BMJ Open

Study Protocol: Validating Screening Tools for Illicit and Prescription Drug Use in Pregnancy Using Hair and Urine Sample Testing

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020248
Article Type:	Protocol
Date Submitted by the Author:	25-Oct-2017
Complete List of Authors:	Coleman-Cowger, V; Battelle Memorial Institute Baltimore Oga, Emmanuel; Battelle Memorial Institute Baltimore Peters, Erica; Battelle Memorial Institute Baltimore Trocin, Kathleen; Battelle Memorial Institute Baltimore Koszowski, Bartosz; Battelle Memorial Institute Baltimore Mark, K; University of Maryland Medical Center
Keywords:	pregnancy, biochemical verification, NIDA Quick Screen/ASSIST, SURP-P, 4P's Plus, substance use screening

Study Protocol: Validating Screening Tools for Illicit and Prescription Drug Use in Pregnancy Using Hair and Urine Sample Testing

Coleman-Cowger, Victoria[1]; Oga, Emmanuel A.[1]; Peters, Erica N.[1]; Trocin, Kathleen; Koszowski, Bartosz[1]; Mark, Katrina[2]

1. Battelle Memorial Institute, Baltimore, MD
2. University of Maryland Medical Center, Baltimore, MD

Authors contact information:
Victoria Coleman-Cowger, PhD: colemancowger@battelle.org

Emmanuel Oga, MD, MPH: oga@battelle.org

Erica Peters, PhD: finan@battelle.org

Kathleen Trocin, MPH: trocin@battelle.org

Bartosz Koszowski, PharmD, PhD: koszowskib@battelle.org

Katrina Mark, MD: kmark@FPI.umaryland.edu

List of Abbreviations

NIDA - National Institute on Drug Abuse

ASSIST – Alcohol, Smoking, and Substance Involvement Screening Test

SURP-P - Substance Use Risk Profile-Pregnancy

EMR – Electronic Medical Records

NICU - Neonatal Intensive Care Unit

WHO – World Health Organization

HIPAA - Health Insurance Portability and Accountability Act

Abstract for Protocol

Introduction: Prescription drug use in the United States (U.S.) has increased by more than 60% in the last 3 decades. Prevalence of prescription drug use among pregnant women is currently estimated around 50%. Prevalence of illicit drug use in the U.S. is 14.6% among pregnant adolescents, 8.6% among pregnant young adults, and 3.2% among pregnant adults. The first step in identifying problematic drug use during pregnancy is screening; however, no specific substance use screener has been universally recommended for use with pregnant women to identify illicit or prescription drug use. This study compares and validates three existing substance use screeners for pregnancy - 4 P's Plus, NIDA Quick Screen/ASSIST, and the Substance Use Risk Profile-Pregnancy (SURP-P) scale.

Methods and Analysis: This is a cross-sectional study designed to evaluate the sensitivity, specificity and usability of existing substance use screeners. Recruitment occurs at two obstetric clinics in Baltimore, Maryland (USA). We are recruiting 500 participants to complete a demographics questionnaire, NIDA Quick Screen/ASSIST, 4 P's Plus, and SURP-P (ordered randomly) during their regularly scheduled prenatal appointment, then again one week later by telephone. Participants consent to multi-drug urine testing, hair drug testing, and allowing access to prescription drug and birth outcome data from electronic medical records (EMR). For each screener, reliability and validity will be assessed. Test-retest reliability analysis will be conducted by examining the results of repeated screener administrations within one week of original screener administrations for consistency via correlation analysis. Furthermore, we will assess if there are differences in the validity of each screener by age, race, and trimester.

Ethics and Dissemination: This study is approved by the Institutional Review Board (IRB) of the University of Maryland (HP-00072042), Baltimore; and Battelle Memorial Institute (0619-100106433). All participants are required to give their informed consent prior to any study procedure.

Keywords: substance use, pregnancy, screening, biochemical verification, NIDA Quick Screen/ASSIST, SURP-P, 4P's Plus

Strengths and limitations of this study

- We will conduct hair and urine analysis to assess for biochemically verified long-term and short-term drug use in large sample of 500 pregnant women.
- The study has the potential to provide validated comparisons of all three pregnancy drug use screeners currently acknowledged by the World Health Organization (WHO), and thus provide a universally acceptable evidence-based screener for drug use in pregnancy.
- The study utilizes electronic medical records (EMR) to capture prescribed drugs and birth outcome data of enrolled participants to assess for associations between drug use in pregnancy and adverse birth outcomes
- This study will rely on a convenience sample from two urban clinics rather than a national sample.

Introduction

Abuse of prescription and illicit drugs in pregnancy is a growing cause of maternal and neonatal morbidity and mortality in the United States (U.S.). According to data from the 2012 and 2013 U.S. National Survey on Drug Use and Health (NSDUH), the rate of current illicit drug use (including non-medical use of prescription drugs) in pregnant adolescents and women was 14.6% among adolescents ages 15 to 17, 8.6% among young adults (18 to 25), and 3.2% among adults (26 to 44)[1]. The consequences of this problem include spontaneous abortions, stillbirths, low birth weight, prematurity, neonatal abstinence syndrome and congenital malformations[2].

Given the relatively high frequency of provider-patient contact during the prenatal period, obstetric care providers have the unique opportunity to identify substance abuse in pregnancy. Furthermore, for pregnant women from socioeconomically disadvantaged groups, obstetricians often serve as primary care physicians and typically are the only contact these women have with the healthcare system[3]. Prenatal screening for drug use is an important way to identify drug abuse in pregnancy, as strongly recommended by the American Congress of Obstetricians and Gynecologists (ACOG)[4]. But, while validated alcohol and tobacco screeners have been recommended by the United States Preventive Services Task Force (USPSTF), there is currently no universally recommended validated screening tool for identifying illicit drug use in pregnancy.

Currently three separate, validated tools exist that screen for use of more than one substance among pregnant women: The Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST); the 4 P's Plus; and the Substance Use Risk Profile – Pregnancy (SURP-P).[5-8] The ASSIST has been validated across several populations, but it has not yet been formally validated with pregnant women[5]. A modified ASSIST, with items on tobacco and alcohol use removed, was incorporated by NIDA to their Quick Screen as a follow-up to the 4-question pre-screener, this is referred to as the NIDA Quick Screen/ASSIST. The 4 P's Plus was designed to identify drug use in pregnancy and has been validated

with pregnant women[7]. The 4P's Plus is brief but is associated with a licensing fee which may be a hindrance to widespread use. The SURP-P is a validated scale composed of three questions that can differentiate between populations of pregnant women at low-risk or high-risk for substance use[8]. The SURP-P is a simple and flexible tool for identifying possible substance use in pregnancy; however, a further screen is required for identifying those who would require treatment.

To bridge this gap and identify the most universally valid and reliable screening tool for drug abuse in pregnancy, this study aims to compare and validate three existing substance use screeners - 4 P's Plus, NIDA Quick Screen/ASSIST, and the Substance Use Risk Profile-Pregnancy (SURP-P) scale - among a cross section of 500 pregnant women presenting to two obstetrics clinics in Baltimore, Maryland (US). The overarching goal of this effort is to determine which screening tool is most effective in identifying prescription drug abuse and illicit drug use among pregnant women and acceptable among patients and clinicians so that evidence-based guidance may be offered.

Methods/Design

Specific Aims

Specific Aims of this study are to: a) conduct validity analyses to determine sensitivity, specificity, usability (test-retest reliability), and how each scale compares to the others and to the gold standard of urine and hair drug testing in identifying prescription and illicit drug use; b) determine the impact of clinic population variables (age, race, trimester of pregnancy) on validity of the three substance use screeners; and c) assess birth outcomes (birth weight, gestational age, head circumference, and Neonatal Intensive Care Unit (NICU) admissions) associated with the most widely used prescription drug and multi-drug exposure.

Study Design

This study is a cross-sectional study that evaluates the sensitivity, specificity and usability of existing substance use screeners. We chose this study design following an extensive search of the literature, an

overall assessment of feasibility and consultation with stakeholders (e.g., clinicians, pregnant women and substance use researchers). We believe that a cross-sectional study such as ours is appropriate for the evaluation of the accuracy and reliability of these screeners. We were also aided by the knowledge that the prevalence of substance use in pregnancy is high[1]. This implies that we are likely to obtain good sensitivity and specificity estimates, with narrow confidence intervals, in a cross-sectional design which is favorable in terms of cost and feasibility.

Setting

The study is being implemented at two urban obstetric clinics which serve diverse populations of pregnant women. The study plans to recruit 500 participants to complete a demographics questionnaire, followed by a randomized order of the NIDA Quick Screen/NIDA-modified ASSIST, 4 P's Plus, and SURP-P. Participants are recruited during their regularly scheduled prenatal appointment, then contacted again one week later by telephone to re-administer the screeners. Participants consent to multi-drug urine testing, hair drug testing, and access to prescription drug and birth outcome data from electronic medical records (EMR).

Recruitment Sites. We are recruiting participants from two obstetric outpatient clinics. Currently all obstetric patients are screened for use of drugs, alcohol and tobacco at their first prenatal visit by medical staff. Additionally, all new obstetric patients receive an in-depth evaluation by a social worker which includes a more detailed assessment of both substance use and mental health disorder history.

In the first clinic, which is the larger of the two clinics, most patients (97%) are publicly insured with medical assistance and are over the age of 20 (80%). This clinic's population is primarily African-American and low-income, all of whom undergo urine toxicological screening for substance use identification. Based on preliminary data obtained from the clinic, about 950 individual obstetric patients are cared for at this clinic annually. In the second (smaller) clinic, approximately 500 pregnant women are cared for annually. Most patients (87%) have commercial insurance and 13% have either medical assistance or Medicare. Most are over the age of 20 years (90%). Due to varying insurance coverage for

urine toxicology screens, patients in this office do not universally undergo urine toxicology screening but all are screened for drug use using various interview techniques by their obstetric care providers at their first prenatal visit. Based on historical data, we expect about 500 individual obstetric patients to be cared for in this clinic across all trimesters of pregnancy in the one year of study recruitment.

Across both study sites, our source population covers a diverse set of participants and captures pregnant women across all socioeconomic categories, insurance types, ethnicities and drug use patterns. This ensures that our study results are generalizable to most populations of pregnant women.

Study Population

In the first clinic, of the estimated 950 individual obstetric patients cared for at this clinic annually, we anticipated approaching 403 (50%), and expected 322 (80%) or more to agree to participate in this study.

In the second clinic, of the approximately 500 pregnant women cared for annually, we expect at least 450 (90%) to meet eligibility criteria. We anticipate approaching 225 pregnant women (50%) and expect 180 (80%) or more to agree to participate in this study.

Expected participation percentages are based on a similar grant-funded study that recruited pregnant smokers from the same population and required consent for urine testing (cotinine) and birth data abstraction from EMR.

Participant eligibility criteria include the following: a) currently pregnant (pre-determined by clinic staff); b) age 18 or older; c) able to speak and understand English sufficiently to provide informed consent; and d) natural hair length at least 3 cm to allow for substance use testing.

If eligibility criteria are met, research staff then obtain informed consent and medical releases for urine collection, hair drug testing, and prescription drug and birth outcome data abstraction from the EMR.

Ethical Approval

This study is approved by the Institutional Review Boards (IRB) of the University of Maryland, Baltimore; and Battelle Memorial Institute. All participants are required to give their informed consent prior to any study procedure. All research staff complete ethics training annually.

Study Procedures

Approach

All patients entering the clinics for prenatal appointments are approached by research staff at check-in and asked to read a brief description of the study to determine their interest in participating (excluding those previously approached). Research staff keep track of which patients have been approached already to avoid repetitive recruitment efforts. The study description includes a section requesting basic demographic information (if they would allow its use for anonymous, grouped analysis) and at the bottom asks potential participants to note their interest and return to clinic staff. There are checkboxes for "not interested" (with additional space beneath for noting reasons for lack of interest) and "interested in learning more." Patients who are not interested in the study are not to be contacted further; however, the basic demographic information provided is used for comparative analyses with study participants to assess for selection bias. If a patient expresses interest, the research staff approaches her as she waits for her prenatal appointment either on the same day or at a future prenatal appointment.

Recruitment

At the enrollment visit, the staff escorts potential participants from the waiting area to a private room, further describes the study and determines whether potential participants meet all eligibility criteria. If eligibility criteria are met, the staff obtains informed consent and HIPAA (Health Insurance Portability and Accountability Act) authorization (for urine collection, hair drug testing, and prescription drug and birth outcome data abstraction from the EMR). Women who refuse to participate are thanked for their time and no further contact is made. The research visit takes 20-30 minutes. Enrolled participants are compensated for their time using a reloadable gift card for their time. The typical patient wait time to see

medical staff at each clinic is 30 minutes to 1 hour, so data collection does not typically interfere with medical visits. See Figure 1 for study procedures.

Self-Report Measures

Participants complete a demographics questionnaire. Afterwards, the NIDA Quick Screen/NIDA-modified Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST), 4 P's Plus, and Substance Use Risk Profile-Pregnancy (SURP-P) surveys are administered on a Wi-Fi enabled iPad Pro through SurveyMonkey (i.e., online survey software). These surveys are assigned to participants in a random sequence; this randomization service is provided by SurveyMonkey. The questions are read aloud by the interviewer and entered directly into SurveyMonkey so that electronic submission is instantaneous and data can be obtained by the research team at any time.

Table 1: Study Instruments

Instrument	Description/Construct	Use in Study
Demographic Questionnaire	20-item questionnaire that collects demographic and general information such as age, marital status, education, employment status, ethnicity and reproductive history	Enrollment
NIDA Quick Screen/ ASSIST	9-item combined NIDA Quick Screen and modified-ASSIST to screen for tobacco, alcohol and illicit drugs	Enrollment, 1-week follow up
4P's Plus	4-item screener for alcohol and general substance use	Enrollment, 1-week follow up
SURP-P	3-item screener for alcohol and substances	Enrollment, 1-week follow up

Biochemical Measures

Participants are asked to consent that urine collected for their prenatal appointment that day is also tested for various drugs by research staff (Table 2). If sufficient urine is unavailable for testing, participants are given bottled water and asked to provide another sample prior to leaving the clinic. Participants must also consent to hair testing, which involves the cutting of approximately 100 strands of hair from the crown of the head (or other body hair if head hair is unavailable). Samples are then shipped to an external laboratory same-day for drug testing utilizing mass spectrometry.

Table 2: Drug detection windows and cutoffs for urine and hair testing

	Drug class	Detection	Confirmation
URINE	Cocaine COC	2-4 Days	300 ng/mL
	Marijuana THC	15-30 Days	50 ng/mL
	Opiates OPI	2-4 Days	2000 ng/mL
	Amphetamines AMP	2-4 Days	1000 ng/mL
	Methamphetamines AMP	3-5 Days	1000 ng/mL
	Phencyclidine PCP	7-14 Days	25 ng/mL
	Benzodiazepines BZO	3-7 Days	300 ng/mL
	Barbiturates BAR	4-7 Days	300 ng/mL
	Methadone MTD	3-5 Days	300 ng/mL
	Tricyclic Antidepressants TCA		1,000 ng/mL
	Oxycodone	2-4 Days	100 ng/mL
	Propoxyphene	1-2 Days	300 ng/mL
	Buprenorphine BUP (Suboxone, Subutex)	2-3 Days	10 ng/mL
HAIR	Marijuana THC	Up to 90 days	
	Amphetamines AMP	Up to 90 days	
	Cocaine COC	Up to 90 days	
	Opiates OPI	Up to 90 days	
	Phencyclidine PCP	Up to 90 days	

All women who screen positive on either biological multi-drug test or any one of the screeners are contacted immediately (for urine and screener results) or within 72 hours (for hair results) to detail the results of her test, encourage the participant to talk with her physician about her substance use, and offer her referrals to community resources for treatment that mirror what is currently given to patients by

medical staff in each clinic. They are encouraged to speak with the on-site clinic social worker who can provide further support.

Birth Outcome Measures

Birth outcome data, including miscarriage, stillbirth, birth weight, gestational age, head circumference, and NICU admissions, as well as a list of drugs prescribed during pregnancy and their dosage are collected by research staff via the EMR and entered into SurveyMonkey.

Participant Follow-Up

After completion of this research visit, participants are contacted once more by telephone one week after completing the surveys to complete the three screeners again to assess test-retest reliability. The average time commitment for the call is about 10-15 minutes, and upon completion $25 is loaded onto the reloadable gift card provided the week prior.

Pilot Study

To examine the recruitment process and determine acceptability from the target population of substance-using pregnant women prior to the start of the study, we conducted a one-month pilot study. Each step of the recruitment process was reviewed to determine where improvements could be made.

We recruited 21 participants from each site for a total of 42 participants (Table 3). Mean age (sd) of participants was 30.1 years (5.64). By race, 11 participants (26.2%) were White, 25 (59.5%) were Black/African American, 4 (9.5%) were Asian, 1 (2.4%) was Hispanic and 1 (2.4%) was Other. About 24.4% tested positive for illicit drugs on urine testing, 22% tested positive on hair sample testing. Seven (7) participants (16.7%) were lost to follow up.

235

236 *Table 3 Pilot Study Participant Characteristics*

Characteristics, N = 42	**Clinical Site**		
	Clinic 1	***Clinic 2***	***Both Sites***
Number of participants	21	21	42
Participant age in years, mean (SD)	27.10 (5.09)	33.05 (4.54)	30.07 (5.64)
Ethnicity n (%)			
African American/Black	18 (85.7)	7 (33.3)	25 (59.5)
Asian	0 (0.0)	4 (19.0)	4 (9.5)
Caucasian/White	2 (9.5)	9 (42.9)	11 (26.2)
Hispanic, Latino or Chicano	0 (0.0)	1 (4.8)	1 (2.4)
Some other group	1 (4.8)	0 (0.0)	1 (2.4)
Trimester n (%)			
1st	2 (9.5)	2 (9.5)	4 (9.5)
2nd	6 (28.6)	6 (28.6)	12 (28.6)
3rd	13 (61.9)	13 (61.9)	26 (61.9)
Urine Results n (%)			
Negative for all substances	15 (71.4)	16 (80.0)	31 (75.6)
Positive for at least 1 substance	6 (28.6)	4 (20.0)	10 (24.4)
Hair Results n (%)			
Negative for all substances	12 (60.0)	18 (85.7)	30 (73.2)
Positive for at least one substance	7 (35.0)	2 (9.5)	9 (22.0)
Invalid	1 (5.0)	1 (4.8)	2 (4.9)
Study Disposition n (%)			
Study completes	19 (90.48)	16 (76.2)	35 (83.3)

Lost to follow-up	2 (9.52)	5 (23.8)	7 (16.7)

Results from the pilot study confirmed the feasibility of this study. Eligibility criteria did not appear too restrictive, given the eligibility rate of 78% (although slightly lower than anticipated) (Figure 2). Overall, there was good comprehension of surveys, a low refusal rate for hair sampling (1 refusal/95 approached, 1.1%), and high study enrollment (Figure 2). The recruitment process took an average of 40 minutes.

Power and Sample Size

The sample size of 500 participants was chosen based on power analyses for the primary study questions. Based on a one-sample binomial approach, with a sample size of 500 participants, we can be 95% confident that the false negative rate in the population is under 10% (assuming no more than 35 individuals test positive in the biologic drug tests without a positive survey screener result). Similarly, we can be 95% confident that the false negative rate in the population is under 5% (assuming no more than 15 individuals test positive in the urine drug test without a positive survey screen result in the study). According to McNemar's Test, if at least 15% of the study participants have disagreement between any pair of survey results, 500 is a sufficient sample size to determine significant disagreement.

After a preliminary sample size of 500 was chosen, a power analysis was conducted to determine the detectable differences in age, race, and trimester with a sample size of 500. The power of the test of proportions was calculated based on the difference in the proportion of false negatives in each age group, race, and trimester of pregnancy. Assuming recruitment of an equal number of women aged 18 to 25 years and women 26 and older, and that their respective positive screener results are 20% and 10%, then the power to detect that difference is 0.88. If the respective screener results are 15% and 20%, then the power is much lower (0.31). If we further assume that recruitment of 23% White women and 77% non-White women and that white women have a false negative rate of 5% and non-White women have a false negative rate of 15%, then the power is high (0.87). Similarly, if we assume recruitment of an equal

number of women in each of the three trimesters of pregnancy and that women in one trimester have a false negative rate of 20% while women in another trimester have a false negative rate of 35%, then the power is high (0.87).

Analysis

For each screener, reliability and validity (convergent/discriminant validity) will be assessed, including calculating correlation coefficients between each pair of screeners and between each screener and the appropriate biologic drug tests. Test-retest reliability analysis will be conducted by examining the results of repeated screener administrations within one week of original screener administrations for consistency via correlation analysis. The sensitivity and specificity of each instrument will be calculated, presented, and interpreted. Each survey instrument will be compared to the gold standard (hair and urine sample drug testing) by comparing the false negative rates to a predetermined limit of acceptability. If the upper one-sided 95% binomial confidence interval around the false negative rate in the sample is less than that limit, then the survey instrument is considered acceptable. The 4P's Plus and SURP-P survey screeners will be compared to both urine and hair testing in the assessment of their sensitivity and specificity in relation to short and long-term drug use, respectively. The NIDA Quick Screen/NIDA-Modified ASSIST screener will be compared to urine testing, while particular questions from the screener regarding long-term drug use will be compared to hair testing.

Furthermore, we will assess if there are differences in the validity of each screener by age, race, and trimester. The false negative rate for each screener will be presented by age, race, and trimester. A two-sided test of proportions will be conducted to test for significant differences in false negative rates between age, race, and trimester for each screener. Chi-square tests (or Fisher's exact tests if subgroup sizes are small) may be conducted to determine whether the distribution of responses on each survey instrument is similar for age, race, and trimester. To examine differences in screener validity by age, race, and trimester, logistic regression models will be fitted to the data. To separately analyze differences in probability of false positive results and false negative results on each survey, data will be stratified by

screener and screener result (positive or negative) for a total of six models. In each model, the dependent variable will be coded 1 for invalid screener result (false negative or false positive) and 0 for valid screener result (true negative or true positive). Independent variables for age, race, and trimester will be added to the models to test whether they have a significant effect on the probability of an invalid screener result. Two-way interaction terms will be included in the model if they are found to be significant effects. In order to stratify results by trimester, if trimester or any two-way interaction term including trimester is a significant effect in the models for any of the screeners, probabilities of false positive/ false negative result will be presented separately by each trimester.

Finally, the prevalence of prescription and illicit drug use will be calculated based on hair test results and self-report. Prevalence of multi-drug exposure will also be calculated. An ANOVA model will be fitted to the data with a fixed effect for drug use (negative, positive, positive for multi-drug exposure) to test for significant differences in birthweight, gestational age, and head circumference based on participant hair drug tests result. Significant differences will be noted and discussed. The relative risk of NICU admission, stillbirth, and miscarriage will be examined. A risk ratio will be calculated and will quantify the percentage difference in these three variables between those with positive hair drug tests versus negative drug test. The risk ratio takes on values between zero and infinity. A risk ratio of one means that there is no difference in NICU admissions, stillbirth, or miscarriage between the participants' biologic drug tests results. A risk ratio very small (close to zero) or very large means a large difference between NICU admissions, stillbirth, or miscarriage based on the hair drug tests results. Approximate 95% confidence intervals for the relative risk will be calculated. The same relative risk ratios and 95% confidence intervals will be calculated for a positive biologic drug tests for multi-drug exposure versus positive for a single-drug exposure. Further, relative risk ratios will be computed with 95% confidence intervals stratified by trimester.

Discussion

Our ongoing research has five aspects of significance. First, the importance of screening pregnant women and the public health impact of the current research is tied directly to the negative health consequences associated with illicit and prescription drug use during pregnancy. Second, it utilizes both urine and hair testing to enable us to examine past 90-day substance use history with precision. Hair analysis provides nearly twice the number of positives due to its longer detection window, but often cannot capture very recent use. Urine analysis supplements hair analysis to allow for the most comprehensive validation of screeners possible. Third, the study compares all three screeners acknowledged by the World Health Organization (WHO) to screen for multiple substances to each other and to the biological screeners (gold standard). This is the first study to conduct a direct, head-to-head comparison of multiple screening tools for prescription and illicit drug use among pregnant women, while also utilizing biologic measures as a gold standard against which to compare. Fourth, the study utilizes electronic medical records (EMR) to capture prescribed drugs and birth outcome data of enrolled participants. The ability to access a participant's prescription drug orders enables better tracking and distinction between prescription drug use and abuse, while birth outcome data allows for determination of associations between specific drug use and birth outcomes. Fifth, the study has the potential to shift clinical practice towards universal standardized substance use screening.

The primary innovation of this project is that it may provide a final evidence-based recommendation for the tool(s) best suited for screening for illicit and prescription drugs among a diverse sample of pregnant women. The provision of this evidence-based guidance to clinicians is a concrete application of findings that is rare in public health research.

Substance use during pregnancy, and specifically prescription and illicit drug use, are high priority topics for the Centers for Disease Control and Prevention (CDC), WHO, The American Congress of Obstetricians and Gynecologists (ACOG), Substance Abuse and Mental Health Services

Administration (SAMHSA), National Institute on Drug Abuse (NIDA), and the National Institutes of Health (NIH). Universal screening has the potential to greatly enhance maternal and infant health outcomes and reduce healthcare costs. Specifically, the current research supports the following Healthy People 2020 public health goals and objectives which include reducing maternal illness and complications due to pregnancy; increasing the proportion of pregnant women who receive adequate prenatal care' increasing abstinence from alcohol, cigarettes, and illicit drugs among pregnant women; and increasing the proportion of women delivering a live birth who received preconception care services and practiced key recommended preconception health behaviors.

This research addresses an important problem by identifying a valid substance use screening instrument for illicit and prescription drugs among pregnant women that is accurate, brief, and acceptable to both patients and health care providers in a primary care setting. Identifying and validating one instrument that functions the closest to the "gold standard" of biologic testing (i.e., urine and hair) and disseminating this information widely will increase the likelihood that primary care clinics nationwide may adopt a quick and easy screener universally. We may find that one instrument does not stand out but that each has its distinct advantages and disadvantages; in this case, the performance of each measure will be detailed with recommendations for which screener may work the best with a given population.

Declarations

Ethics approval and consent to participate

This study is approved by the Institutional Review Boards (IRB) of the University of Maryland (HP-00072042), Baltimore; and Battelle Memorial Institute (0619-100106433). All participants are required to give their informed consent prior to any study procedure.

Availability of data and materials

The data that support the findings of this study are available on request from the corresponding author [VHCC]. The data are not publicly available due to them containing information that could compromise research participant privacy/consent.

Competing interests

The authors declare that there are no competing interests.

Authors' contributions

VCC conceived the study. VCC, EO, EP, KT, BK and KM participated in the drafting of the manuscript and each approved the final draft

Funding

The research reported in this article is supported by the National Institute on Drug Abuse (NIDA) of the National Institutes of Health (NIH) grant under Award Number R01DA041328 (PI-Coleman-Cowger). The content is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health.

REFERENCES

1. **Substance Abuse and Mental Health Services Administration (SAMHSA). Results from the 2012 National Survey on Drug Use and Health: Summary of national findings**. In: *NSDUH Series H-46, HHS Publication No(SMA) 13-4795.* edn.: Substance Abuse and Mental Health Services Administration Rockville, MD; 2013.
2. **World Health Organization. Guidelines for the identification and management of substance use and substance use disorders in pregnancy**. 2014.
3. Scholle SH, Kelleher K: **Assessing primary care performance in an obstetrics/gynecology clinic**. *Women & health* 2003, **37**(1):15-30.
4. **ACOG Committee on Health Care for Underserved Women. ACOG Committee Opinion No. 524: opioid abuse, dependence, and addiction in pregnancy**. *Obstetric Anesthesia Digest* 2013, **33**(2):79-80.
5. Group W: **The alcohol, smoking and substance involvement screening test (ASSIST): development, reliability and feasibility**. *Addiction* 2002, **97**(9):1183-1194.
6. Smith PC, Schmidt SM, Allensworth-Davies D, Saitz R: **A single-question screening test for drug use in primary care**. *Archives of internal medicine* 2010, **170**(13):1155-1160.
7. Chasnoff I, Wells A, McGourty R, Bailey L: **Validation of the 4P's Plus© screen for substance use in pregnancy validation of the 4P's Plus**. *Journal of Perinatology* 2007, **27**(12):744.
8. Yonkers KA, Gotman N, Kershaw T, Forray A, Howell HB, Rounsaville BJ: **Screening for prenatal substance use: development of the Substance Use Risk Profile-Pregnancy scale**. *Obstetrics and gynecology* 2010, **116**(4):827.

FIGURE 1: Study procedures

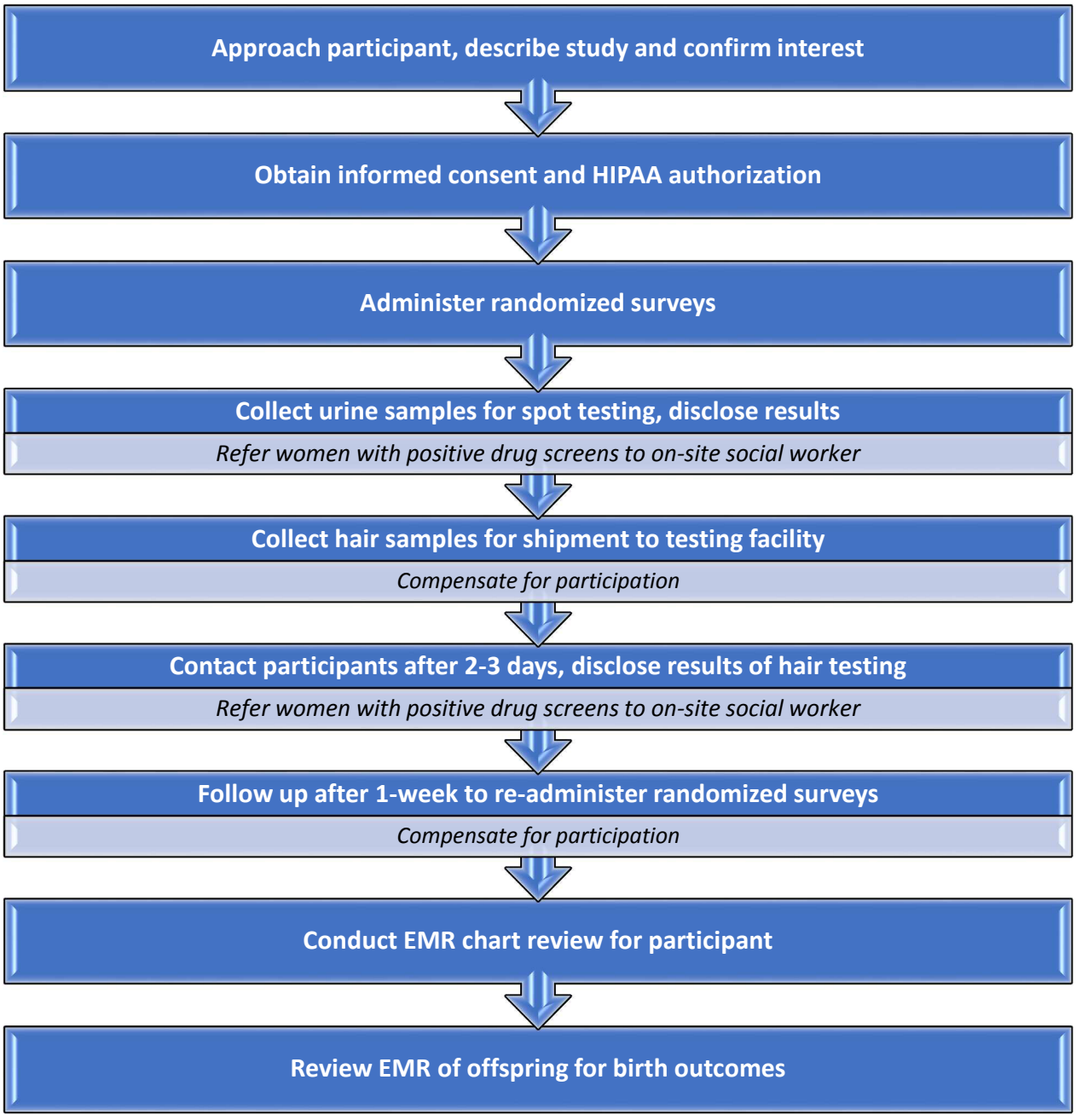

Figure 1: Study procedures

FIGURE 2: Pilot study participation

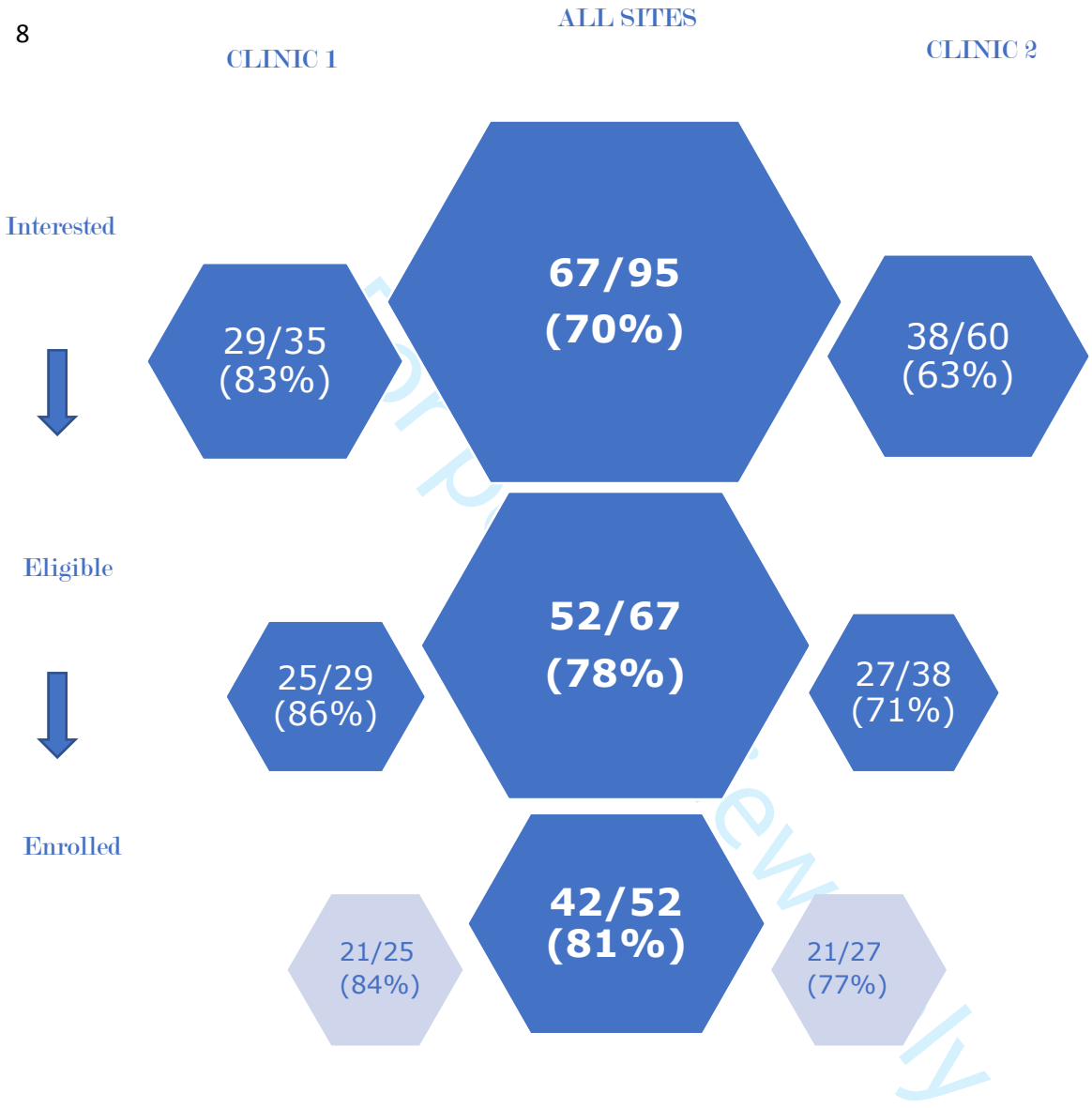

Figure 2: Pilot study participation

BMJ Open

Comparison and Validation of Screening Tools for Substance Use in Pregnancy: A Cross-Sectional Study conducted in Maryland Prenatal Clinics

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020248.R1
Article Type:	Protocol
Date Submitted by the Author:	20-Dec-2017
Complete List of Authors:	Coleman-Cowger, V; Battelle Memorial Institute Baltimore Oga, Emmanuel; Battelle Memorial Institute Baltimore Peters, Erica; Battelle Memorial Institute Baltimore Trocin, Kathleen; Battelle Memorial Institute Baltimore Koszowski, Bartosz; Battelle Memorial Institute Baltimore Mark, K; University of Maryland Medical Center
Primary Subject Heading:	Public health
Secondary Subject Heading:	Addiction, Public health
Keywords:	pregnancy, biochemical verification, NIDA Quick Screen/ASSIST, SURP-P, 4P's Plus, substance use screening

Comparison and Validation of Screening Tools for Substance Use in Pregnancy: A Cross-Sectional Study conducted in Maryland Prenatal Clinics

Coleman-Cowger, Victoria[1]; Oga, Emmanuel A.[1]; Peters, Erica N.[1]; Trocin, Kathleen; Koszowski, Bartosz[1]; Mark, Katrina[2]

1. Battelle Memorial Institute, Baltimore, MD
2. University of Maryland Medical Center, Baltimore, MD

Authors contact information:
Victoria Coleman-Cowger, PhD: colemancowger@battelle.org
Emmanuel Oga, MD, MPH: oga@battelle.org
Erica Peters, PhD: finan@battelle.org
Kathleen Trocin, MPH: trocin@battelle.org
Bartosz Koszowski, PharmD, PhD: koszowskib@battelle.org
Katrina Mark, MD: kmark@FPI.umaryland.edu

List of Abbreviations

NIDA - National Institute on Drug Abuse
ASSIST – Alcohol, Smoking, and Substance Involvement Screening Test
SURP-P - Substance Use Risk Profile-Pregnancy
EMR – Electronic Medical Records
NICU - Neonatal Intensive Care Unit
WHO – World Health Organization
HIPAA - Health Insurance Portability and Accountability Act

Abstract for Protocol

Introduction: Prescription drug use in the United States (U.S.) has increased by more than 60% in the last 3 decades. Prevalence of prescription drug use among pregnant women is currently estimated around 50%. Prevalence of illicit drug use in the U.S. is 14.6% among pregnant adolescents, 8.6% among pregnant young adults, and 3.2% among pregnant adults. The first step in identifying problematic drug use during pregnancy is screening; however, no specific substance use screener has been universally recommended for use with pregnant women to identify illicit or prescription drug use. This study compares and validates three existing substance use screeners for pregnancy - 4 P's Plus, NIDA Quick Screen/ASSIST, and the Substance Use Risk Profile-Pregnancy (SURP-P) scale.

Methods and Analysis: This is a cross-sectional study designed to evaluate the sensitivity, specificity and usability of existing substance use screeners. Recruitment occurs at two obstetric clinics in Baltimore, Maryland (USA). We are recruiting 500 participants to complete a demographics questionnaire, NIDA Quick Screen/ASSIST, 4 P's Plus, and SURP-P (ordered randomly) during their regularly scheduled prenatal appointment, then again one week later by telephone. Participants consent to multi-drug urine testing, hair drug testing, and allowing access to prescription drug and birth outcome data from electronic medical records (EMR). For each screener, reliability and validity will be assessed. Test-retest reliability analysis will be conducted by examining the results of repeated screener administrations within one week of original screener administrations for consistency via correlation analysis. Furthermore, we will assess if there are differences in the validity of each screener by age, race, and trimester.

Ethics and Dissemination: This study is approved by the Institutional Review Board (IRB) of the University of Maryland (HP-00072042), Baltimore; and Battelle Memorial Institute (0619-100106433). All participants are required to give their informed consent prior to any study procedure.

Keywords: substance use, pregnancy, screening, biochemical verification, NIDA Quick Screen/ASSIST, SURP-P, 4P's Plus

Strengths and limitations of this study

- This study will provide insight into the substance use screener(s) that works best to identify illicit drug use and prescription drug misuse during pregnancy, utilizing hair and urine analysis for biochemical verification of long-term and short-term substance use in a convenience sample of 500 pregnant women.
- The study will provide evidence of screener usefulness and acceptability in prenatal clinic settings that could inform United States Preventive Services Task Force (USPSTF) recommendations for substance use screening during pregnancy.
- The study utilizes electronic medical records (EMR) to capture prescribed drugs and birth outcome data of enrolled participants to assess for associations between substance use in pregnancy and adverse birth outcomes.
- A limitation of this study is the reliance on a convenience sample from two urban clinics rather than a national sample.
- Findings from this study will not be generalizable to pregnant adolescents who were not included in our study sample.

Introduction

Abuse of prescription and illicit drugs in pregnancy is a growing cause of maternal and neonatal morbidity and mortality in the United States (U.S.). According to data from the 2012 and 2013 U.S. National Survey on Drug Use and Health (NSDUH), the rate of current illicit drug use (including non-medical use of prescription drugs) in pregnant adolescents and women was 14.6% among adolescents ages 15 to 17, 8.6% among young adults (18 to 25), and 3.2% among adults (26 to 44)[1]. The consequences of this problem include spontaneous abortions, stillbirths, low birth weight, prematurity, neonatal abstinence syndrome and congenital malformations[2].

Given the relatively high frequency of provider-patient contact during the prenatal period, obstetric care providers have the unique opportunity to identify substance abuse in pregnancy. Furthermore, for pregnant women from socioeconomically disadvantaged groups, obstetricians often serve as primary care physicians and typically are the only contact these women have with the healthcare system[3]. Prenatal screening for drug use is an important way to identify drug abuse in pregnancy, as strongly recommended by the American Congress of Obstetricians and Gynecologists (ACOG)[4]. But, while validated alcohol and tobacco screeners have been recommended by the United States Preventive Services Task Force (USPSTF), there is currently no universally recommended validated screening tool for identifying illicit drug use in pregnancy.

Currently three separate, validated tools exist that screen for use of more than one substance among pregnant women: The Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST); the 4 P's Plus; and the Substance Use Risk Profile – Pregnancy (SURP-P).[5-8] The ASSIST has been validated across several populations, but it has not yet been formally validated with pregnant women[5]. A modified ASSIST, with items on tobacco and alcohol use removed, was incorporated by NIDA to their Quick Screen as a follow-up to the 4-question pre-screener, this is referred to as the NIDA Quick Screen/ASSIST. The 4 P's Plus was designed to identify drug use in pregnancy and has been validated

with pregnant women[7]. The 4P's Plus is brief but is associated with a licensing fee which may be a hindrance to widespread use. The SURP-P is a validated scale composed of three questions that can differentiate between populations of pregnant women at low-risk or high-risk for substance use[8]. The SURP-P is a simple and flexible tool for identifying possible substance use in pregnancy; however, a further screen is required for identifying those who would require treatment.

To bridge this gap and identify the most universally valid and reliable screening tool for drug abuse in pregnancy, this study aims to compare and validate three existing substance use screeners - 4 P's Plus, NIDA Quick Screen/ASSIST, and the Substance Use Risk Profile-Pregnancy (SURP-P) scale - among a cross section of 500 pregnant women presenting to two obstetrics clinics in Baltimore, Maryland (US). The overarching goal of this effort is to determine which screening tool is most effective in identifying prescription drug abuse and illicit drug use among pregnant women and acceptable among patients and clinicians so that evidence-based guidance may be offered.

Methods/Design

Specific Aims

Specific Aims of this study are to: a) conduct validity analyses to determine sensitivity, specificity, usability (test-retest reliability), and how each scale compares to the others and to the gold standard of urine and hair drug testing in identifying prescription and illicit drug use; b) determine the impact of clinic population variables (age, race, trimester of pregnancy) on validity of the three substance use screeners; and c) assess birth outcomes (birth weight, gestational age, head circumference, and Neonatal Intensive Care Unit (NICU) admissions) associated with the most widely used prescription drug and multi-drug exposure.

Study Design

This study is a cross-sectional study that evaluates the sensitivity, specificity and usability of existing substance use screeners. We chose this study design following an extensive search of the literature, an

overall assessment of feasibility and consultation with stakeholders (e.g., clinicians, pregnant women and substance use researchers). We believe that a cross-sectional study such as ours is appropriate for the evaluation of the accuracy and reliability of these screeners. We were also aided by the knowledge that the prevalence of substance use in pregnancy is high[1]. This implies that we are likely to obtain good sensitivity and specificity estimates, with narrow confidence intervals, in a cross-sectional design which is favorable in terms of cost and feasibility.

Setting

The study is being implemented at two urban obstetric clinics which serve diverse populations of pregnant women. The study plans to recruit 500 participants to complete a demographics questionnaire, followed by a randomized order of the NIDA Quick Screen/NIDA-modified ASSIST, 4 P's Plus, and SURP-P. Participants are recruited during their regularly scheduled prenatal appointment, then contacted again one week later by telephone to re-administer the screeners. Participants consent to multi-drug urine testing, hair drug testing, and access to prescription drug and birth outcome data from electronic medical records (EMR).

Recruitment Sites. We are recruiting participants from two obstetric outpatient clinics from January 2017-January 2018. Currently all obstetric patients are screened for use of drugs, alcohol and tobacco at their first prenatal visit by medical staff. Additionally, all new obstetric patients receive an in-depth evaluation by a social worker which includes a more detailed assessment of both substance use and mental health disorder history.

In the first clinic, which is the larger of the two clinics, most patients (97%) are publicly insured with medical assistance and are over the age of 20 (80%). This clinic's population is primarily African-American and low-income, all of whom undergo urine toxicological screening for substance use identification. Based on preliminary data obtained from the clinic, about 950 individual obstetric patients are cared for at this clinic annually. In the second (smaller) clinic, approximately 500 pregnant women are cared for annually. Most patients (87%) have commercial insurance and 13% have either medical

assistance or Medicare. Most are over the age of 20 years (90%). Due to varying insurance coverage for urine toxicology screens, patients in this office do not universally undergo urine toxicology screening but all are screened for drug use using various interview techniques by their obstetric care providers at their first prenatal visit. Based on historical data, we expect about 500 individual obstetric patients to be cared for in this clinic across all trimesters of pregnancy in the one year of study recruitment.

Across both study sites, our source population covers a diverse set of participants and captures pregnant women across all socioeconomic categories, insurance types, ethnicities and drug use patterns. This ensures that our study results are generalizable to most populations of pregnant women.

Study Population

In the first clinic, of the estimated 950 individual obstetric patients cared for at this clinic annually, we anticipated approaching 403 (50%), and expected 322 (80%) or more to agree to participate in this study.

In the second clinic, of the approximately 500 pregnant women cared for annually, we expect at least 450 (90%) to meet eligibility criteria. We anticipate approaching 225 pregnant women (50%) and expect 180 (80%) or more to agree to participate in this study.

Expected participation percentages are based on a similar grant-funded study that recruited pregnant smokers from the same population and required consent for urine testing (cotinine) and birth data abstraction from EMR.

Participant eligibility criteria include the following: a) currently pregnant (pre-determined by clinic staff); b) age 18 or older; c) able to speak and understand English sufficiently to provide informed consent; and d) natural hair length at least 3 cm to allow for substance use testing.

If eligibility criteria are met, research staff then obtain informed consent and medical releases for urine collection, hair drug testing, and prescription drug and birth outcome data abstraction from the EMR.

Ethics and Dissemination

This study is approved by the Institutional Review Boards (IRB) of the University of Maryland, Baltimore (HP-00072042); and Battelle Memorial Institute (0619-100106433). All participants are required to give their informed consent prior to any study procedure. All research staff complete ethics training via the Collaborative Institutional Training Initiative (CITI) annually.

Study Procedures

Approach

All patients entering the clinics for prenatal appointments are approached by research staff at check-in and asked to read a brief description of the study to determine their interest in participating (excluding those previously approached). Research staff keep track of which patients have been approached already to avoid repetitive recruitment efforts. The study description includes a section requesting basic demographic information (if they would allow its use for anonymous, grouped analysis) and at the bottom asks potential participants to note their interest and return to clinic staff. There are checkboxes for "not interested" (with additional space beneath for noting reasons for lack of interest) and "interested in learning more." Patients who are not interested in the study are not to be contacted further; however, the basic demographic information provided is used for comparative analyses with study participants to assess for selection bias. If a patient expresses interest, the research staff approaches her as she waits for her prenatal appointment either on the same day or at a future prenatal appointment.

Recruitment

At the enrollment visit, the staff escorts potential participants from the waiting area to a private room, further describes the study and determines whether potential participants meet all eligibility criteria. If eligibility criteria are met, the staff obtains informed consent and HIPAA (Health Insurance Portability and Accountability Act) authorization (for urine collection, hair drug testing, and prescription drug and birth outcome data abstraction from the EMR). Women who refuse to participate are thanked for their time and no further contact is made. The research visit takes 20-30 minutes. Enrolled participants are

compensated for their time using a reloadable gift card for their time. The typical patient wait time to see medical staff at each clinic is 30 minutes to 1 hour, so data collection does not typically interfere with medical visits. See Figure 1 for study procedures.

Self-Report Measures

Participants complete a demographics questionnaire. Afterwards, the NIDA Quick Screen/NIDA-modified Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST), 4 P's Plus, and Substance Use Risk Profile-Pregnancy (SURP-P) surveys are administered on a Wi-Fi enabled iPad Pro through SurveyMonkey (i.e., online survey software) See Table 1 for description of surveys and the timing of administration during the study. These surveys are assigned to participants in a random sequence; this randomization service is provided by SurveyMonkey. The questions are read aloud by the interviewer and entered directly into SurveyMonkey so that electronic submission is instantaneous, and data can be obtained by the research team at any time.

Table 1: Study Instruments

Instrument	Description/Construct	Use in Study
Demographic Questionnaire	20-item questionnaire that collects demographic and general information such as age, marital status, education, employment status, ethnicity and reproductive history	Enrollment
NIDA Quick Screen/ ASSIST	9-item combined NIDA Quick Screen and modified-ASSIST to screen for tobacco, alcohol and illicit drugs	Enrollment, 1-week follow up
4P's Plus	4-item screener for alcohol and general substance use	Enrollment, 1-week follow up
SURP-P	3-item screener for alcohol and substances	Enrollment, 1-week follow up

Biochemical Measures

Participants are asked to consent that urine collected for their prenatal appointment that day is also tested for various drugs by research staff (Table 2). If sufficient urine is unavailable for testing, participants are given bottled water and asked to provide another sample prior to leaving the clinic. Participants must also consent to hair testing, which involves the cutting of approximately 100 strands of hair from the crown of the head (or other body hair if head hair is unavailable). Samples are then shipped to an external laboratory same-day for drug testing utilizing mass spectrometry.

Table 2: Drug detection windows and cutoffs for urine and hair testing

	Drug class	Detection	Confirmation
URINE	Cocaine COC	2-4 Days	300 ng/mL
	Marijuana THC	15-30 Days	50 ng/mL
	Opiates OPI	2-4 Days	2000 ng/mL
	Amphetamines AMP	2-4 Days	1000 ng/mL
	Methamphetamines AMP	3-5 Days	1000 ng/mL
	Phencyclidine PCP	7-14 Days	25 ng/mL
	Benzodiazepines BZO	3-7 Days	300 ng/mL
	Barbiturates BAR	4-7 Days	300 ng/mL
	Methadone MTD	3-5 Days	300 ng/mL
	Tricyclic Antidepressants TCA		1,000 ng/mL
	Oxycodone	2-4 Days	100 ng/mL
	Propoxyphene	1-2 Days	300 ng/mL
	Buprenorphine BUP (Suboxone, Subutex)	2-3 Days	10 ng/mL
HAIR	Marijuana THC	Up to 90 days	
	Amphetamines AMP	Up to 90 days	
	Cocaine COC	Up to 90 days	
	Opiates OPI	Up to 90 days	
	Phencyclidine PCP	Up to 90 days	

All women who screen positive on either biological multi-drug test or any one of the screeners are contacted immediately (for urine and screener results) or within 72 hours (for hair results) to detail the results of her test, encourage the participant to talk with her physician about her substance use, and offer her referrals to community resources for treatment that mirror what is currently given to patients by

medical staff in each clinic. They are encouraged to speak with the on-site clinic social worker who can provide further support.

Birth Outcome Measures

Birth outcome data, including miscarriage, stillbirth, birth weight, gestational age, head circumference, and NICU admissions, as well as a list of drugs prescribed during pregnancy and their dosage are collected by research staff via the EMR and entered into SurveyMonkey.

Participant Follow-Up

After completion of this research visit, participants are contacted once more by telephone one week after completing the surveys to complete the three screeners again to assess test-retest reliability. The average time commitment for the call is about 10-15 minutes, and upon completion $25 is loaded onto the reloadable gift card provided the week prior.

Pilot Study

To examine the recruitment process and determine acceptability from the target population of substance-using pregnant women prior to the start of the study, we conducted a one-month pilot study. Each step of the recruitment process was reviewed to determine where improvements could be made.

We recruited 21 participants from each site for a total of 42 participants (Table 3). Mean age (sd) of participants was 30.1 years (5.64). By race, 11 participants (26.2%) were White, 25 (59.5%) were Black/African American, 4 (9.5%) were Asian, 1 (2.4%) was Hispanic and 1 (2.4%) was Other. About 24.4% tested positive for illicit drugs on urine testing, 22% tested positive on hair sample testing. Seven (7) participants (16.7%) were lost to follow up.

Table 3 Pilot Study Participant Characteristics

Characteristics, N = 42	**Clinical Site**		
	Clinic 1	***Clinic 2***	***Both Sites***
Number of participants	21	21	42
Participant age in years, mean (SD)	27.10 (5.09)	33.05 (4.54)	30.07 (5.64)
Ethnicity n (%)			
African American/Black	18 (85.7)	7 (33.3)	25 (59.5)
Asian	0 (0.0)	4 (19.0)	4 (9.5)
Caucasian/White	2 (9.5)	9 (42.9)	11 (26.2)
Hispanic, Latino or Chicano	0 (0.0)	1 (4.8)	1 (2.4)
Some other group	1 (4.8)	0 (0.0)	1 (2.4)
Trimester n (%)			
1st	2 (9.5)	2 (9.5)	4 (9.5)
2nd	6 (28.6)	6 (28.6)	12 (28.6)
3rd	13 (61.9)	13 (61.9)	26 (61.9)
Urine Results n (%)			
Negative for all substances	15 (71.4)	16 (80.0)	31 (75.6)
Positive for at least 1 substance	6 (28.6)	4 (20.0)	10 (24.4)
Hair Results n (%)			
Negative for all substances	12 (60.0)	18 (85.7)	30 (73.2)
Positive for at least one substance	7 (35.0)	2 (9.5)	9 (22.0)
Invalid	1 (5.0)	1 (4.8)	2 (4.9)
Study Disposition n (%)			

Study completes	19 (90.48)	16 (76.2)	35 (83.3)
Lost to follow-up	2 (9.52)	5 (23.8)	7 (16.7)

Results from the pilot study confirmed the feasibility of this study. Eligibility criteria did not appear too restrictive, given the eligibility rate of 78% (although slightly lower than anticipated) (Figure 2). Overall, there was good comprehension of surveys, a low refusal rate for hair sampling (1 refusal/95 approached, 1.1%), and high study enrollment (Figure 2). The recruitment process took an average of 40 minutes.

Power and Sample Size

The sample size of 500 participants was chosen based on power analyses for the primary study questions. Based on a one-sample binomial approach, with a sample size of 500 participants, we can be 95% confident that the false negative rate in the population is under 10% (assuming no more than 35 individuals test positive in the biologic drug tests without a positive survey screener result). Similarly, we can be 95% confident that the false negative rate in the population is under 5% (assuming no more than 15 individuals test positive in the urine drug test without a positive survey screen result in the study). According to McNemar's Test, if at least 15% of the study participants have disagreement between any pair of survey results, 500 is a sufficient sample size to determine significant disagreement.

After a preliminary sample size of 500 was chosen, a power analysis was conducted to determine the detectable differences in age, race, and trimester with a sample size of 500. The power of the test of proportions was calculated based on the difference in the proportion of false negatives in each age group, race, and trimester of pregnancy. Assuming recruitment of an equal number of women aged 18 to 25 years and women 26 and older, and that their respective positive screener results are 20% and 10%, then the power to detect that difference is 0.88. If the respective screener results are 15% and 20%, then the power is much lower (0.31). If we further assume that recruitment of 23% White women and 77% non-White women and that white women have a false negative rate of 5% and non-White women have a false

negative rate of 15%, then the power is high (0.87). Similarly, if we assume recruitment of an equal number of women in each of the three trimesters of pregnancy and that women in one trimester have a false negative rate of 20% while women in another trimester have a false negative rate of 35%, then the power is high (0.87).

Analysis

For each screener, reliability and validity (convergent/discriminant validity) will be assessed, including calculating correlation coefficients between each pair of screeners and between each screener and the appropriate biologic drug tests. Test-retest reliability analysis will be conducted by examining the results of repeated screener administrations within one week of original screener administrations for consistency via correlation analysis. The sensitivity and specificity of each instrument will be calculated, presented, and interpreted. Each survey instrument will be compared to the gold standard (hair and urine sample drug testing) by comparing the false negative rates to a predetermined limit of acceptability. If the upper one-sided 95% binomial confidence interval around the false negative rate in the sample is less than that limit, then the survey instrument is considered acceptable. The 4P's Plus and SURP-P survey screeners will be compared to both urine and hair testing in the assessment of their sensitivity and specificity in relation to short and long-term drug use, respectively. The NIDA Quick Screen/NIDA-Modified ASSIST screener will be compared to urine testing, while particular questions from the screener regarding long-term drug use will be compared to hair testing.

Furthermore, we will assess if there are differences in the validity of each screener by age, race, and trimester. The false negative rate for each screener will be presented by age, race, and trimester. A two-sided test of proportions will be conducted to test for significant differences in false negative rates between age, race, and trimester for each screener. Chi-square tests (or Fisher's exact tests if subgroup sizes are small) may be conducted to determine whether the distribution of responses on each survey instrument is similar for age, race, and trimester. To examine differences in screener validity by age, race, and trimester, logistic regression models will be fitted to the data. To separately analyze differences in

probability of false positive results and false negative results on each survey, data will be stratified by screener and screener result (positive or negative) for a total of six models. In each model, the dependent variable will be coded 1 for invalid screener result (false negative or false positive) and 0 for valid screener result (true negative or true positive). Independent variables for age, race, and trimester will be added to the models to test whether they have a significant effect on the probability of an invalid screener result. Two-way interaction terms will be included in the model if they are found to be significant effects. In order to stratify results by trimester, if trimester or any two-way interaction term including trimester is a significant effect in the models for any of the screeners, probabilities of false positive/ false negative result will be presented separately by each trimester.

Finally, the prevalence of prescription and illicit drug use will be calculated based on hair test results and self-report. Prevalence of multi-drug exposure will also be calculated. An ANOVA model will be fitted to the data with a fixed effect for drug use (negative, positive, positive for multi-drug exposure) to test for significant differences in birthweight, gestational age, and head circumference based on participant hair drug tests result. Significant differences will be noted and discussed. The relative risk of NICU admission, stillbirth, and miscarriage will be examined. A risk ratio will be calculated and will quantify the percentage difference in these three variables between those with positive hair drug tests versus negative drug test. The risk ratio takes on values between zero and infinity. A risk ratio of one means that there is no difference in NICU admissions, stillbirth, or miscarriage between the participants' biologic drug tests results. A risk ratio very small (close to zero) or very large means a large difference between NICU admissions, stillbirth, or miscarriage based on the hair drug tests results. Approximate 95% confidence intervals for the relative risk will be calculated. The same relative risk ratios and 95% confidence intervals will be calculated for a positive biologic drug tests for multi-drug exposure versus positive for a single-drug exposure. Further, relative risk ratios will be computed with 95% confidence intervals stratified by trimester.

Discussion

Our ongoing research has five aspects of significance. First, the importance of screening pregnant women and the public health impact of the current research is tied directly to the negative health consequences associated with illicit and prescription drug use during pregnancy. Second, it utilizes both urine and hair testing to enable us to examine past 90-day substance use history with precision. Hair analysis provides nearly twice the number of positives due to its longer detection window, but often cannot capture very recent use. Urine analysis supplements hair analysis to allow for the most comprehensive validation of screeners possible. Third, the study compares three screeners acknowledged by the World Health Organization (WHO) to screen for multiple substances to each other and to the biological screeners (gold standard). This is the first study to conduct a direct, head-to-head comparison of multiple screening tools for prescription and illicit drug use among pregnant women, while also utilizing biologic measures as a gold standard against which to compare. Fourth, the study utilizes electronic medical records (EMR) to capture prescribed drugs and birth outcome data of enrolled participants. The ability to access a participant's prescription drug orders enables better tracking and distinction between prescription drug use and abuse, while birth outcome data allows for determination of associations between specific drug use and birth outcomes. Fifth, the study has the potential to shift clinical practice towards universal standardized substance use screening.

Despite the significant contributions of this work, it is not without limitations. Though the study will enroll a large sample of pregnant women, it is a convenience sample from two prenatal clinics in an urban area. We have attempted to increase generalizability by enrolling women from two clinics with different population characteristics: one clinic serves low-income, Medicaid-eligible, primarily African-American women and the other serves privately-insured, primarily White women. Second, there is a possibility of selection bias. Incentive may be more appealing to those who have lower socioeconomic status, individuals with more time may be those willing to take the study, and pregnant women who use substances may not want to participate. For the latter point, we have obtained a Certificate of

Confidentiality and ensured participants that their data will not be shared with anyone including clinic staff. Finally, our study is limited to adults. Though our initial protocol included adolescents, the IRB did not allow for "no-benefit" studies enrolling pregnant adolescents. This is an important area for further exploration, given that pregnant adolescents report higher substance use rates than pregnant adults in national surveys.

The primary innovation of this project is that it may provide a final evidence-based recommendation for the tool(s) best suited for screening for illicit and prescription drugs among a diverse sample of pregnant women. The provision of this evidence-based guidance to clinicians is a concrete application of findings that is rare in public health research. We recognize that screening is a first step; also important is the need for a public health focus on treatment of substance use during pregnancy to enhance the odds of a successful pregnancy outcome. Barriers to treatment that are imperative to address are the potential legal repercussions of identifying substance use during pregnancy that exist in some states[9] and unintentional breach of confidentiality.[10] There is a strong need for a re-examination of state policies so that women are not punished for having a treatment need.

Substance use during pregnancy, and specifically prescription and illicit drug use, are high priority topics for the Centers for Disease Control and Prevention (CDC), WHO, The American Congress of Obstetricians and Gynecologists (ACOG), Substance Abuse and Mental Health Services Administration (SAMHSA), National Institute on Drug Abuse (NIDA), and the National Institutes of Health (NIH). Universal screening has the potential to greatly enhance maternal and infant health outcomes and reduce healthcare costs. Specifically, the current research supports the following Healthy People 2020 public health goals and objectives which include reducing maternal illness and complications due to pregnancy; increasing the proportion of pregnant women who receive adequate prenatal care' increasing abstinence from alcohol, cigarettes, and illicit drugs among pregnant women; and increasing the proportion of women delivering a live birth who received preconception care services and practiced key recommended preconception health behaviors.

This research addresses an important problem by identifying a valid substance use screening instrument for illicit and prescription drugs among pregnant women that is accurate, brief, and acceptable to both patients and health care providers in a primary care setting. Identifying and validating one instrument that functions the closest to the "gold standard" of biologic testing (i.e., urine and hair) and disseminating this information widely will increase the likelihood that primary care clinics nationwide may adopt a quick and easy screener universally. We may find that one instrument does not stand out but that each has its distinct advantages and disadvantages; in this case, the performance of each measure will be detailed with recommendations for which screener may work the best with a given population.

Declarations

Ethics approval and consent to participate

This study is approved by the Institutional Review Boards (IRB) of the University of Maryland (HP-00072042), Baltimore; and Battelle Memorial Institute (0619-100106433). All participants are required to give their informed consent prior to any study procedure.

Availability of data and materials

The data that support the findings of this study are available on request from the corresponding author [VHCC]. The data are not publicly available due to them containing information that could compromise research participant privacy/consent.

Competing interests

The authors declare that there are no competing interests.

Authors' contributions

VCC conceived the study. VCC, EO, EP, KT, BK and KM participated in the drafting of the manuscript and each approved the final draft

Funding

The research reported in this article is supported by the National Institute on Drug Abuse (NIDA) of the National Institutes of Health (NIH) grant under Award Number R01DA041328 (PI-Coleman-Cowger). The content is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health.

REFERENCES

1. Substance Abuse and Mental Health Services Administration (SAMHSA). Results from the 2012 National Survey on Drug Use and Health: Summary of national findings. In: *NSDUH Series H-46, HHS Publication No.(SMA) 13-4795.* Substance Abuse and Mental Health Services Administration Rockville, MD; 2013.
2. World Health Organization. Guidelines for the identification and management of substance use and substance use disorders in pregnancy. 2014.
3. Scholle SH, Kelleher K. Assessing primary care performance in an obstetrics/gynecology clinic. *Women & health.* 2003;37(1):15-30.
4. ACOG Committee on Health Care for Underserved Women. ACOG Committee Opinion No. 524: opioid abuse, dependence, and addiction in pregnancy. *Obstetric Anesthesia Digest.* 2013;33(2):79-80.
5. Group W. The alcohol, smoking and substance involvement screening test (ASSIST): development, reliability and feasibility. *Addiction.* 2002;97(9):1183-1194.
6. Smith PC, Schmidt SM, Allensworth-Davies D, Saitz R. A single-question screening test for drug use in primary care. *Archives of internal medicine.* 2010;170(13):1155-1160.
7. Chasnoff I, Wells A, McGourty R, Bailey L. Validation of the 4P's Plus© screen for substance use in pregnancy validation of the 4P's Plus. *Journal of Perinatology.* 2007;27(12):744.
8. Yonkers KA, Gotman N, Kershaw T, Forray A, Howell HB, Rounsaville BJ. Screening for prenatal substance use: development of the Substance Use Risk Profile-Pregnancy scale. *Obstetrics and gynecology.* 2010;116(4):827.
9. Guttmacher Institute. State Laws and Policies:Substance Use during Pregnancy. 2017; https://www.guttmacher.org/state-policy/explore/substance-use-during-pregnancy. Accessed December 18, 2017.
10. Knight KR. *Addicted. pregnant. poor.* Duke University Press; 2015.

Figure 1: Study procedures

Figure 2: Pilot study participation

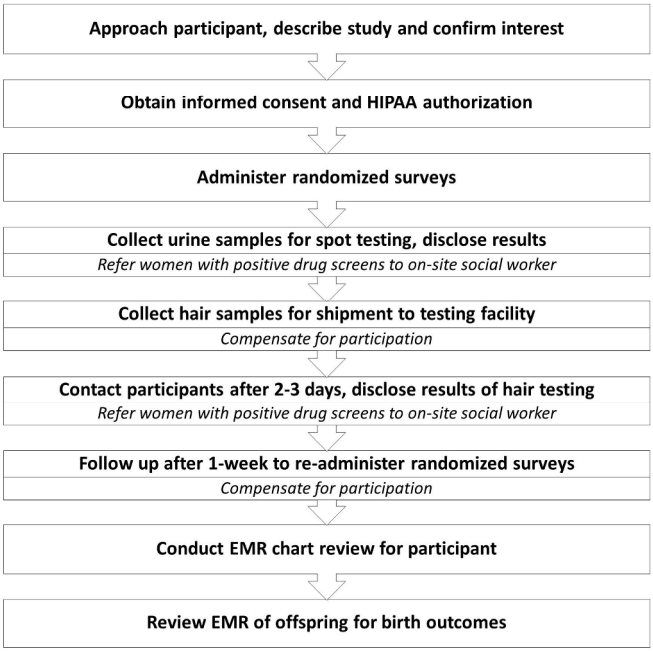

FIGURE 1: Study procedures

227x162mm (300 x 300 DPI)

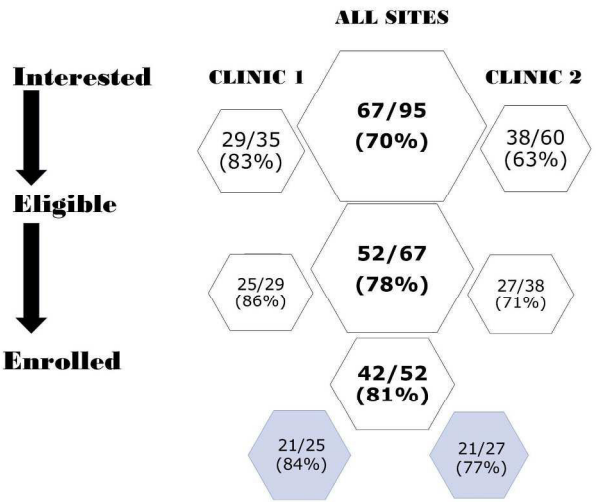

Figure 2: Pilot study participation

227x162mm (300 x 300 DPI)