the study was to assess the validity of self-reported drug use during pregnancy for seven common drugs of abuse in a unique population of pregnant women receiving prenatal care at a substance-abuse treatment clinic at the University of New Mexico, simple kappa, PABAK coefficients, sensitivity and specificity were estimated and reported. The values of PABAK indicated perfect agreement for methadone use and substantial agreement for all other classes of drugs. Sensitivity of self-reports for methadone use during pregnancy was 100% whereas it was slightly lower for buprenorphine, 83%. While the sensitivity of self-reports was poor, with only slightly more than half of the opioid and marijuana users reporting their use in the self-reports, it was even poorer for cocaine, benzodiazepines and amphetamines.

The value of simple kappa coefficients indicated a varied level of agreement between the two measures for the reported classes of drugs of abuse. The simple kappa coefficients indicated perfect agreement for methadone. In other words, there was no discordance in self-reported methadone use and the results of urine toxicology screens implying that these women were indeed completely honest in admitting their methadone use during pregnancy. The levels of agreement varied from almost perfect for buprenorphine use,

substantial agreement for marijuana use, moderate agreement for cocaine, amphetamine, opioids, and fair agreement for benzodiazepine use. However, as discussed previously, one of the major limitations of simple kappa coefficient is its dependence on prevalence. Thus, in order to overcome this limitation of the simple kappa coefficient, PABAK coefficients were also calculated for the above-mentioned classes of drugs. In concordance with the simple kappa, the value of PABAK for methadone indicated perfect agreement. However, for all other classes of drugs, the value of PABAK between 0.6-0.8 indicated moderate to almost perfect agreement between the two measures. The similar values of PABAK coefficients according to the classes of drugs and frequency of use implied no differences in the self-reporting. Thus, the measure of agreement revealed a good agreement between self-reports and urine toxicology screens.

Since kappa values are a function of sensitivity, specificity, prevalence and bias index (ref1), sensitivity and specificity of self-reports for individual drug classes were also reported. The 100% sensitivity and specificity of self-report for methadone use explains no discordance in self-reports and urine toxicology screens. Sensitivity of self-report for buprenorphine, being another choice of therapy, was lower than expected at 83%. One plausible explanation for this could be the illicit use of buprenorphine (purchased without prescription) by the enrolled patients. Sensitivity of self-report for opioids (prescription and heroin) and marijuana was slightly higher (58.3% and 57.9%, respectively) than other classes of drugs (ranging from 37.5% to 47.4%) indicating substantial underreporting of drug use by the enrolled patients for marijuana, cocaine, opioids, benzodiazepines, and amphetamines. Hence, large epidemiological studies merely relying on self-reports as a measure of drug use assessment would be able to capture less than

half of actual users. Moreover, as these women were recruited from a specialized clinic, a higher degree of accuracy in reporting drug use during pregnancy is expected in this population as opposed to the general OB/GYN population as these women are on therapy for their addictions and have nothing to conceal. Thus, we expect sensitivity of self-report to be even lower in the general population due to the social stigma attached with drug use during pregnancy. Specificity of self-reports for all classes of drugs was close to 100% except for opioids perhaps due to their short detection windows in the urine testing (See Table 1). We here would like to acknowledge that even though detection window of amphetamines is also short but the specificity of self-reports for amphetamine use was high compared to opioids due to very few amphetamine regular users. We would also like to acknowledge that the opioids given during labor and delivery will not affect the results of this study as toxicology screens conducted before labor and delivery were included in this study.

5.2. Previous Literature on Self-Reports and Urinalysis

The number of prior studies assessing agreement between self-reports and urine toxicology screens for drug use during pregnancy are very limited and focused only on cocaine and marijuana use during pregnancy. In the study conducted by Marroun et. al., the value of agreement between self-reports and urinalysis for prenatal cannabis use (0.77) corresponded with the results of our study (0.76). However, the sensitivity of self-reports was 36%, much lower than our results (58%). A plausible explanation for this finding in the Marroun's study was that it was conducted in a general Dutch population with high educational levels (45% reported higher than secondary education) and low prevalence of cannabis use during pregnancy (2.3%). Hence due to the social stigma

attached with drug use during pregnancy, these women were less likely to admit their cannabis use. Moreover, the results of the Marroun's study should be interpreted with caution because of the methodological limitations in the study design. Self-reports, collected at the time of enrollment retrospectively assessed cannabis use either before pregnancy or during the last three months from the interview. The authors mentioned that self-reports were usually collected in the first trimester of pregnancy, however, no detailed information or data was provided on this. Urine samples in the Marroun study on the other hand, were collected in early, mid and late pregnancy with 1-3 urine samples per individual throughout pregnancy. In addition, the time interval between interviews and urine collection was not reported.

Our measure of agreement is also similar to a study reported by Yonkers[88] et.al. for cocaine (0.73 and 0.70) and marijuana use (0.76 and 0.74) during pregnancy in a substance abuse treatment population of 168 women. However, Yonkers found self-reports to be highly sensitive, 78% sensitivity for marijuana and 86% for cocaine in a population comprised of majority of African-American. The time frame used in this study was the past month from the interview whereas we captured last trimester of pregnancy comprising of three months indicating higher sensitivity of self-report for shorter time windows. Moreover, Yonkers utilized audio softwares for data collection on the substance use, thus women were more open about their addiction as opposed to data collected by healthcare/research staff. In the study conducted by Lindsay[87] et. al. in an urban sample of pregnant women, only 47% of the women with positive urine test acknowledged their cocaine use during the interviews thus revealing low sensitivity of self-reports for cocaine use during pregnancy in an urban sample of pregnant women

seeking prenatal care. Although sensitive of self-report for cocaine use in this population (47%) is in complete agreement with the sensitivity of self-report in our population (47.3%), however, the two populations are completely different and hence comparison may not be valid.

In another study by Christmas[84] et. al., the sensitivity of self-reports for any drug use during pregnancy was 41% with only 17 out 41 women admitting drug use in the interviews. In that study, authors mentioned that the interviews were conducted during the first prenatal visit; however, no information was provided about variability in the timing of this prenatal visit. In addition, urine collection process and time-period were not documented explicitly. The study evaluated self-reports for a number of drug classes, however, only cumulative results for any drug use were presented, thus making the comparison with the results of our study difficult.

Horrigan[85] et. al. compared three measures, self-reports, urinalysis and substance abuse subtle screening inventory (SASSI) in order to determine which combination of measures would yield maximum sensitivity. SASSI consisted of two separate screens, first comprised of 26 face valid items known as risk prediction scales and second is made up of 52 true/false empirically derived items. Self-reports identified 15.2% of users, whereas urinalysis identified 17.3% of users, and SASSRI yielding highest sensitivity by identifying 43.4% of users. Although highly sensitive than urine screens and self-reports, SASSI is a complex tool, which requires 15-20 minutes for the respondents for complete, thus limiting its acceptability. Moreover, the circumstances under which self-reports, and urine collection process took place were not elaborated in the article.

The results of our study, however, are in complete disagreement with Bibb[86] et. al. They compared the validity of maternal interview versus drug screens conducted using newborn meconium, newborn urine, and maternal urine samples for cocaine and cannabis use. In that study, urinalysis or meconium screen were less sensitive for identification of cannabis use than self-reports, as 5.7% of mothers admitted cannabis use, while only 2.5% of maternal urine samples and 1% of meconium samples were positive. Our explanation for this is the data collection time frames employed in this study, as interviews and urine samples were collected after delivery. Interviews captured drug use during pregnancy and urine analysis due to limited detection windows cannot capture drug use throughout the pregnancy and as a result must have missed early pregnancy drug use. With respect to cocaine use, maternal interview, urine and meconium drug screens, all identified an equal number of users (3.4%). It is suspected that these women may have underreported their cocaine use in the interviews as urinalysis and meconium analysis may not have captured all the users owning to methodological limitations in the urine collection timing and that meconium formation does not take place until the second part of pregnancy. Long window of whole pregnancy captured by interviews and one urine sample at the end of pregnancy make the results of this study unreliable.

In summary, only one study has previously evaluated agreement between self-reports and urine screens in pregnant population enrolled at a substance abuse treatment program. Only two drugs were evaluated in that study (cocaine and marijuana) and only simple kappa coefficients were estimated. Our study not only presents multiple measures of agreement such as simple kappa, PABAK, sensitivity and specificity, but has a more

robust study design with multiple urine screens in a trimester and evaluation of some commonly abused classes of drugs.

In addition, the results of our study revealed higher sensitivity of self-report for opioids and marijuana compared to the other classes of drugs. While none of prior study has reported class-wise differences in the sensitivity of self-report in a population of pregnant women, a very limited number of studies have estimated this in diverse populations. A study conducted by Musshoff[103] et. al. compared the results of urinalysis and hair analysis with self-reported drug use data for opiates, cocaine, amphetamines, methadone, and cannabinoids in drug users from a psychiatric clinic. The authors found that except for opiates, all other drugs showed low correlation between self-reports and urinalysis. In another study conducted by Ledgerwood [104] et.al., the sensitivity of self-reports compared to hair analysis varied depending on the drug class with opioids self-report exhibiting highest sensitivity (78%) and methamphetamine(44%) lowest sensitivity in the same cohort of middle age men. Another study conducted by Glintborg[105] et. al. assessed the reliability of self-reports for amphetamines, barbiturates, benzodiazepines, cannabinoids, cocaine, methadone and opioids use in a random sample of 100 hospitalized elderly medical patients. This study used a combination of urine and blood sample for drug analysis and found 100% sensitivity of self-reports for opiates, which however, should be interpreted with caution as sample of this study was hospitalized and this could include only prescription opioids. The sensitivity of self-report for cannabinoids in this population was 0% as none of the 5% users acknowledged its use. No conclusion can be drawn from this as cannabinoids have relatively longer detection windows (upto 30 days) and self-reports assessed use in the preceding one week

timeframe. Sensitivity of benzodiazepines was also relatively low (53%) and remaining drugs were almost non-prevalent in this population. In another study conducted by Magura[106] et. al. in 250 methadone treatment clients, they found opioid self-reports to be least sensitive, whereas benzodiazepines and cocaine self-reports were moderately and highly sensitive, respectively. In another study conducted by Lu et. al[107]. on a population of arrestees that participated in the Arrestee Drug Monitoring program, the authors found least amount of underreporting among the marijuana users with 64% sensitive self-reports using urinalysis, followed by methamphetamine (56%), cocaine (48%) and least sensitive self-reports for opiates (45%).

In summary, some prior studies have demonstrated higher sensitivity of self-reports for opioids and marijuana use compared to other classes of drugs; however others have completely nullified this notion. Moreover, differences in the study population and cohort characteristics, interview timings, framing of questions, and research methods make each study unique and it is difficult to derive conclusions from the existing literature for our unique population of pregnant women.

5.3. Limitations and Strengths

It is important to acknowledge that the urine toxicology screens conducted during the third trimester of pregnancy were compared with the self-reported drug use which captured the timeframe between the baseline and follow-up interviews. The average time between the baseline and follow-up interviews was16.8 ±7.6 (range: 4-32) weeks. The time-period captured by the interviews and detection window of urine screens do not overlap completely and this may lead to potential misclassification bias. For instance, the

results of urine screens conducted in the third trimester may lead to misclassification of a user who accurately reported drug use in the early second trimester but then discontinued use later on. This might lead to disagreement between the two measures owning to the detection window of the urine screens. Moreover, detection windows for each drug class vary with opioids (oxycodone, morphine) having the shortest detection window of 1-4 days and cannabinoids having the longest window of upto 30 days for chronic users. However, the majority of patients were enrolled during the second trimester of pregnancy (median gestational age at recruitment: 21weeks) approaching the timeframe captured by urine screens. Moreover, the majority of women had multiple urine screens during the third trimester (4.7 screens on average per patient) increasing the detection window. Our results also demonstrated that the number of screens did not significantly affect the level of agreement between the two measures.

We would also like to acknowledge that buprenorphine, methadone, and prescription opioids were separate urine toxicology screens. This implies that someone who takes buprenorphine would be negative for methadone or opiates but positive for buprenorphine would require a separate test to be ordered.

We would also like to acknowledge that the use of brand names for the classes of drugs evaluated in the study can contribute to the accuracy of self-reports. However, street/common names of commonly abused drugs were mentioned in the questionnaire, for example, buprenorphine and subutex were used for buprenorphine. Moreover, common examples of prescription opioids were provided in the questionnaire to help the respondents to accurately recall their use.

Another limitation includes limited generalizability of the results of the study as the sample was enrolled from a specialized prenatal clinic providing care to women with past or current history of substance abuse. Thus, results might not be generalizable to the general population of pregnant women. However, we expect that the sensitivity of self-reported drug use would be even poorer in the general population than in our study due to the profound stigma attached with drug use during pregnancy. In the Milagro clinic, all patients have already acknowledged their substance abuse and are actively seeking care for it. Thus, it was surprising to find out that even in this population sensitivity of self-report was rather poor for most drug classes except opioids and marijuana. This underlines the need to supplement maternal reports with other methods in order to accurately capture drug use in epidemiological studies.

We would also like to acknowledge that owning to the small sample size and low prevalence of certain classes of drugs, such as inhalants and ecstasy; we were not able to draw conclusions regarding the accuracy of their report during pregnancy. Low prevalence of some other classes was adjusted for by estimating PABAK coefficients in addition to a simple kappa coefficient. In addition, we cannot differentiate the accuracy of self-report for use/abuse of prescription opioids and heroin since the urine toxicology screen employed in the Milagro clinic did not differentiate between these subtypes of opioids.

Unique strength of the study includes a well-characterized cohort comprising mostly of ethnic minority. While prior studies were majorly comprised of either White or African-American population we were able to capture a large proportion of Hispanic minority with economically disadvantaged background. In addition, the study was able to assess

the validity of self-reported drug use for multiple drug classes in a same cohort. Other studies focused on one or two drug classes or lumped all drugs into a single ‘any drug use’ category.

5.4. Implications and Future Recommendations

Accurate identification of illicit drug use during pregnancy offers unique opportunities for both healthcare providers and pregnant women. This is an ideal time for healthcare providers to help pregnant women in overcoming their drug addiction problems and thus develop a healthy lifestyle, not only for themselves but also for their babies. However, the social stigma attached with this issue might lead to the withholding of important drug use information to the healthcare providers and hence appropriate care and therapy may not be made available to these women. Accurate identification of drug use during pregnancy is complicated and ensuring the accuracy of self-reports is important from clinical and public health perspective, as it is the most inexpensive and commonly employed method of data collection.

The result of this study indicate substantial agreement between urine toxicology screens and self-reports; however, this should be interpreted with caution as agreement takes both sensitivity and specificity into account and high specificity, a function of effective urine screening process, may alter the results. Thus special attention should be paid to the sensitivity of self-reports for drug use and our results reveal low sensitivity of self-reports for all major classes of drugs despite the fact that these women were enrolled from a specialty clinic and were aware of the purpose of the urine collection process. Moreover,

as evident from the subgroup analyses of cocaine, marijuana, and opioids, the sensitivity of self-reports was even lower among the regular users compared to occasional users. A plausible explanation for this could be decreased willingness among regular users to admit drug use during pregnancy owning to chronicity or severity of their addiction. Hence, studies only relying on self-reports for drug use information during pregnancy are very likely to suffer from reporting bias, more specifically underreporting. An approach involving a combination of both self-reports and urinalysis for accurate identification of prenatal drug use is suggested based on the evidence found in this study. However, in circumstances when self-reports is the only mode of data collection, few measures can that can improve the accuracy of self-reports include assuring confidentiality and anonymity of the collected information, using skillful and emphatic interviewers or use of audio-visual aids in case of sensitive information, using a short time interval between the occurrence of actual event and the data collection process[12].

Future studies focusing on different classes of drugs with adequate urine samples per semester (weekly/fortnightly) in different populations are required. In addition, studies examining the factors affecting the accuracy of self-reports in different populations, mode and timing of data collection, stratification of study population as chronic and occasional users would add to the gap in the literature.

Conclusions

In conclusion, the findings of our study revealed that pregnant women from a substance-abuse treatment clinic highly underreported their drug use during pregnancy. While opioids and marijuana were more accepted than other drugs in this population, the level of underreporting was substantial. We expect even lower sensitivity of self-reports for

prenatal drug use in the general population. We found empirical evidence that it is difficult to obtain valid prenatal drug use information by merely relying on self-reports. It is highly likely that epidemiological studies merely relying on self-reported drug use data would result in biased results and prevalence. Therefore, in order to improve the data collection process on prenatal substance use, we suggest a 2-step approach involving both self-reports and urinalysis.

Table 6: Description of the Study Population

Patient Characteristics (N=102)	
Maternal age (range: 18-41 years)	26.4 ± 4.9
Gestational age at enrollment (weeks)	20.9 ± 7.9
Gestational age at delivery (weeks)	38.6 ± 2.1
	%
Ethnicity: Latina	77.5
Race:	
White	86.1
American Indian	8.9
Black/African American	4.0
Other	1.0
Marital Status:	
Single, never married	53.9
Married/ not-married, living with partner	38.2
Divorced/Separated/Widowed	7.9
Educational Level	
Less than high school graduate	51.5
High school graduate or GED	27.7
College or higher	20.8
Employed	11.8
Health Insurance Status:	
No insurance	3.9
Employer or self-purchased insurance	1.0
Medicaid or other public insurance	95.1
Primary language: English	97.1
Place of birth: US	98.0
Smoking	
Any maternal smoking during pregnancy	79.4
No maternal smoking during pregnancy	20.5
Paternal smoking	63.7
Alcohol use	
Abstainers	44.8
Early pregnancy	35.4
Continuous use	19.8

Table 7: Maternal Medical and Reproductive History (N=102)

Medical and Reproductive Characteristics	
	%
Presence of chronic conditions	37.3
Hepatitis	14.7
Depression	15.7
Anxiety	18.6
Migraine	5.0
Asthma	3.9
Diabetes/ thyroid/ heart disease	1.0
Primigravida	25.4
Unplanned pregnancy	85.3
Healthy BMI (18.0-24.9)	52.9
Use of prenatal vitamins/ iron	88.2
History of adverse perinatal outcomes	58.8
Stillbirth	2.9
Termination	25.5
Miscarriage	46.1
Ectopic pregnancy	5.0
Complications in pregnancy	15.7
Bleeding	8.0
High blood pressure	2.0
Other	6.0

Table 8: Prevalence of Drug Use in the Milagro Population N(%)

	Self-Report	Toxicology screens	Total prevalence*
Marijuana	13 (15.7)	19 (22.9)	21 (25.3)
Cocaine	10 (12.0)	19 (22.9)	20 (24.1)
Methadone	53 (63.8)	53 (63.8)	53 (63.8)
Buprenorphine	15 (51.7)	18 (62.1)	18 (62.1)
Opioids**	20 (24.4)	24 (29.3)	30 (36.5)
Benzodiazepines	5 (6.0)	5 (6.0)	5 (6.0)
Amphetamines	3 (3.6)	8 (9.6)	8 (9.6)

*Total prevalence calculated using either measure, self-report or urine toxicology screen
** Opioids include both prescription opioids and heroin

Table 9: Frequency of Drug Use during Pregnancy by Major Classes*, N (%)

	No use	Occasionally	Once a month	Once every 2 to 3 weeks	Once a week	Almost everyday
Marijuana	73 (83.9)	6 (6.9)	2 (2.3)	1 (1.2)	1 (1.2)	4 (4.6)
Methadone	31 (35.6)	--	--	--	--	56 (64.4)
Buprenorphine[¥]	71 (81.6)	--	--	--	1 (1.2)	15 (17.2)
Amphetamine	84 (96.5)	1 (1.2)	--	2 (2.3)	--	--
Benzodiazepine	82 (94.2)	1 (1.2)	--	2 (2.3)	1 (1.2)	1 (1.2)
Cocaine	81 (93.1)	2 (2.3)	1 (1.2)	2 (2.3)	1 (1.2)	--
Opioids	72 (82.7)	8 (9.2)	2 (2.3)	--	2 (2.3)	3 (3.5)
Ecstasy	87 (100)	1 (1.2)	--	--	--	--
Inhalants	86 (98.8)	--	--	--	--	--

* Based on self-reports
[¥] Sample size was 29 for buprenorphine as only these patients had urine toxicology screen results for it.

Table 10: Simple Kappa Coefficient, PABAK and Other Measures of Agreement.

	N*	Prevalence**	**Kappa**	C.I.	Expected agreement	Observed agreement	Proportion of positive agreement	Proportion of negative agreement	Prevalence index[¥]	Bias index[α]	**PABAK[ε]**
Marijuana	83	21	**0.61**	0.41; 0.83	0.69	0.88	0.69	0.93	0.61	0.07	**0.76**
Occasional use	77		**0.51**	0.21; 0.75	0.76	0.88	0.57	0.93	0.73	0.09	**0.75**
Regular use	76		**0.51**	0.23; 0.79	0.78	0.88	0.53	0.93	0.75	0.09	**0.76**
Cocaine	83	20	**0.55**	0.32; 0.78	0.71	0.87	0.62	0.92	0.65	0.11	**0.73**
Occasional use	79		**0.41**	0.14; 0.68	0.76	0.86	0.48	0.92	0.73	0.11	**0.72**
Regular use	77		**0.40**	0.12; 0.67	0.87	0.79	0.44	0.93	0.77	0.13	**0.74**
Methadone	83	53	**1**		0.54	1.00	1.00	1.00	0.28	0	**1.00**
Buprenorphine	29	18	**0.79**	0.57; 1.00	0.50	0.90	0.91	0.88	0.14	0.14	**0.79**
Opioids	82	30	**0.50**	0.29; 071	0.61	0.81	0.64	0.87	0.46	0.05	**0.61**
Occasional use	77		**0.40**	0.16; 0.64	0.65	0.79	0.53	0.87	0.56	0.05	**0.58**
Regular use	65		**0.2**	-0.07 ; 0.49	0.79	0.83	0.27	0.90	0.77	0.14	**0.66**
Benzodiazepines	83	5	**0.36**	-0.04 ; 0.76	0.89	0.93	0.4	0.96	0.88	0	**0.86**
Amphetamines	83	8	**0.52**	0.16; 0.88	0.87	0.94	0.55	0.97	0.87	0.06	**0.88**

* N is the sample with both self-reports and urine toxicology screens available

** Prevalence represents total prevalence estimated using either measure, self-report or urine screens

[ε] PABAK: Prevalence-adjusted-bias-adjusted kappa

[¥] Prevalence Index: Difference in the proportion of positive and negative agreements

[α] Bias Index: Extent to which the raters disagree on the proportion of positive or negative cases.

Table 11: Sensitivity and Specificity of Self-Reports for Drug Use during Pregnancy

	Simple kappa	PABAK	Sensitivity (%)	Specificity (%)
Marijuana	0.61	0.76	57.90	96.90
Cocaine	0.55	0.73	47.37	98.44
Methadone	1	1	100.00	100.00
Buprenorphine	0.79	0.79	83.33	100.00
Opioids	0.50	0.61	58.33	89.66
Benzodiazepines	0.36	0.86	40.00	96.15
Amphetamines	0.52	0.88	37.50	100.00

Table 12: The Effect of Number of Urine Toxicology Screens on Class-Specific Agreement

	Odds Ratio (C.I.)*
Marijuana	0.9 (0.8;1.2)
Cocaine	1.0 (0.8;1.2)
Opioids	1.1 (0.9;1.3)
Buprenorphine	1.1 (0.7; 1.6)
Benzodiazepine	1.0 (0.8; 1.3)
Amphetamine	0.8 (0.6; 1.1)

* Odds of agreement are modeled for one unit increase in urine toxicology screens.

APPENDIX: DESCRIPTION OF STUDIES

AUTHOR	POPULATION/SAMPLE	COMPARATORS	RESULTS	COMMENTS
ADVERSE EFFECTS OF IN-UTERO DRUG EXPOSURE				
Mirochnick et. al[16].	95 term cocaine (and/or marijuana) exposed infants, confirmed by benzoylecgonine in meconium.	Cocaine exposure and impaired fetal growth.	Significant negative correlations between meconium benzoylecgonine concentration and birth weight, length and head circumference	Restricted to full term babies only.

Harsham et. al[17]	31 in utero exposed infants; predominantly black. Reference populations: the National Center for Health Statistics (NCHS); 1991 Pediatric Nutrition Surveillance System (PNSS-all); Black infants in PNSS (PNSS-black)	In utero cocaine exposure and postnatal growth pattern (for 1 year).	Mean weight of the exposed infants at birth was significantly lower than three reference populations ($p<0.01$); mean length was significantly lower than NCHS and PNSS-all infants ($p<0.01$). Mean length for exposed infants was significantly less than NCHS, PNSS ($p<0.01$) and PNSS-Black ($p<0.05$)	Limited to infants in foster care eliminated environmental effects.
Shankaran et. al.	365 cocaine-exposed and 771 non-exposed infants;	Cocaine exposure and SGA status (defined as birth weight <10 percentile on the Alexander curves)	Significantly lower growth parameters in cocaine exposed infants. Similar weight at 6 years. Significant interaction between prenatal growth exposure and SGA status at 6 years.	Self-reported non-exposure status confirmed by meconium analysis. Matching: gestational age, sex and race.

Bandstra et. al[20].	476 neonates with in-utero cocaine exposure	Cocaine exposure and fetal growth and gestational age	Cocaine associated growth deficits were observed	African-American population
Eyler et.al[21].	154 cocaine exposed and 154 control mothers	Cocaine exposure and neonatal outcomes	Significantly decreased head circumference among exposed infants.	Under-studied rural public health population
Minnes[22] et. al.	156 prenatally exposed to cocaine 6 year old children; 131 high risk controls	Cocaine exposure and growth parameters along with dysmorphic outcomes from birth through 6 years of age.	Heavier prenatal exposure to cocaine negatively affected height and height for weight z scores at age 6.	Considered various levels of exposure: non-exposed, light, heavy, units per week and meconium threshold.
Bauer et. al[23]	Cocaine-exposed infants:	Cocaine exposure	Cocaine exposed: 1.2 weeks	Multisite study with

	717 and 7442 non-exposed infants	during pregnancy and medical conditions in newborn infants from birth through hospital discharge.	younger, less in weight, smaller height and head circumference. (all P<.001). CNS and ANS symptoms more frequent in the exposed group as well as more infections.	large sample size.
Ostrea et. al.[6]	44% of 2964 infants positive for drugs.	Death outcome in drug exposed infants	High perinatal morbidity and high mortality in low weight drug exposed infants.	High odds ratio for sudden infant death syndrome among cocaine exposed (1.9)
Cherukuri et. al.[9]	55 crack using mothers and 55 non-drug exposed	Crack exposure and fetal effects	Significant differences in preterm births, intrauterine growth retardation and premature rupture of membranes.	Mild neurobehavioral symptoms among crack exposed infants.

Fries et. al. [25]	32 with prenatal cocaine exposure referred to genetics evaluation	Cocaine, alcohol exposure and Fetal effects	Distinctive phenotypes along with premature births, small-for-gestational age and smaller head circumference	Distinctive diagnosis for fetal cocaine syndrome
Dogra et. al [26]	40 cocaine exposed neonates and 34 controls	Increased neurosonographic abnormalities in cocaine exposed	35% cocaine exposed had neurosonographic abnormality compared to none in controls	Degenerative changes or focal infarctions in basal ganglia of cocaine exposed neonates.
Behnke et. al. [28]	311 cocaine exposed infants matched to controls	Hospital costs associated with prenatal cocaine exposures	Cocaine exposed: increased healthcare services utilization, longer lengths of stay, and higher charges.	Population with minimum access to drug rehabilitation services
Bandstra[30] et. al.	Longitudinal Analysis of 451 full-term children; 242 cocaine-exposed and	Total, expressive, and receptive language at ages 3, 5 and 12 using age-appropriate	PCE was associated with lower expressive and total language scores.	Children with prenatal substance exposure, except alcohol, tobacco and marijuana were

	209 non-exposed.	versions of the Clinical Evaluation of Language Fundamentals (CELF)		excluded from the analysis.
Lewis et. al[31].	175 PCE exposed children and 175 non-exposed children followed to 10 years of age	Language subscales of the Test of Language Development-Intermediate 3rd Edition (TOLD-I:3) and phonological processing measured by the Comprehensive Test of Phonological Processing (CTOPP)	PCE cocaine effects were observed for aspects of language including syntax semantics and phonological processing (p=0.001)	Study sample limited to African-American children in disadvantaged neighborhoods.
Schuetze et. al[32]	Heart rate and respiratory sinus arrhythmia (RSA) in cocaine exposed and control infants	Prenatal cocaine exposure and autonomic regulation at 7 months of age	Cocaine exposed had higher heart rate and significant suppression of RSA.	Supported autonomic dysregulation with cocaine exposure.

Lewis et. al.[33]	209 in utero cocaine exposed and 189 non-exposed children	Prenatal cocaine and polydrug exposure on language development of preschool kids. Compared on receptive, expressive and total language scores at 1, 2, 4 and 6 years of age.	Significant negative effect of cocaine on all language domains. Cocaine-exposed children demonstrated linguistic deficits compared with non-exposed peers and did not catch up.	Controlled for cigarette and environmental factors.
Bandstra et. al. [34]	200 cocaine-exposed and 176 non-exposed African-American children	Longitudinal effects of severity of prenatal cocaine exposure on language functioning through age 7 years	Greater severity of PCE was associated with greater deficits for language performance (D = -0.071, 95% CI = -0.133, -0.009; p = 0.026).	Factors included in the analysis were fetal growth, gestational age, and IQ as intercorrelated response variables and child's age, gender, and prenatal alcohol, tobacco, and marijuana exposure as covariates

Bateman et. al.[19]	240 healthy infants exposed to cocaine exposed in third trimester. No exposure (n = 136), low cocaine exposure (n = 52), and high cocaine exposure (n = 52) by hair analysis	Relationship between head circumference, birth weight, and cocaine dose in prenatally exposed to cocaine infants.	Mean birth weight, length, and head circumference of infants with high cocaine exposure differed significantly from those with low exposure.	Birth weight, sex, and high cocaine exposure significantly associated with newborn head circumference
Morrow et. al[36].	253 cocaine-exposed and 223 non-cocaine-exposed	Influence of PCE on children's language functioning at six time points from 4 months to 3 years of age.	Cocaine-exposed children had lower overall language skills than non-cocaine-exposed children (D = -0.151; 95% CI = -0.269, -0.033; p =.012).	Remained stable after evaluating confounding by prenatal substance exposures and sociodemographic factors.
Bandstra et. al. [37]	Urban sample of 236 cocaine-exposed and 207 noncocaine-exposed)	Longitudinal effects of in utero cocaine exposure on language	Significant association between prenatal cocaine exposure and deficits in total	Supported cocaine-specific effect on indicators of language

		functioning at 3, 5 and 7 years of age	language functioning (D=-0.17; 95% CI=-0.32, -0.03; P=.019).	functioning during early childhood through age 7 years.
Singer et. al[38].	131 Non-exposed children; 66 heavily exposed and 68 lightly exposed	Association of level of fetal cocaine exposure to developmental precursors of speech-language skills at 1 year of age	Heavily exposed infants had lower auditory comprehension scores; more likely to be classified as mildly delayed by total language score.	PCE led to attentional disabilities underlying auditory comprehension.
Mentis et. al[39].	5 prenatally cocaine exposed children	Language development and cocaine exposure	Compromised language development in cocaine exposed infants	Analysis of 30 minute language sample
Morrow et. al. [42]	212 cocaine exposed and 197 non-cocaine exposed children enrolled at birth	Risk of developing learning disability or impaired intellectual functioning by age of	No differences in the estimate of relative risk for impaired intellectual functioning between exposed	Results remained stable with adjustment for multiple child and care-giver covariates

		7 years	and non-exposed. However, cocaine-exposed children had 2.8 times greater risk of developing a LD by age 7 than non-cocaine-exposed children (95%CI = 1.05,7.67; p = .038; IQ ≥ 70 cutoff)	
Fajemirokunet. al. [51]	110 babies born to 108 women who used opiates in later pregnancy	Opioids exposure and neonatal outcomes	Significantly high likelihood among neonates born to heroin using mother to need morphine than methadone (40% vs 19%); high NAS score (5.8 vs 4.7); longer stay in neonatal units (17.2 days vs 11.8)	Heroin using women with higher incidence of NAS
Binder et. al[52].	47 heroin, 32 methadone and 38 buprenorphine addicted pregnant women	Effect of maintenance therapy on neonatal outcomes and NAS.	Infants among heroin using women had lowest birth weight, highest proportion of IUGR and placental change ($p < 0.05$). The severity and	Methadone notably protracted the newborn's abstinence syndrome.

			course of NAS were most severe ($p < 0.001$) in newborns of women from the methadone group.	
Alrettaz et. al.[53]	86 neonates born to mothers enrolled in methadone maintenance program.	Neonatal impact of methadone use during pregnancy.	24% babies were premature, 27% babies were growth retarded (<3rd centile), and 13% had microcephaly (<3rd centile). 62 % developed NAS requiring pharmacological treatment for median 47 days.	Child services involved in 56% cases and 42% neonates were placed outside mother homes.
Greig et. al.[59]	44 pregnant women on methadone and 88 non-	Maternal and neonatal outcomes of pregnant women	The MSP group higher relative risk of premature delivery, lower birth weight,	No difference in congenital abnormalities in the

	methadone controls	enrolled in a Methadone Substitution Programme (MSP)	smaller head circumferences	two groups; although controls had higher caesarean sections.
Broussard et. al. [62]	17,449 mothers of kids among cases with various birth defects and 6701 control mothers	Association between opioid use early in pregnancy and birth defects.	Statistically significant association of opioid use with conoventricular septal defects atrioventricular septal defects, hypoplastic left heart syndrome, spina bifida or gastroschisis	Commonly reported opioids were codeine hydrocodone oxycodone and meperidine
Bracken et. al. [63]	1472 cases and 3001 controls	Exposure to prescription opioids during pregnancy and congenital malformations in	Case mothers were more likely than controls to have used a prescription drug (odds ratio [o] = 13, P less than .0001), particularly an	Incidence of congenital malformations was 52 per 1000 live births and 44% reported using atleast one

		newborns	antidepressant (o = 7.6), narcotic analgesic (o = 3.6), or tranquilizer (o = 23); all p <0.01	prescription drug during pregnancy.
Saxen et.al[64].	599 children with clefts	Clefts and maternal drug consumption during pregnancy	More frequent consumption of drugs among mothers of infants with cleft lip.	Analgesic, chemotherapeutic and antineurotic more frequent among cases.
Clearly et. al[66].	117 pregnant women on methadone maintenance program	Perinatal outcomes involving NAS between methadone users only and concomitant polydrug users.	Methadone-only exposed neonate had shorter hospitalization than those exposed to methadone and concomitant drugs (median 5.0 days versus 6.0 days, P = 0.03)	The incidence and duration of the NAS is not associated with maternal methadone dose but maternal polydrug use.
Hurt et. al. [108]	PFC activation during task	Gestational cocaine	Similar performance on n-	Exposed and non-

	performance in 25 cocaine exposed infants and 24 non-exposed.	exposure and its effects on prefrontal cortex (PFC) with functional magnetic resonance imaging (fMRI)	back task (P >/= .4), indicating increased demands on working memory with greater task difficulty. Region of interest images showed similar activation for both groups.	exposed were similar in performance on the executive function task and in fMRI activation patterns during task performance.
Beeghly et. al. [109]	160 low-income, urban children from a prospective study who completed a standardized language assessment at 6 and 9.5 years. PCE through neonatal meconium assays and maternal self-report.	Relationship between prenatal cocaine exposure and contextual variables on children's language at age 6 and 9.5 years.	PCE children had lower receptive language than unexposed children at 6 but not at 9.5 years. Also, lower expressive language for lower birth weight, and lower expressive and total language if they were female	Age, birth weight, and gender moderated the relation between PCE and language development

SELF-REPORT FOR RISKY BEHAVIORS AMONG PREGNANT WOMEN				
Gollenberg[77]et. al.	82 women enrolled in the periconceptional period	Smoking, caffeine, fish, vitamins, alcohol.	Poor to moderate agreement.	Though higher for regular behaviors like caffeine and smoking.
Gilligan et. al.[72]	Women attending antenatal care clinic.	Smoking and cotinine levels in urine	17% self-reported non-smokers misreported their smoking status.	Error in smoking assessment was substantial
Shipton et.al.[73]	3475 pregnant women	SR smoking and blood cotinine levels	SR led to an underestimation of 25%	SR leads to failure to detect 2400 smokers each year for cessation services.

Britton et.al.[74]				
Klebanoff et.al.[78]	448 pregnant women enrolled in a perinatal project	SR smoking and maternal cotinine assays	High kappa coefficient (k=0.83)	Pregnant women accurately reported their smoking status.
Secker-Walker[79]	Pregnant women	SR smoking, CO levels, and cotinine/creatinine levels	Moderate correlations between the measures	Urine cotinine/creatinine most accurate.
Ford et.al[75].	Postal questionnaires to 4875 mothers retrospectively	Questionnaires, cotinine levels and obstetric bookings	Underreporting in SR and obstetric bookings	22% cotinine validated denied smoking.

Burstyn et.al.[76]	92 smokers and 285 non-smokers (self-reported)	Self-reports and urinary cotinine assays	Poor sensitivity (47%). High specificity (95%)	Non-random sample of women with live births
Rice et. al[80].	126 mothers with school aged children	Pre and perinatal maternal self-reports with medical records	Good agreement for smoking (k=0.80) whereas poor for alcohol use (0.17)	Authors concluded high validity of self-reports
Fox et.al[81].	700 pregnant women in a randomized clinical trial for a smoking cessation intervention.	Alcohol and smoking self-reports with thiocyanate levels in saliva	Moderate agreement for both smoking and alcohol drinking patterns during pregnancy	Results may be biased as participants were aware about the objectives of sample collection.
		SR alcohol, smoking	Low agreement for alcohol	Maternal

SR: Self-report; CO: Carbon monoxide;

Hessol et. al[82].	350 Latina women	and various behavioral and medical factors with medical data	and smoking than for medical conditions	characteristics did not predicted patterns of disagreement.
SELF-REPORT VS URINE SCREENS AMONG PREGNANT WOMEN				
Marroun et. al. [83]	Generation R study, a population based birth cohort to collect data on a sample of parents and children from early pregnancy.	SR, UA Cannabis	**Moderate agreement** (Yule's Y=0.77) Sensitivity and specificity of self- report: 0.36 and 0.99. Sensitivity and specificity of urinalysis: 0.46 and 0.98	Neither gold standard. Two-step approach starting with self-report.
Christmas et. al.[84]	302 urban pregnant women attending university based	SR, UA	**50 % self-reported use**	Two-step approach

UA: Urinanalysis

	obstetric clinic.	Alcohol/ any illicit substance	**had positive toxicology screens;** **41.5 % with positive toxicology screen admitted to current use.**	Questionnaire administered at the initial visit; no such information about urine screens (conducted during the study period)
Horrigan et. al.[85]	1276 pregnant women	SR, SASSRI, UA Cannabinoids, cocaine, opiates, amphetamines, barbiturates, phencyclidines,benzodiazepines.	**Self-report : 15.4 %** **Urinalysis: 17.3 %** **SASSRI: 43.4%**	No screening procedure is better than the other.
Bibb et. al[86].	580 mother infant pairs	SR, UA, MA	**THC: Self-report: 5.7%; maternal urine: 2.5 %**	In addition to interview, maternal urine and newborn

		Tetrahydrocannabinol (THC), phencyclidine cocaine	**and meconium: 1%.** **Cocaine: Equal sensitivity (3.4%)**	meconium should also be incorporated.
Lindsay et. al[87].	5200 women consented for urine assay	SR and UA Cocaine	**47% with positive urine assay acknowledged use in self-report**	Prevalence of cocaine use: 5%
Yonkers et. al[88].	168 women substance abuse treatment program women	SR, UA	Marijuana (k)=0.74 caine (k)=0.70	Good agreement
Meconium				
Ostrea et. al[90].	Prospectively in 58 women	SR, meconium and hair analysis	Lowest sensitivity of self-reports for cocaine and opioids.	Meconium best screening method among the three evaluated in this study

SASSRI: Substance Abuse Subtle Screening Inventory ; MA: Meconium analysis

		Cocaine, opiate, cannabinoid		
Tassiopoulos et. al. [91]	281 HIV-infected women and children	SR and meconium Marijuana Cocaine	**Moderate Agreement** Sensitivity for self-reported marijuana use: 80% (kappa=0.61, p-value <0.01), cocaine: 67% (kappa=0.80, p-value <0.01).	Sensitivity of SR compared to urine or blood toxicology: 31% for cocaine, 30% for marijuana and 20% for opiates
Gray et. al. [92]	Pregnant women in the IDEAL study Infant Development, Environment and Lifestyle (IDEAL)	SR and meconium Methamphetamine Cannabis	Self-report more sensitive than meconium analysis for amphetamine and cannabis exposure	Majority ceased the use in first/second trimester

Lozano et. al.[93]	974 mother-infants dyads	SR and meconium Cannabis	5. 3% identified by meconium whereas only 1.7% self-reported.	Check for SR timing?
SR: Self-report; UA: Urinalysis; MA: Meconium analysis		**Umbilical Cord**		
Wright et. al.[89]	28 women admited drug use	SR , UCA Methamphetamines Cocaine Marijuana	K=0.19 Sensitivity and Specificity of cord levels: 32% and 85%	Cord had highest sensitivity for marijuana
		Hair		
			High prevalence according	Not a single person

Bessa et. al. [94]	Pregnant adolescents	SR and Hair Marijuana Cocaine	to hair analysis and none reported	accepted in self-report
Grant et. al [95]	405 post partum women	SR and hair Cocaine	Hair analysis identified 87% self reported use. 14% with negative SR were positive with hair analysis.	Both measures for gestational cocaine exposure.

SR: Self-report; UA: Urinalysis; MA: Meconium analysis

REFERENCES:

1. Abuse NIoD. http://www.drugabuse.gov/publications/drugfacts/nationwide-trends.

2. McGinnis JM, Foege WH. Mortality and morbidity attributable to use of addictive substances in the United States. Proceedings of the Association of American Physicians 2003;111:109-18.

3. Rehm J, Mathers C, Popova S, Thavorncharoensap M, Teerawattananon Y, Patra J. Alcohol and Global Health 1 Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. Lancet 2009;373:2223-33.

4. Abuse NIoD. http://www.drugabuse.gov/related-topics/trends-statistics#costs.

5. {http://www.samhsa.gov/data/nsduh/2k11results/nsduhresults2011.htm#2.6}.

6. Ostrea EM, Ostrea AR, Simpson PM. Mortality within the first 2 years in infants exposed to cocaine, opiate, or cannabinoid during gestation. Pediatrics 1997;100:79-83.

7. Wiss DA, CPT DI. PREGNANCY AND SUBSTANCE ABUSE. 2012.

8. Minnes S, Lang A, Singer L. Prenatal tobacco, marijuana, stimulant, and opiate exposure: Outcomes and practice implications. Addiction science & clinical practice 2011;6:57.

9. Cherukuri R, MINKOFF H, FELDMAN J, PAREKH A, GLASS L. A cohort study of alkaloidal cocaine (" crack") in pregnancy. Obstetrics & Gynecology 1988;72:147-51.

10. Jones HE, Johnson RE. Pregnancy and substance abuse. Current Opinion in Psychiatry 2001;14:187-93.

11. Harrison L. The validity of self-reported drug use in survey research: an overview and critique of research methods. NIDA Res Monogr 1997;167:17-36.

12. Del Boca FK, Noll JA. Truth or consequences: the validity of self- report data in health services research on addictions. Addiction 2000;95:347-60.

13. Rouse BA, Kozel NJ, Richards LG. Self-report methods of estimating drug use: Meeting current challenges to validity: National Institute on Drug Abuse Rockville, MD; 1985.

14. Wagner CL, Katikaneni LD, Cox TH, Ryan RM. The impact of prenatal drug exposure on the neonate. Obstetrics and gynecology clinics of North America 1998;25:169-94.

15. MacGregor SN, Keith L, Chasnoff I, et al. Cocaine use during pregnancy: adverse perinatal outcome. American journal of obstetrics and gynecology 1987;157:686.

16. Mirochnick M, Frank DA, Cabral H, Turner A, Zuckerman B. Relation between meconium concentration of the cocaine metabolite benzoylecgonine and fetal growth. The Journal of pediatrics 1995;126:636.

17. Harsham J, Keller JH, Disbrow D. Growth patterns of infants exposed to cocaine and other drugs in utero. Journal of the American Dietetic Association 1994;94:999-1007.

18. Gouin K, Murphy K, Shah PS. Effects of cocaine use during pregnancy on low birthweight and preterm birth: systematic review and metaanalyses. American journal of obstetrics and gynecology 2011;204:340. e1-. e12.

19. Bateman DA, Chiriboga CA. Dose-response effect of cocaine on newborn head circumference. Pediatrics 2000;106:e33-e.

20. Bandstra ES, Morrow CE, Anthony JC, et al. Intrauterine growth of full-term infants: impact of prenatal cocaine exposure. Pediatrics 2001;108:1309-19.

21. Eyler FD, Behnke M, Conlon M, Woods NS, Wobie K. Birth outcome from a prospective, matched study of prenatal crack/cocaine use: I. Interactive and dose effects on health and growth. Pediatrics 1998;101:229-36.

22. Minnes S, Robin NH, Alt AA, et al. Dysmorphic and anthropometric outcomes in 6-year-old prenatally cocaine-exposed children. Neurotoxicology and teratology 2006;28:28-38.

23. Bauer CR, Langer JC, Shankaran S, et al. Acute neonatal effects of cocaine exposure during pregnancy. Archives of pediatrics & adolescent medicine 2005;159:824.

24. Chasnoff IJ. Cocaine, pregnancy, and the neonate. Women & Health 1989;15:23-35.

25. Fries MH, Kuller JA, Norton ME, et al. Facial features of infants exposed prenatally to cocaine. Teratology 2005;48:413-20.

26. Dogra VS, Shyken J, Menon P, Poblete J, Lewis D, Smeltzer J. Neurosonographic abnormalities associated with maternal history of cocaine use in neonates of appropriate size for their gestational age. American journal of neuroradiology 1994;15:697-702.

27. Robin NH, Zackai EH. Unusual craniofacial dysmorphia due to prenatal alcohol and cocaine exposure. Teratology 2005;50:160-4.

28. Behnke M, Eyler FD, Conlon M, Casanova OQ, Woods NS. How fetal cocaine exposure increases neonatal hospital costs. Pediatrics 1997;99:204.

29. Mayes LC. Developing brain and in utero cocaine exposure: effects on neural ontogeny. Development and Psychopathology 1999;11:685-714.

30. Bandstra ES, Morrow CE, Accornero VH, Mansoor E, Xue L, Anthony JC. Estimated effects of< i> in utero</i> cocaine exposure on language development through early adolescence. Neurotoxicology and teratology 2011;33:25-35.

31. Lewis BA, Minnes S, Short EJ, et al. The effects of prenatal cocaine on language development at 10 years of age. Neurotoxicol Teratol 2011;33:17-24.

32. Schuetze P, Eiden RD, Coles CD. Prenatal cocaine and other substance exposure: Effects on infant autonomic regulation at 7 months of age. Developmental psychobiology 2007;49:276-89.

33. Lewis BA, Kirchner HL, Short EJ, et al. Prenatal cocaine and tobacco effects on children's language trajectories. Pediatrics 2007;120:e78-e85.

34. Bandstra ES, Vogel AL, Morrow CE, Xue L, Anthony JC. Severity of prenatal cocaine exposure and child language functioning through age seven years: a longitudinal latent growth curve analysis. Substance use & misuse 2004;39:25-59.

35. Nelson S, Lerner E, Needlman R, Salvator A, Singer LT. Cocaine, anemia, and neurodevelopmental outcomes in children: a longitudinal study. Journal of developmental and behavioral pediatrics: JDBP 2004;25:1.

36. Morrow CE, Bandstra ES, Anthony JC, Ofir AY, Xue L, Reyes MB. Influence of prenatal cocaine exposure on early language development: longitudinal findings from four months to three years of age. Journal of developmental and behavioral pediatrics: JDBP 2003;24:39.

37. Bandstra ES, Morrow CE, Vogel AL, et al. Longitudinal influence of prenatal cocaine exposure on child language functioning. Neurotoxicology and teratology 2002.

38. Singer LT, Arendt R, Minnes S, Salvator A, Siegel AC, Lewis BA. Developing language skills of cocaine-exposed infants. Pediatrics 2001;107:1057-64.

39. Mentis M, Lundgren K. Effects of prenatal exposure to cocaine and associated risk factors on language development. Journal of Speech, Language and Hearing Research 1995;38:1303.

40. Bender SL, Word CO, Diclemente RJ, Crittenden MR. The developmental implications of prenatal and/or postnatal crack cocaine exposure in preschool children: a preliminary report. Journal of developmental and behavioral pediatrics 1995.

41. Hawley TL, Halle TG, Drasin RE, Thomas NG. Children of addicted mothers. American Journal of Orthopsychiatry 1995;65:364-79.

42. Morrow CE, Culbertson JL, Accornero VH, Xue L, Anthony JC, Bandstra ES. Learning disabilities and intellectual functioning in school-aged children with prenatal cocaine exposure. Developmental neuropsychology 2006;30:905-31.

43. Mesulam M-M. The human frontal lobes: Transcending the default mode through contingent encoding. Principles of frontal lobe function 2002:8-30.

44. Huijbregts SC, Warren AJ, de Sonneville LM, Swaab-Barneveld H. Hot and cool forms of inhibitory control and externalizing behavior in children of mothers who smoked during pregnancy: An exploratory study. Journal of abnormal child psychology 2008;36:323-33.

45. Handley SJ, Capon A, Beveridge M, Dennis I, Evans JSB. : Working memory, inhibitory control and the development of children's reasoning. Thinking & Reasoning 2004;10:175-95.

46. Barkley RA. Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. Psychological bulletin 1997;121:65.

47. Muris P, Meesters C, Blijlevens P. Self-reported reactive and regulative temperament in early adolescence: Relations to internalizing and externalizing problem behavior and “Big Three” personality factors. Journal of Adolescence 2007;30:1035-49.

48. Frank DA, Augustyn M, Knight WG, Pell T, Zuckerman B. Growth, development, and behavior in early childhood following prenatal cocaine exposure. JAMA: the journal of the American Medical Association 2001;285:1613-25.

49. Lutiger B, Graham K, Einarson TR, Koren G. Relationship between gestational cocaine use and pregnancy outcome: A meta- analysis. Teratology 1991;44:405-14.

50. Kaltenbach K, Berghella V, Finnegan L. Opioid dependence during pregnancy: effects and management. Obstetrics and gynecology clinics of North America 1998;25:139-51.

51. Fajemirokun-Odudeyi O, Sinha C, Tutty S, et al. Pregnancy outcome in women who use opiates. European Journal of Obstetrics & Gynecology and Reproductive Biology 2006;126:170-5.

52. Binder T, Vavrinkova B. Prospective randomised comparative study of the effect of buprenorphine, methadone and heroin on the course of pregnancy, birthweight of newborns, early postpartum adaptation and course of the neonatal abstinence syndrome (NAS) in women followed up in the outpatient department. Neuro endocrinology letters 2008;29:80-6.

53. Arlettaz R, Kashiwagi M, Das- Kundu S, Fauchère JC, Lang A, Bucher HU. Methadone maintenance program in pregnancy in a Swiss perinatal center (II): neonatal outcome and social resources. Acta obstetricia et gynecologica Scandinavica 2005;84:145-50.

54. Kandall SR, Albin S, Gartner LM, Lee K-S, Eidelman A, Lowinson J. The narcotic-dependent mother: fetal and neonatal consequences. Early Human Development 1977;1:159-69.

55. Lam S, To W, Duthie S, Ma H. Narcotic addiction in pregnancy with adverse maternal and perinatal outcome. Australian and New Zealand journal of obstetrics and gynaecology 1992;32:216-21.

56. Little B, Snell L, Klein V, Gilstrap L, Knoll K, Breckenridge J. Maternal and fetal effects of heroin addiction during pregnancy. J Reprod Med 1990;35:159-62.

57. Vargas G, Pildes R, Vidyasagar D, Keith L. Effect of maternal heroin addiction on 67 liveborn neonates. Withdrawal symptoms, small body size, and small head circumference were frequent findings. Clin Pediatr 1975;14:751-3.

58. Whiting M, Whitman S, Bergner L, Patrick S. Addiction and low birth weight: a quasi-experimental study. American Journal of Public Health 1978;68:676-8.

59. Greig E, Ash A, Douiri A. Maternal and neonatal outcomes following methadone substitution during pregnancy. Archives of Gynecology and Obstetrics 2012:1-9.

60. MSSW DEJ. Wanted: a public health approach to prescription opioid abuse and diversion. Pharmacoepidemiology and drug safety 2006;15:632-4.

61. Hall AJ, Logan JE, Toblin RL, et al. Patterns of abuse among unintentional pharmaceutical overdose fatalities. JAMA: the journal of the American Medical Association 2008;300:2613-20.

62. Broussard CS, Rasmussen SA, Reefhuis J, et al. Maternal treatment with opioid analgesics and risk for birth defects. American journal of obstetrics and gynecology 2011;204:314. e1-. e11.

63. Bracken MB, HOLFORD TR. Exposure to prescribed drugs in pregnancy and association with congenital malformations. Obstetrics & Gynecology 1981;58:336-44.

64. SAXÉN I. Associations between oral clefts and drugs taken during pregnancy. International Journal of Epidemiology 1975;4:37-44.

65. Shaw GM, Todoroff K, Velie EM, Lammer EJ. Maternal illness, including fever, and medication use as risk factors for neural tube defects. Teratology 1998;57:1-7.

66. Cleary BJ, Eogan M, O'Connell MP, et al. Methadone and perinatal outcomes: a prospective cohort study. Addiction 2012.

67. Kuczkowski K. Marijuana in pregnancy. ANNALS-ACADEMY OF MEDICINE SINGAPORE 2004;33:336-9.

68. Fried P, Smith A. A literature review of the consequences of prenatal marihuana exposure: an emerging theme of a deficiency in aspects of executive function. Neurotoxicology and teratology 2001;23:1-11.

69. Arria AM, Derauf C, LaGasse LL, et al. Methamphetamine and other substance use during pregnancy: preliminary estimates from the Infant Development, Environment, and Lifestyle (IDEAL) study. Maternal and child health journal 2006;10:293-302.

70. McElhatton PR. The effects of benzodiazepine use during pregnancy and lactation. Reproductive toxicology (Elmsford, NY) 1994;8:461.

71. ROUNDS M. The fetal safety of benzodiazepines: an updated meta-analysis. J Obstet Gynaecol Can 2011;33:46-8.

72. Gilligan C, SANSON- FISHER R, Eades S, Wenitong M, Panaretto K, D'ESTE C. Assessing the accuracy of self- reported smoking status and impact of passive smoke exposure among pregnant Aboriginal and Torres Strait Islander women using cotinine biochemical validation. Drug and alcohol review 2010;29:35-40.

73. Shipton D, Tappin DM, Vadiveloo T, Crossley JA, Aitken DA, Chalmers J. Reliability of self reported smoking status by pregnant women for estimating smoking prevalence: a retrospective, cross sectional study. BMJ: British Medical Journal 2009;339.

74. Britton GRA, Brinthaupt J, Stehle JM, James GD. Comparison of Self- Reported Smoking and Urinary Cotinine Levels in a Rural Pregnant Population. Journal of Obstetric, Gynecologic, & Neonatal Nursing 2006;33:306-11.

75. Ford R, Tappin D, Schluter P, Wild C. Smoking during pregnancy: how reliable are maternal self reports in New Zealand? Journal of epidemiology and community health 1997;51:246-51.

76. Burstyn I, Kapur N, Shalapay C, et al. Evaluation of the accuracy of self-reported smoking in pregnancy when the biomarker level in an active smoker is uncertain. Nicotine & Tobacco Research 2009;11:670-8.

77. Gollenberg A, Mumford S, Cooney M, Sundaram R, Louis G. Validity of retrospectively reported behaviors during the periconception window. The Journal of reproductive medicine 2011;56:130.

78. Klebanoff MA, Levine RJ, Clemens JD, DerSimonian R, Wilkins DG. Serum cotinine concentration and self-reported smoking during pregnancy. American Journal of Epidemiology 1998;148:259-62.

79. Secker-Walker RH, Vacek PM, Flynn BS, Mead PB. Exhaled carbon monoxide and urinary cotinine as measures of smoking in pregnancy. Addictive behaviors 1997;22:671-84.

80. Rice F, Lewis A, Harold G, et al. Agreement between maternal report and antenatal records for a range of pre and peri-natal factors: the influence of maternal and child characteristics. Early Human Development 2007;83:497-504.

81. Fox NL, Sexton M, Hebel JR, Thompson B. The reliability of self-reports of smoking and alcohol consumption by pregnant women. Addictive behaviors 1989;14:187-95.

82. Hessol NA, Missett B, Fuentes-Afflick E. Lower agreement on behavioral factors than on medical conditions in self-reported data among pregnant Latina women. Archives of medical research 2004;35:241-5.

83. El Marroun H, Tiemeier H, Jaddoe V, et al. Agreement between maternal cannabis use during pregnancy according to self-report and urinalysis in a population-based cohort: the Generation R Study. European addiction research 2011;17:37-43.

84. Christmas JT, Knisely JS, Dawson KS, Dinsmoor MJ, Weber SE, Schnoll SH. Comparison of questionnaire screening and urine toxicology for detection of pregnancy complicated by substance use. Obstetrics and gynecology 1992;80:750-4.

85. Horrigan TJ, Piazza N. The Substance Abuse Subtle Screening Inventory minimizes the need for toxicology screening of prenatal patients. Journal of substance abuse treatment 1999;17:243-7.

86. Bibb KW, Stewart D, Walker JR, Cook VD, Wagener RE. Drug screening in newborns and mothers using meconium samples, paired urine samples, and interviews. Journal of perinatology: official journal of the California Perinatal Association 1995;15:199.

87. Lindsay MK, Carmichael S, Peterson H, Risby J, Williams H, Klein L. Correlation between self-reported cocaine use and urine toxicology in an inner-city prenatal population. Journal of the National Medical Association 1997;89:57.

88. Yonkers KA, Howell HB, Gotman N, Rounsaville BJ. Self-report of illicit substance use versus urine toxicology results from at-risk pregnant women. Journal of Substance Use 2011;16:372-80.

89. Wright TE, Milam KA, Rougee L, Tanaka MD, Collier AC. Agreement of umbilical cord drug and cotinine levels with maternal self-report of drug use and smoking during pregnancy. Journal of Perinatology 2010;31:324-9.

90. Ostrea Jr E, Knapp DK, Tannenbaum L, et al. Estimates of illicit drug use during pregnancy by maternal interview, hair analysis, and meconium analysis. The Journal of pediatrics 2001;138:344.

91. Tassiopoulos K, Read JS, Brogly S, et al. Substance use in HIV-Infected women during pregnancy: Self-report versus meconium analysis. AIDS and behavior 2010;14:1269-78.

92. Gray TR, LaGasse LL, Smith LM, et al. Identification of prenatal amphetamines exposure by maternal interview and meconium toxicology in the Infant Development, Environment and Lifestyle (IDEAL) study. Therapeutic drug monitoring 2009;31:769.

93. Lozano J, García- Algar O, Marchei E, et al. Prevalence of gestational exposure to cannabis in a Mediterranean city by meconium analysis. Acta Paediatrica 2007;96:1734-7.

94. Bessa MA, Mitsuhiro SS, Chalem E, Barros MM, Guinsburg R, Laranjeira R. Underreporting of use of cocaine and marijuana during the third trimester of gestation among pregnant adolescents. Addictive behaviors 2010;35:266-9.

95. Grant T, Brown Z, Callahan C, Barr H, Streissguth AP. Cocaine exposure during pregnancy: improving assessment with radioimmunoassay of maternal hair. OBSTETRICS AND GYNECOLOGY-NEW YORK- 1994;83:524-.

96. Bakhireva LN, Cano S, Rayburn WF, et al. Advanced gestational age increases serum carbohydrate-deficient transferrin levels in abstinent pregnant women. Alcohol and Alcoholism 2012;47:683-7.

97. MEDTOX.

98. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977:159-74.

99. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med 2005;37:360-3.

100. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Physical therapy 2005;85:257-68.

101. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. Journal of clinical epidemiology 1993;46:423-9.

102. Altman DG, Bland JM. Diagnostic tests. 1: Sensitivity and specificity. BMJ: British Medical Journal 1994;308:1552.

103. Musshoff F, Driever F, Lachenmeier K, Lachenmeier D, Banger M, Madea B. Results of hair analyses for drugs of abuse and comparison with self-reports and urine tests. Forensic science international 2006;156:118-23.

104. Ledgerwood DM, Goldberger BA, Risk NK, Lewis CE, Price RK. Comparison between self-report and hair analysis of illicit drug use in a community sample of middle-age men. Addictive behaviors 2008;33:1131.

105. Glintborg B, Olsen L, Poulsen H, Linnet K, Dalhoff K. Reliability of self-reported use of amphetamine, barbiturates, benzodiazepines, cannabinoids, cocaine, methadone, and opiates among acutely hospitalized elderly medical patients. Clinical Toxicology 2008;46:239-42.

106. Magura S, Goldsmith D, Casriel C, Goldstein PJ, Lipton DS. The validity of methadone clients' self-reported drug use. Substance use & misuse 1987;22:727-49.

107. Lu NT, Taylor BG, Riley KJ. THE VALIDITY OF ADULT ARRESTEE SELF-REPORTS OF CRACK COCAINE USE 1*. The American journal of drug and alcohol abuse 2001;27:399-419.

108. Hurt H, Giannetta JM, Korczykowski M, et al. Functional magnetic resonance imaging and working memory in adolescents with gestational cocaine exposure. The Journal of pediatrics 2008;152:371-7.

109. Beeghly M, Martin B, Rose-Jacobs R, et al. Prenatal cocaine exposure and children's language functioning at 6 and 9.5 years: moderating effects of child age, birthweight, and gender. Journal of pediatric psychology 2006;31:98-115.