

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email editorial.bmjopen@bmj.com

BMJ Open

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

Journal:	<i>BMJ Open</i>
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

SCHOLARONE™
Manuscripts

Peer Review Only

1 Study Protocol: Validating Screening Tools for Illicit and Prescription Drug Use in

2 Pregnancy Using Hair and Urine Sample Testing

3 Coleman-Cowger, Victoria¹; Oga, Emmanuel A. ¹; Peters, Erica N. ¹; Trocin, Kathleen; Koszowski,
4 Bartosz¹; Mark, Katrina²

5 1. Battelle Memorial Institute, Baltimore, MD

6 2. University of Maryland Medical Center, Baltimore, MD

7 Authors contact information:

8 Victoria Coleman-Cowger, PhD: colemancowger@battelle.org

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

10 Erica Peters, PhD: finan@battelle.org

11 Kathleen Trocin, MPH: trocin@battelle.org

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

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

14 List of Abbreviations

15 NIDA - National Institute on Drug Abuse

16 ASSIST – Alcohol, Smoking, and Substance Involvement Screening Test

17 SURP-P - Substance Use Risk Profile-Pregnancy

18 EMR – Electronic Medical Records

19 NICU - Neonatal Intensive Care Unit

20 WHO – World Health Organization

21 HIPAA - Health Insurance Portability and Accountability Act

22 Abstract for Protocol

23 **Introduction:** Prescription drug use in the United States (U.S.) has increased by more than 60% in the
24 last 3 decades. Prevalence of prescription drug use among pregnant women is currently estimated around
25 50%. Prevalence of illicit drug use in the U.S. is 14.6% among pregnant adolescents, 8.6% among
26 pregnant young adults, and 3.2% among pregnant adults. The first step in identifying problematic drug
27 use during pregnancy is screening; however, no specific substance use screener has been universally
28 recommended for use with pregnant women to identify illicit or prescription drug use. This study
29 compares and validates three existing substance use screeners for pregnancy - 4 P's Plus, NIDA Quick
30 Screen/ASSIST, and the Substance Use Risk Profile-Pregnancy (SURP-P) scale.

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

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

44 **Keywords:** substance use, pregnancy, screening, biochemical verification, NIDA Quick Screen/ASSIST,
45 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.

66 Introduction

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

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

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

1
2
3 90 with pregnant women[7]. The 4P's Plus is brief but is associated with a licensing fee which may be a
4
5 91 hindrance to widespread use. The SURP-P is a validated scale composed of three questions that can
6
7 92 differentiate between populations of pregnant women at low-risk or high-risk for substance use[8]. The
8
9 93 SURP-P is a simple and flexible tool for identifying possible substance use in pregnancy; however, a
10
11 94 further screen is required for identifying those who would require treatment.
12
13

14 95 To bridge this gap and identify the most universally valid and reliable screening tool for drug abuse in
15
16 96 pregnancy, this study aims to compare and validate three existing substance use screeners - 4 P's Plus,
17
18 97 NIDA Quick Screen/ASSIST, and the Substance Use Risk Profile-Pregnancy (SURP-P) scale - among a
19
20 98 cross section of 500 pregnant women presenting to two obstetrics clinics in Baltimore, Maryland (US).
21
22 99 The overarching goal of this effort is to determine which screening tool is most effective in identifying
23
24 100 prescription drug abuse and illicit drug use among pregnant women and acceptable among patients and
25
26 101 clinicians so that evidence-based guidance may be offered.
27
28
29

30 102 **Methods/Design**

31 103 **Specific Aims**

32
33 104 Specific Aims of this study are to: a) conduct validity analyses to determine sensitivity, specificity,
34
35 105 usability (test-retest reliability), and how each scale compares to the others and to the gold standard of
36
37 106 urine and hair drug testing in identifying prescription and illicit drug use; b) determine the impact of
38
39 107 clinic population variables (age, race, trimester of pregnancy) on validity of the three substance use
40
41 108 screeners; and c) assess birth outcomes (birth weight, gestational age, head circumference, and Neonatal
42
43 109 Intensive Care Unit (NICU) admissions) associated with the most widely used prescription drug and
44
45 110 multi-drug exposure.
46
47
48
49

50 111 **Study Design**

51
52 112 This study is a cross-sectional study that evaluates the sensitivity, specificity and usability of existing
53
54 113 substance use screeners. We chose this study design following an extensive search of the literature, an
55
56
57
58
59
60

1
2
3 114 overall assessment of feasibility and consultation with stakeholders (e.g., clinicians, pregnant women and
4
5 115 substance use researchers). We believe that a cross-sectional study such as ours is appropriate for the
6
7 116 evaluation of the accuracy and reliability of these screeners. We were also aided by the knowledge that
8
9 117 the prevalence of substance use in pregnancy is high[1]. This implies that we are likely to obtain good
10
11 118 sensitivity and specificity estimates, with narrow confidence intervals, in a cross-sectional design which is
12
13 119 favorable in terms of cost and feasibility.

16 120 Setting

17
18 121 The study is being implemented at two urban obstetric clinics which serve diverse populations of
19
20 122 pregnant women. The study plans to recruit 500 participants to complete a demographics questionnaire,
21
22 123 followed by a randomized order of the NIDA Quick Screen/NIDA-modified ASSIST, 4 P's Plus, and
23
24 124 SURP-P. Participants are recruited during their regularly scheduled prenatal appointment, then contacted
25
26 125 again one week later by telephone to re-administer the screeners. Participants consent to multi-drug urine
27
28 126 testing, hair drug testing, and access to prescription drug and birth outcome data from electronic medical
29
30 127 records (EMR).

31
32
33
34 128 Recruitment Sites. We are recruiting participants from two obstetric outpatient clinics. Currently all
35
36 129 obstetric patients are screened for use of drugs, alcohol and tobacco at their first prenatal visit by medical
37
38 130 staff. Additionally, all new obstetric patients receive an in-depth evaluation by a social worker which
39
40 131 includes a more detailed assessment of both substance use and mental health disorder history.

41
42
43 132 In the first clinic, which is the larger of the two clinics, most patients (97%) are publicly insured with
44
45 133 medical assistance and are over the age of 20 (80%). This clinic's population is primarily African-
46
47 134 American and low-income, all of whom undergo urine toxicological screening for substance use
48
49 135 identification. Based on preliminary data obtained from the clinic, about 950 individual obstetric patients
50
51 136 are cared for at this clinic annually. In the second (smaller) clinic, approximately 500 pregnant women are
52
53 137 cared for annually. Most patients (87%) have commercial insurance and 13% have either medical
54
55 138 assistance or Medicare. Most are over the age of 20 years (90%). Due to varying insurance coverage for

1
2
3 139 urine toxicology screens, patients in this office do not universally undergo urine toxicology screening but
4
5 140 all are screened for drug use using various interview techniques by their obstetric care providers at their
6
7 141 first prenatal visit. Based on historical data, we expect about 500 individual obstetric patients to be cared
8
9 142 for in this clinic across all trimesters of pregnancy in the one year of study recruitment.

11
12 143 Across both study sites, our source population covers a diverse set of participants and captures pregnant
13
14 144 women across all socioeconomic categories, insurance types, ethnicities and drug use patterns. This
15
16 145 ensures that our study results are generalizable to most populations of pregnant women.

18 19 146 **Study Population**

20
21 147 In the first clinic, of the estimated 950 individual obstetric patients cared for at this clinic annually, we
22
23 148 anticipated approaching 403 (50%), and expected 322 (80%) or more to agree to participate in this study.

24
25
26 149 In the second clinic, of the approximately 500 pregnant women cared for annually, we expect at least 450
27
28 150 (90%) to meet eligibility criteria. We anticipate approaching 225 pregnant women (50%) and expect 180
29
30 151 (80%) or more to agree to participate in this study.

31
32
33 152 Expected participation percentages are based on a similar grant-funded study that recruited pregnant
34
35 153 smokers from the same population and required consent for urine testing (cotinine) and birth data
36
37 154 abstraction from EMR.

38
39
40 155 Participant eligibility criteria include the following: a) currently pregnant (pre-determined by clinic staff);
41
42
43 156 b) age 18 or older; c) able to speak and understand English sufficiently to provide informed consent; and
44
45
46 157 d) natural hair length at least 3 cm to allow for substance use testing.

47
48 158 If eligibility criteria are met, research staff then obtain informed consent and medical releases for urine
49
50 159 collection, hair drug testing, and prescription drug and birth outcome data abstraction from the EMR.

160 Ethical Approval

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

164 Study Procedures

165 *Approach*

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

177 *Recruitment*

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

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

187

188

189 *Self-Report Measures*

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

197 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

198

199 *Biochemical Measures*

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

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

	Drug class	Detection	Confirmation
U	Cocaine COC	2-4 Days	300 ng/mL
	Marijuana THC	15-30 Days	50 ng/mL
	Opiates OPI	2-4 Days	2000 ng/mL
R	Amphetamines AMP	2-4 Days	1000 ng/mL
	Methamphetamines AMP	3-5 Days	1000 ng/mL
I	Phencyclidine PCP	7-14 Days	25 ng/mL
	Benzodiazepines BZO	3-7 Days	300 ng/mL
N	Barbiturates BAR	4-7 Days	300 ng/mL
	Methadone MTD	3-5 Days	300 ng/mL
E	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
H	Marijuana THC	Up to 90 days	
	Amphetamines AMP	Up to 90 days	
A	Cocaine COC	Up to 90 days	
	Opiates OPI	Up to 90 days	
I	Phencyclidine PCP	Up to 90 days	
R			

207
 208 All women who screen positive on either biological multi-drug test or any one of the screeners
 209 are contacted immediately (for urine and screener results) or within 72 hours (for hair results) to detail the
 210 results of her test, encourage the participant to talk with her physician about her substance use, and offer
 211 her referrals to community resources for treatment that mirror what is currently given to patients by

1
2
3 212 medical staff in each clinic. They are encouraged to speak with the on-site clinic social worker who can
4
5 213 provide further support.
6
7

8 214 *Birth Outcome Measures*

9

10 215 Birth outcome data, including miscarriage, stillbirth, birth weight, gestational age, head circumference,
11
12 216 and NICU admissions, as well as a list of drugs prescribed during pregnancy and their dosage are
13
14 217 collected by research staff via the EMR and entered into SurveyMonkey.
15
16

17 218 *Participant Follow-Up*

18

19 219 After completion of this research visit, participants are contacted once more by telephone one week after
20
21 220 completing the surveys to complete the three screeners again to assess test-retest reliability. The average
22
23 221 time commitment for the call is about 10-15 minutes, and upon completion \$25 is loaded onto the
24
25 222 reloadable gift card provided the week prior.
26
27

28 223 *Pilot Study*

29

30 224 To examine the recruitment process and determine acceptability from the target population of substance-
31
32 225 using pregnant women prior to the start of the study, we conducted a one-month pilot study. Each step of
33
34 226 the recruitment process was reviewed to determine where improvements could be made.
35
36

37 227 We recruited 21 participants from each site for a total of 42 participants (Table 3). Mean age (sd)
38
39 228 of participants was 30.1 years (5.64). By race, 11 participants (26.2%) were White, 25 (59.5%) were
40
41 229 Black/African American, 4 (9.5%) were Asian, 1 (2.4%) was Hispanic and 1 (2.4%) was Other. About
42
43 230 24.4% tested positive for illicit drugs on urine testing, 22% tested positive on hair sample testing. Seven
44
45 231 (7) participants (16.7%) were lost to follow up.
46
47

48 232

49
50
51 233

52
53
54 234

235

236 *Table 3 Pilot Study Participant Characteristics*

Characteristics, N = 42	Clinical Site		
	<i>Clinic 1</i>	<i>Clinic 2</i>	<i>Both Sites</i>
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 (%)			
1 st	2 (9.5)	2 (9.5)	4 (9.5)
2 nd	6 (28.6)	6 (28.6)	12 (28.6)
3 rd	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)
-------------------	----------	----------	----------

237

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

243 *Power and Sample Size*

244 The sample size of 500 participants was chosen based on power analyses for the primary study questions.
 245 Based on a one-sample binomial approach, with a sample size of 500 participants, we can be 95%
 246 confident that the false negative rate in the population is under 10% (assuming no more than 35
 247 individuals test positive in the biologic drug tests without a positive survey screener result). Similarly, we
 248 can be 95% confident that the false negative rate in the population is under 5% (assuming no more than
 249 15 individuals test positive in the urine drug test without a positive survey screen result in the study).
 250 According to McNemar's Test, if at least 15% of the study participants have disagreement between any
 251 pair of survey results, 500 is a sufficient sample size to determine significant disagreement.

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

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

264 Analysis

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

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

1
2
3 286 screener and screener result (positive or negative) for a total of six models. In each model, the dependent
4
5 287 variable will be coded 1 for invalid screener result (false negative or false positive) and 0 for valid
6
7 288 screener result (true negative or true positive). Independent variables for age, race, and trimester will be
8
9 289 added to the models to test whether they have a significant effect on the probability of an invalid screener
10
11 290 result. Two-way interaction terms will be included in the model if they are found to be significant effects.
12
13 291 In order to stratify results by trimester, if trimester or any two-way interaction term including trimester is
14
15 292 a significant effect in the models for any of the screeners, probabilities of false positive/ false negative
16
17 293 result will be presented separately by each trimester.
18
19
20 294 Finally, the prevalence of prescription and illicit drug use will be calculated based on hair test results and
21
22 295 self-report. Prevalence of multi-drug exposure will also be calculated. An ANOVA model will be fitted to
23
24 296 the data with a fixed effect for drug use (negative, positive, positive for multi-drug exposure) to test for
25
26 297 significant differences in birthweight, gestational age, and head circumference based on participant hair
27
28 298 drug tests result. Significant differences will be noted and discussed. The relative risk of NICU
29
30 299 admission, stillbirth, and miscarriage will be examined. A risk ratio will be calculated and will quantify
31
32 300 the percentage difference in these three variables between those with positive hair drug tests versus
33
34 301 negative drug test. The risk ratio takes on values between zero and infinity. A risk ratio of one means that
35
36 302 there is no difference in NICU admissions, stillbirth, or miscarriage between the participants' biologic
37
38 303 drug tests results. A risk ratio very small (close to zero) or very large means a large difference between
39
40 304 NICU admissions, stillbirth, or miscarriage based on the hair drug tests results. Approximate 95%
41
42 305 confidence intervals for the relative risk will be calculated. The same relative risk ratios and 95%
43
44 306 confidence intervals will be calculated for a positive biologic drug tests for multi-drug exposure versus
45
46 307 positive for a single-drug exposure. Further, relative risk ratios will be computed with 95% confidence
47
48 308 intervals stratified by trimester.
49
50
51
52
53
54
55
56
57
58
59
60

309 Discussion

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

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

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

1
2
3 333 Administration (SAMHSA), National Institute on Drug Abuse (NIDA), and the National Institutes of
4
5 334 Health (NIH). Universal screening has the potential to greatly enhance maternal and infant health
6
7 335 outcomes and reduce healthcare costs. Specifically, the current research supports the following Healthy
8
9 336 People 2020 public health goals and objectives which include reducing maternal illness and complications
10
11 337 due to pregnancy; increasing the proportion of pregnant women who receive adequate prenatal care;
12
13 338 increasing abstinence from alcohol, cigarettes, and illicit drugs among pregnant women; and increasing
14
15 339 the proportion of women delivering a live birth who received preconception care services and practiced
16
17 340 key recommended preconception health behaviors.

18
19
20
21 341 This research addresses an important problem by identifying a valid substance use screening
22
23 342 instrument for illicit and prescription drugs among pregnant women that is accurate, brief, and acceptable
24
25 343 to both patients and health care providers in a primary care setting. Identifying and validating one
26
27 344 instrument that functions the closest to the “gold standard” of biologic testing (i.e., urine and hair) and
28
29 345 disseminating this information widely will increase the likelihood that primary care clinics nationwide
30
31 346 may adopt a quick and easy screener universally. We may find that one instrument does not stand out but
32
33 347 that each has its distinct advantages and disadvantages; in this case, the performance of each measure will
34
35 348 be detailed with recommendations for which screener may work the best with a given population.

36
37
38 349
39
40 350
41
42 351
43
44 352
45
46 353
47
48 354
49
50 355
51
52 356
53
54 357
55
56
57
58
59
60

1
2
3
4 3585
6 3597
8 3609
10 36111
12 36213
14 363 **Declarations**15 364 *Ethics approval and consent to participate*

16
17
18 365 This study is approved by the Institutional Review Boards (IRB) of the University of Maryland (HP-
19 366 00072042), Baltimore; and Battelle Memorial Institute (0619-100106433). All participants are required to
20 367 give their informed consent prior to any study procedure.

21
22 368 *Availability of data and materials*

23
24 369 The data that support the findings of this study are available on request from the corresponding author
25 370 [VHCC]. The data are not publicly available due to them containing information that could compromise
26 371 research participant privacy/consent.

27
28 372 *Competing interests*

29
30 373 The authors declare that there are no competing interests.

31
32
33 374 *Authors' contributions*

34
35 375 VCC conceived the study. VCC, EO, EP, KT, BK and KM participated in the drafting of the manuscript and
36 376 each approved the final draft

37
38 377 *Funding*

39
40 378 The research reported in this article is supported by the National Institute on Drug Abuse (NIDA) of the

41
42 379 National Institutes of Health (NIH) grant under Award Number R01DA041328 (PI-Coleman-Cowger).

43
44 380 The content is solely the responsibility of the authors and does not represent the official views of the

45
46 381 National Institutes of Health.

47
48
49 38250
51
52 38353
54
55 38456
57
58
59
60

385

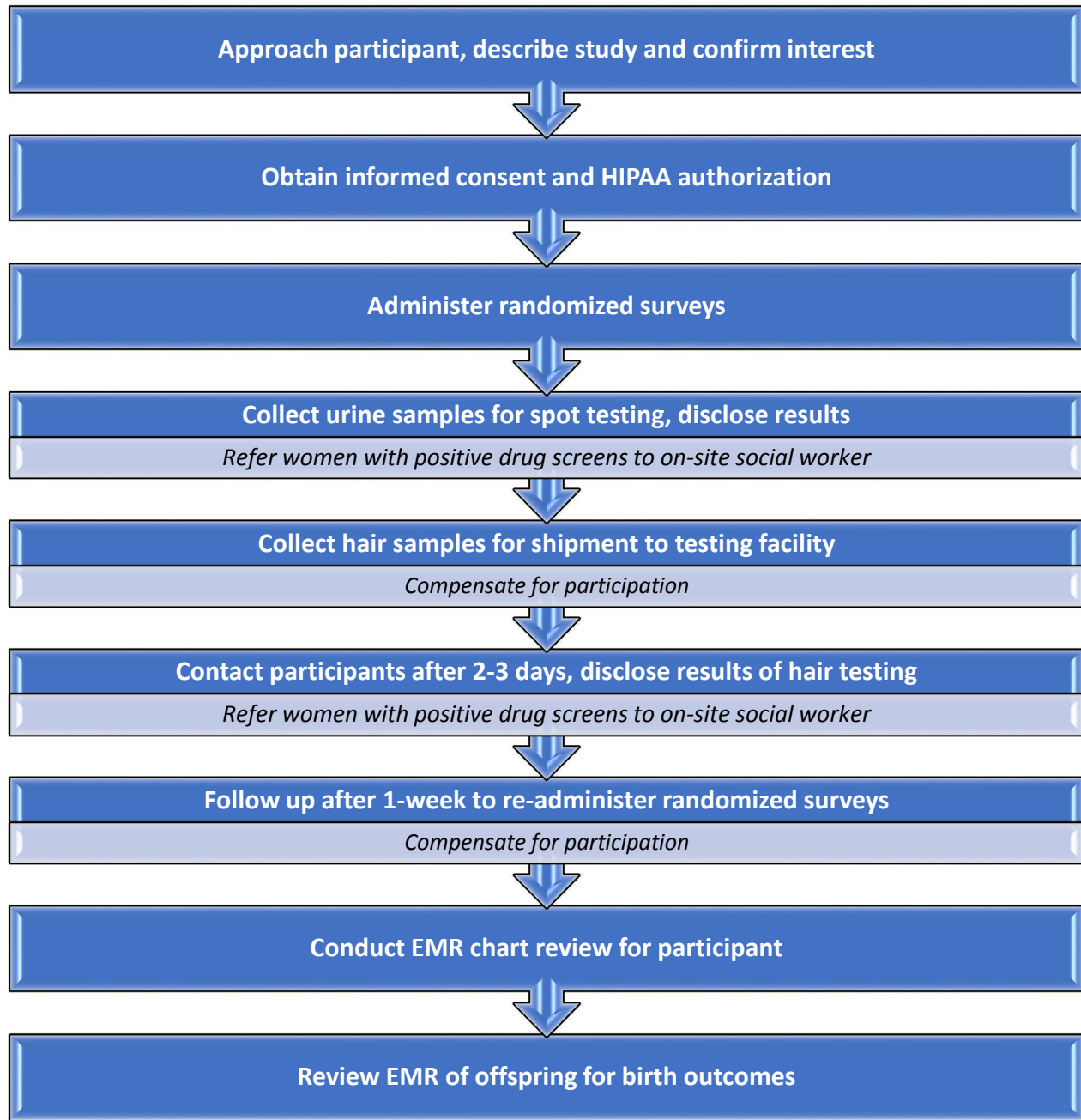
386

387 REFERENCES

- 388 1. **Substance Abuse and Mental Health Services Administration (SAMHSA). Results from the**
389 **2012 National Survey on Drug Use and Health: Summary of national findings.** In: *NSDUH*
390 *Series H-46, HHS Publication No(SMA) 13-4795.* edn.: Substance Abuse and Mental Health
391 Services Administration Rockville, MD; 2013.
- 392 2. **World Health Organization. Guidelines for the identification and management of**
393 **substance use and substance use disorders in pregnancy.** 2014.
- 394 3. Scholle SH, Kelleher K: **Assessing primary care performance in an obstetrics/gynecology**
395 **clinic.** *Women & health* 2003, **37**(1):15-30.
- 396 4. **ACOG Committee on Health Care for Underserved Women. ACOG Committee Opinion**
397 **No. 524: opioid abuse, dependence, and addiction in pregnancy.** *Obstetric Anesthesia Digest*
398 2013, **33**(2):79-80.
- 399 5. Group W: **The alcohol, smoking and substance involvement screening test (ASSIST):**
400 **development, reliability and feasibility.** *Addiction* 2002, **97**(9):1183-1194.
- 401 6. Smith PC, Schmidt SM, Allensworth-Davies D, Saitz R: **A single-question screening test for**
402 **drug use in primary care.** *Archives of internal medicine* 2010, **170**(13):1155-1160.
- 403 7. Chasnoff I, Wells A, McGourty R, Bailey L: **Validation of the 4P's Plus© screen for substance**
404 **use in pregnancy validation of the 4P's Plus.** *Journal of Perinatology* 2007, **27**(12):744.
- 405 8. Yonkers KA, Gotman N, Kershaw T, Forray A, Howell HB, Rounsaville BJ: **Screening for**
406 **prenatal substance use: development of the Substance Use Risk Profile-Pregnancy scale.**
407 *Obstetrics and gynecology* 2010, **116**(4):827.

408

1 **FIGURE 1: Study procedures**

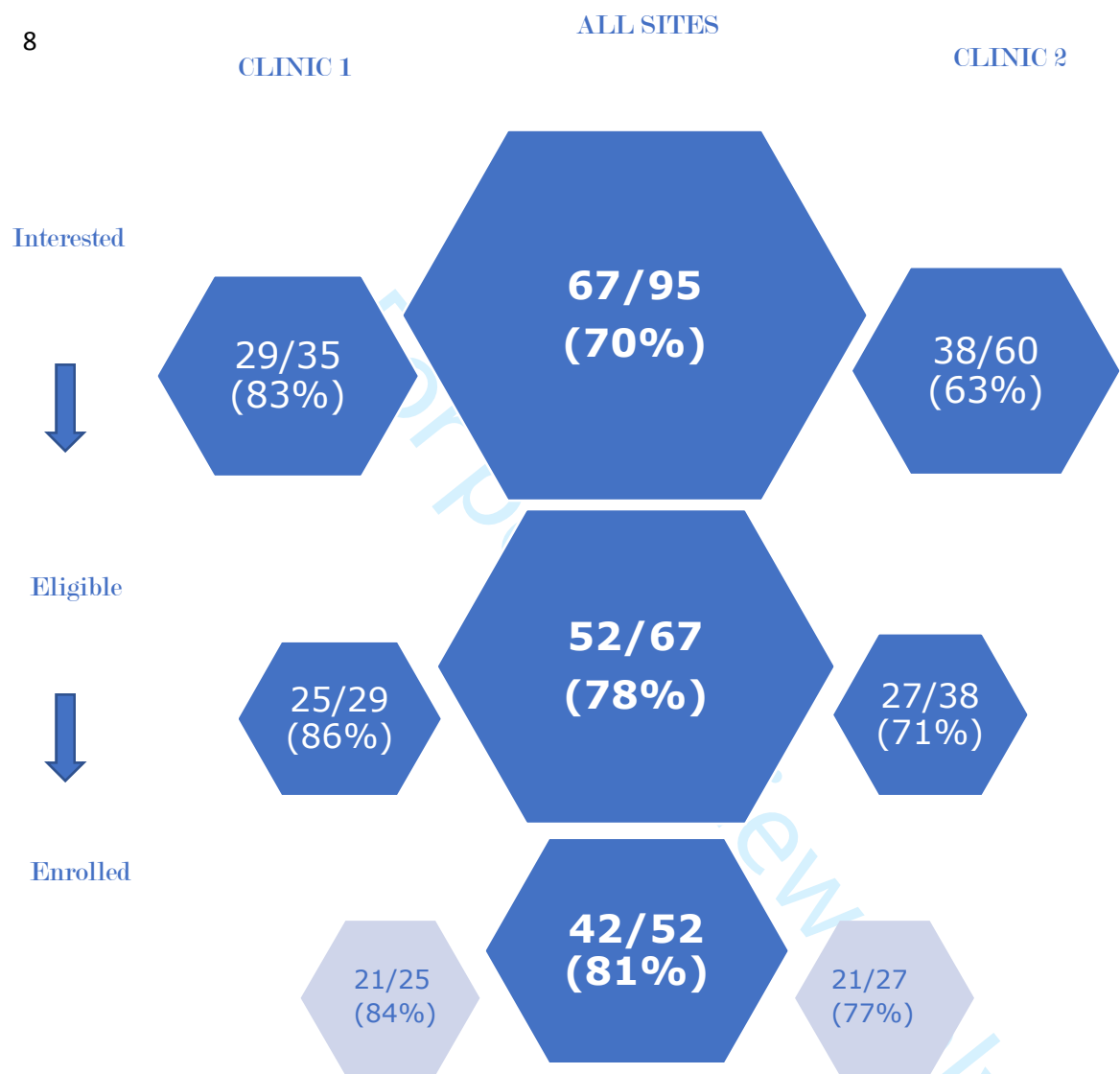


2
3 **Figure 1: Study procedures**

4
5

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

6 **FIGURE 2: Pilot study participation**



9
10 **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:	<i>BMJ Open</i>
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

SCHOLARONE™
Manuscripts

1
2
3
4 1 Comparison and Validation of Screening Tools for Substance Use in Pregnancy: A
5
6
7 2 Cross-Sectional Study conducted in Maryland Prenatal Clinics
8
9

10 3 Coleman-Cowger, Victoria¹; Oga, Emmanuel A. ¹; Peters, Erica N. ¹; Trocin, Kathleen; Koszowski,
11
12 4 Bartosz¹; Mark, Katrina²
13
14

- 15 5 1. Battelle Memorial Institute, Baltimore, MD
16
17 6 2. University of Maryland Medical Center, Baltimore, MD
18
19

20 7 Authors contact information:

21 8 Victoria Coleman-Cowger, PhD: colemancowger@battelle.org

22 9 Emmanuel Oga, MD, MPH: oga@battelle.org

23
24 10 Erica Peters, PhD: finan@battelle.org

25
26 11 Kathleen Trocin, MPH: trocina@battelle.org

27
28 12 Bartosz Koszowski, PharmD, PhD: koszowskib@battelle.org

29
30 13 Katrina Mark, MD: kmark@FPI.umaryland.edu
31
32

33
34
35
36 14 List of Abbreviations

37 15 NIDA - National Institute on Drug Abuse

38 16 ASSIST – Alcohol, Smoking, and Substance Involvement Screening Test

39 17 SURP-P - Substance Use Risk Profile-Pregnancy

40 18 EMR – Electronic Medical Records

41 19 NICU - Neonatal Intensive Care Unit

42 20 WHO – World Health Organization

43 21 HIPAA - Health Insurance Portability and Accountability Act
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

22 Abstract for Protocol

23 **Introduction:** Prescription drug use in the United States (U.S.) has increased by more than 60% in the
24 last 3 decades. Prevalence of prescription drug use among pregnant women is currently estimated around
25 50%. Prevalence of illicit drug use in the U.S. is 14.6% among pregnant adolescents, 8.6% among
26 pregnant young adults, and 3.2% among pregnant adults. The first step in identifying problematic drug
27 use during pregnancy is screening; however, no specific substance use screener has been universally
28 recommended for use with pregnant women to identify illicit or prescription drug use. This study
29 compares and validates three existing substance use screeners for pregnancy - 4 P's Plus, NIDA Quick
30 Screen/ASSIST, and the Substance Use Risk Profile-Pregnancy (SURP-P) scale.

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

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

44 **Keywords:** substance use, pregnancy, screening, biochemical verification, NIDA Quick Screen/ASSIST,
45 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.

67 Introduction

68 Abuse of prescription and illicit drugs in pregnancy is a growing cause of maternal and neonatal
69 morbidity and mortality in the United States (U.S.). According to data from the 2012 and 2013 U.S.
70 National Survey on Drug Use and Health (NSDUH), the rate of current illicit drug use (including non-
71 medical use of prescription drugs) in pregnant adolescents and women was 14.6% among adolescents
72 ages 15 to 17, 8.6% among young adults (18 to 25), and 3.2% among adults (26 to 44)¹. The
73 consequences of this problem include spontaneous abortions, stillbirths, low birth weight, prematurity,
74 neonatal abstinence syndrome and congenital malformations².

75 Given the relatively high frequency of provider-patient contact during the prenatal period, obstetric care
76 providers have the unique opportunity to identify substance abuse in pregnancy. Furthermore, for
77 pregnant women from socioeconomically disadvantaged groups, obstetricians often serve as primary care
78 physicians and typically are the only contact these women have with the healthcare system³. Prenatal
79 screening for drug use is an important way to identify drug abuse in pregnancy, as strongly recommended
80 by the American Congress of Obstetricians and Gynecologists (ACOG)⁴. But, while validated alcohol and
81 tobacco screeners have been recommended by the United States Preventive Services Task Force
82 (USPSTF), there is currently no universally recommended validated screening tool for identifying illicit
83 drug use in pregnancy.

84 Currently three separate, validated tools exist that screen for use of more than one substance among
85 pregnant women: The Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST); the 4 P's
86 Plus; and the Substance Use Risk Profile – Pregnancy (SURP-P).⁵⁻⁸ The ASSIST has been validated
87 across several populations, but it has not yet been formally validated with pregnant women⁵. A modified
88 ASSIST, with items on tobacco and alcohol use removed, was incorporated by NIDA to their Quick
89 Screen as a follow-up to the 4-question pre-screener, this is referred to as the NIDA Quick
90 Screen/ASSIST. The 4 P's Plus was designed to identify drug use in pregnancy and has been validated

1
2
3 91 with pregnant women⁷. The 4P's Plus is brief but is associated with a licensing fee which may be a
4
5 92 hindrance to widespread use. The SURP-P is a validated scale composed of three questions that can
6
7 93 differentiate between populations of pregnant women at low-risk or high-risk for substance use⁸. The
8
9 94 SURP-P is a simple and flexible tool for identifying possible substance use in pregnancy; however, a
10
11 95 further screen is required for identifying those who would require treatment.
12
13

14 96 To bridge this gap and identify the most universally valid and reliable screening tool for drug abuse in
15
16 97 pregnancy, this study aims to compare and validate three existing substance use screeners - 4 P's Plus,
17
18 98 NIDA Quick Screen/ASSIST, and the Substance Use Risk Profile-Pregnancy (SURP-P) scale - among a
19
20 99 cross section of 500 pregnant women presenting to two obstetrics clinics in Baltimore, Maryland (US).
21
22 100 The overarching goal of this effort is to determine which screening tool is most effective in identifying
23
24 101 prescription drug abuse and illicit drug use among pregnant women and acceptable among patients and
25
26 102 clinicians so that evidence-based guidance may be offered.
27
28
29

30 103 **Methods/Design**

31 104 **Specific Aims**

32
33 105 Specific Aims of this study are to: a) conduct validity analyses to determine sensitivity, specificity,
34
35 106 usability (test-retest reliability), and how each scale compares to the others and to the gold standard of
36
37 107 urine and hair drug testing in identifying prescription and illicit drug use; b) determine the impact of
38
39 108 clinic population variables (age, race, trimester of pregnancy) on validity of the three substance use
40
41 109 screeners; and c) assess birth outcomes (birth weight, gestational age, head circumference, and Neonatal
42
43 110 Intensive Care Unit (NICU) admissions) associated with the most widely used prescription drug and
44
45 111 multi-drug exposure.
46
47
48
49

50 112 **Study Design**

51
52 113 This study is a cross-sectional study that evaluates the sensitivity, specificity and usability of existing
53
54 114 substance use screeners. We chose this study design following an extensive search of the literature, an
55
56
57
58
59

1
2
3 115 overall assessment of feasibility and consultation with stakeholders (e.g., clinicians, pregnant women and
4
5 116 substance use researchers). We believe that a cross-sectional study such as ours is appropriate for the
6
7 117 evaluation of the accuracy and reliability of these screeners. We were also aided by the knowledge that
8
9 118 the prevalence of substance use in pregnancy is high¹. This implies that we are likely to obtain good
10
11 119 sensitivity and specificity estimates, with narrow confidence intervals, in a cross-sectional design which is
12
13 120 favorable in terms of cost and feasibility.

16 121 **Setting**

17
18 122 The study is being implemented at two urban obstetric clinics which serve diverse populations of
19
20 123 pregnant women. The study plans to recruit 500 participants to complete a demographics questionnaire,
21
22 124 followed by a randomized order of the NIDA Quick Screen/NIDA-modified ASSIST, 4 P's Plus, and
23
24 125 SURP-P. Participants are recruited during their regularly scheduled prenatal appointment, then contacted
25
26 126 again one week later by telephone to re-administer the screeners. Participants consent to multi-drug urine
27
28 127 testing, hair drug testing, and access to prescription drug and birth outcome data from electronic medical
29
30 128 records (EMR).

31
32
33
34 129 Recruitment Sites. We are recruiting participants from two obstetric outpatient clinics from January 2017-
35
36 130 January 2018. Currently all obstetric patients are screened for use of drugs, alcohol and tobacco at their
37
38 131 first prenatal visit by medical staff. Additionally, all new obstetric patients receive an in-depth evaluation
39
40 132 by a social worker which includes a more detailed assessment of both substance use and mental health
41
42 133 disorder history.

43
44
45 134 In the first clinic, which is the larger of the two clinics, most patients (97%) are publicly insured with
46
47 135 medical assistance and are over the age of 20 (80%). This clinic's population is primarily African-
48
49 136 American and low-income, all of whom undergo urine toxicological screening for substance use
50
51 137 identification. Based on preliminary data obtained from the clinic, about 950 individual obstetric patients
52
53 138 are cared for at this clinic annually. In the second (smaller) clinic, approximately 500 pregnant women are
54
55 139 cared for annually. Most patients (87%) have commercial insurance and 13% have either medical

1
2
3 140 assistance or Medicare. Most are over the age of 20 years (90%). Due to varying insurance coverage for
4
5 141 urine toxicology screens, patients in this office do not universally undergo urine toxicology screening but
6
7 142 all are screened for drug use using various interview techniques by their obstetric care providers at their
8
9 143 first prenatal visit. Based on historical data, we expect about 500 individual obstetric patients to be cared
10
11 144 for in this clinic across all trimesters of pregnancy in the one year of study recruitment.

12
13
14 145 Across both study sites, our source population covers a diverse set of participants and captures pregnant
15
16 146 women across all socioeconomic categories, insurance types, ethnicities and drug use patterns. This
17
18 147 ensures that our study results are generalizable to most populations of pregnant women.

19 20 21 148 **Study Population**

22
23 149 In the first clinic, of the estimated 950 individual obstetric patients cared for at this clinic annually, we
24
25 150 anticipated approaching 403 (50%), and expected 322 (80%) or more to agree to participate in this study.

26
27
28 151 In the second clinic, of the approximately 500 pregnant women cared for annually, we expect at least 450
29
30 152 (90%) to meet eligibility criteria. We anticipate approaching 225 pregnant women (50%) and expect 180
31
32 153 (80%) or more to agree to participate in this study.

33
34
35 154 Expected participation percentages are based on a similar grant-funded study that recruited pregnant
36
37 155 smokers from the same population and required consent for urine testing (cotinine) and birth data
38
39 156 abstraction from EMR.

40
41
42 157 Participant eligibility criteria include the following: a) currently pregnant (pre-determined by clinic staff);
43
44
45 158 b) age 18 or older; c) able to speak and understand English sufficiently to provide informed consent; and
46
47
48 159 d) natural hair length at least 3 cm to allow for substance use testing.

49
50
51 160 If eligibility criteria are met, research staff then obtain informed consent and medical releases for urine
52
53 161 collection, hair drug testing, and prescription drug and birth outcome data abstraction from the EMR.

162 Ethics and Dissemination

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

167 Study Procedures

168 *Approach*

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

180 *Recruitment*

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

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

190 *Self-Report Measures*

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

199 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

200

201 *Biochemical Measures*

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

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

	Drug class	Detection	Confirmation
U	Cocaine COC	2-4 Days	300 ng/mL
	Marijuana THC	15-30 Days	50 ng/mL
	Opiates OPI	2-4 Days	2000 ng/mL
R	Amphetamines AMP	2-4 Days	1000 ng/mL
	Methamphetamines AMP	3-5 Days	1000 ng/mL
I	Phencyclidine PCP	7-14 Days	25 ng/mL
	Benzodiazepines BZO	3-7 Days	300 ng/mL
N	Barbiturates BAR	4-7 Days	300 ng/mL
	Methadone MTD	3-5 Days	300 ng/mL
E	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
H	Marijuana THC	Up to 90 days	
	Amphetamines AMP	Up to 90 days	
A	Cocaine COC	Up to 90 days	
	Opiates OPI	Up to 90 days	
I	Phencyclidine PCP	Up to 90 days	
R			

209
 210 All women who screen positive on either biological multi-drug test or any one of the screeners
 211 are contacted immediately (for urine and screener results) or within 72 hours (for hair results) to detail the
 212 results of her test, encourage the participant to talk with her physician about her substance use, and offer
 213 her referrals to community resources for treatment that mirror what is currently given to patients by

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

216 *Birth Outcome Measures*

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

220 *Participant Follow-Up*

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

225 *Pilot Study*

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

229 We recruited 21 participants from each site for a total of 42 participants (Table 3). Mean age (sd)
230 of participants was 30.1 years (5.64). By race, 11 participants (26.2%) were White, 25 (59.5%) were
231 Black/African American, 4 (9.5%) were Asian, 1 (2.4%) was Hispanic and 1 (2.4%) was Other. About
232 24.4% tested positive for illicit drugs on urine testing, 22% tested positive on hair sample testing. Seven
233 (7) participants (16.7%) were lost to follow up.

234

235

236

237

238

239 *Table 3 Pilot Study Participant Characteristics*

Characteristics, N = 42	Clinical Site		
	<i>Clinic 1</i>	<i>Clinic 2</i>	<i>Both Sites</i>
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 (%)			
1 st	2 (9.5)	2 (9.5)	4 (9.5)
2 nd	6 (28.6)	6 (28.6)	12 (28.6)
3 rd	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)

240

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

246 *Power and Sample Size*

247 The sample size of 500 participants was chosen based on power analyses for the primary study questions.
 248 Based on a one-sample binomial approach, with a sample size of 500 participants, we can be 95%
 249 confident that the false negative rate in the population is under 10% (assuming no more than 35
 250 individuals test positive in the biologic drug tests without a positive survey screener result). Similarly, we
 251 can be 95% confident that the false negative rate in the population is under 5% (assuming no more than
 252 15 individuals test positive in the urine drug test without a positive survey screen result in the study).
 253 According to McNemar's Test, if at least 15% of the study participants have disagreement between any
 254 pair of survey results, 500 is a sufficient sample size to determine significant disagreement.

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

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

267 Analysis

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

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

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

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

312 Discussion

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

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

1
2
3 337 Confidentiality and ensured participants that their data will not be shared with anyone including clinic
4
5 338 staff. Finally, our study is limited to adults. Though our initial protocol included adolescents, the IRB did
6
7 339 not allow for “no-benefit” studies enrolling pregnant adolescents. This is an important area for further
8
9 340 exploration, given that pregnant adolescents report higher substance use rates than pregnant adults in
10
11 341 national surveys.

12
13
14 342 The primary innovation of this project is that it may provide a final evidence-based
15
16 343 recommendation for the tool(s) best suited for screening for illicit and prescription drugs among a diverse
17
18 344 sample of pregnant women. The provision of this evidence-based guidance to clinicians is a concrete
19
20 345 application of findings that is rare in public health research. We recognize that screening is a first step;
21
22 346 also important is the need for a public health focus on treatment of substance use during pregnancy to
23
24 347 enhance the odds of a successful pregnancy outcome. Barriers to treatment that are imperative to address
25
26 348 are the potential legal repercussions of identifying substance use during pregnancy that exist in some
27
28 349 states⁹ and unintentional breach of confidentiality.¹⁰ There is a strong need for a re-examination of state
29
30 350 policies so that women are not punished for having a treatment need.

31
32
33
34 351 Substance use during pregnancy, and specifically prescription and illicit drug use, are high
35
36 352 priority topics for the Centers for Disease Control and Prevention (CDC), WHO, The American Congress
37
38 353 of Obstetricians and Gynecologists (ACOG), Substance Abuse and Mental Health Services
39
40 354 Administration (SAMHSA), National Institute on Drug Abuse (NIDA), and the National Institutes of
41
42 355 Health (NIH). Universal screening has the potential to greatly enhance maternal and infant health
43
44 356 outcomes and reduce healthcare costs. Specifically, the current research supports the following Healthy
45
46 357 People 2020 public health goals and objectives which include reducing maternal illness and complications
47
48 358 due to pregnancy; increasing the proportion of pregnant women who receive adequate prenatal care;
49
50 359 increasing abstinence from alcohol, cigarettes, and illicit drugs among pregnant women; and increasing
51
52 360 the proportion of women delivering a live birth who received preconception care services and practiced
53
54 361 key recommended preconception health behaviors.

1
2
3 362 This research addresses an important problem by identifying a valid substance use screening
4
5 363 instrument for illicit and prescription drugs among pregnant women that is accurate, brief, and acceptable
6
7 364 to both patients and health care providers in a primary care setting. Identifying and validating one
8
9 365 instrument that functions the closest to the “gold standard” of biologic testing (i.e., urine and hair) and
10
11 366 disseminating this information widely will increase the likelihood that primary care clinics nationwide
12
13 367 may adopt a quick and easy screener universally. We may find that one instrument does not stand out but
14
15 368 that each has its distinct advantages and disadvantages; in this case, the performance of each measure will
16
17 369 be detailed with recommendations for which screener may work the best with a given population.
18
19
20
21 370
22 371
23
24 372
25
26 373
27
28 374
29
30 375
31
32 376
33
34 377
35
36 378
37
38 379
39
40 380
41
42 381
43
44 382
45
46 383
47
48 384
49
50 385
51
52 386
53
54 387
55
56 388
57
58
59
60

389

390 Declarations

391 *Ethics approval and consent to participate*

392 This study is approved by the Institutional Review Boards (IRB) of the University of Maryland (HP-
393 00072042), Baltimore; and Battelle Memorial Institute (0619-100106433). All participants are required to
394 give their informed consent prior to any study procedure.

395 *Availability of data and materials*

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

399 *Competing interests*

400 The authors declare that there are no competing interests.

401 *Authors' contributions*

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

404 *Funding*

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

409

410

411

412

413

414 REFERENCES

- 415 1. Substance Abuse and Mental Health Services Administration (SAMHSA). Results
416 from the 2012 National Survey on Drug Use and Health: Summary of national
417 findings. In: *NSDUH Series H-46, HHS Publication No.(SMA) 13-4795*. Substance
418 Abuse and Mental Health Services Administration Rockville, MD; 2013.
- 419 2. World Health Organization. Guidelines for the identification and management of
420 substance use and substance use disorders in pregnancy. 2014.
- 421 3. Scholle SH, Kelleher K. Assessing primary care performance in an
422 obstetrics/gynecology clinic. *Women & health*. 2003;37(1):15-30.
- 423 4. ACOG Committee on Health Care for Underserved Women. ACOG Committee
424 Opinion No. 524: opioid abuse, dependence, and addiction in pregnancy.
425 *Obstetric Anesthesia Digest*. 2013;33(2):79-80.
- 426 5. Group W. The alcohol, smoking and substance involvement screening test
427 (ASSIST): development, reliability and feasibility. *Addiction*. 2002;97(9):1183-
428 1194.
- 429 6. Smith PC, Schmidt SM, Allensworth-Davies D, Saitz R. A single-question screening
430 test for drug use in primary care. *Archives of internal medicine*.
431 2010;170(13):1155-1160.
- 432 7. Chasnoff I, Wells A, McGourty R, Bailey L. Validation of the 4P's Plus© screen for
433 substance use in pregnancy validation of the 4P's Plus. *Journal of Perinatology*.
434 2007;27(12):744.
- 435 8. Yonkers KA, Gotman N, Kershaw T, Forray A, Howell HB, Rounsaville BJ. Screening
436 for prenatal substance use: development of the Substance Use Risk Profile-
437 Pregnancy scale. *Obstetrics and gynecology*. 2010;116(4):827.
- 438 9. Guttmacher Institute. State Laws and Policies: Substance Use during Pregnancy.
439 2017; [https://www.guttmacher.org/state-policy/explore/substance-use-during-](https://www.guttmacher.org/state-policy/explore/substance-use-during-pregnancy)
440 [pregnancy](https://www.guttmacher.org/state-policy/explore/substance-use-during-pregnancy). Accessed December 18, 2017.
- 441 10. Knight KR. *Addicted. pregnant. poor*. Duke University Press; 2015.

442
443

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

444 Figure 1: Study procedures

445

446

For peer review only

1
2
3 447 Figure 2: Pilot study participation
4 448
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

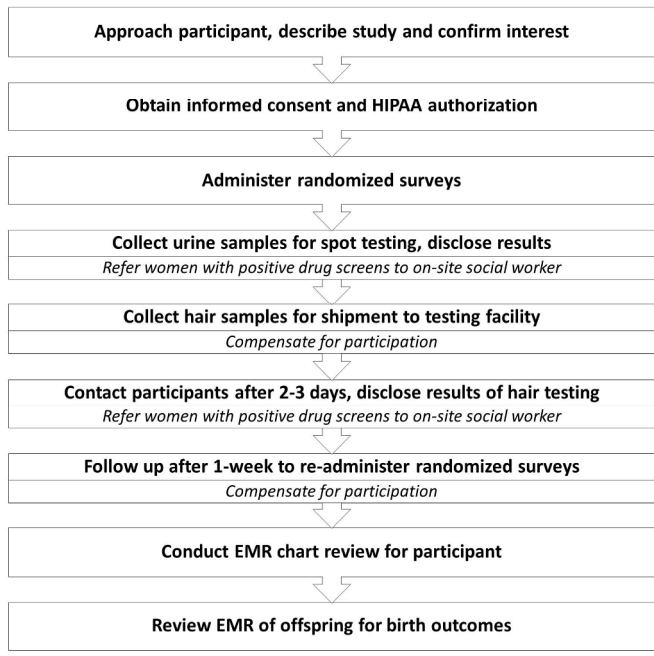


FIGURE 1: Study procedures

227x162mm (300 x 300 DPI)

ew only

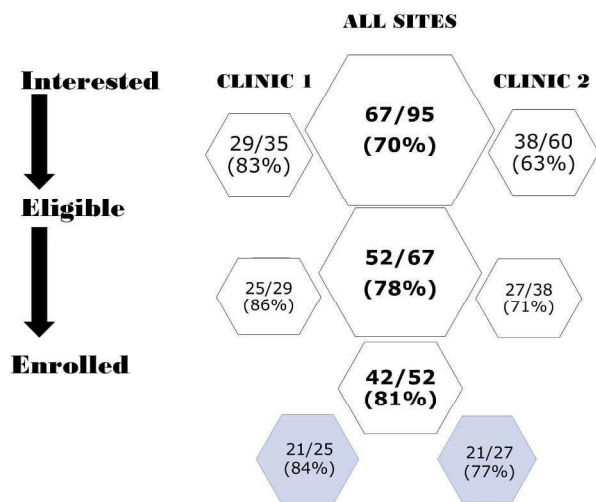


Figure 2: Pilot study participation

227x162mm (300 x 300 DPI)