



Published in final edited form as:

Addiction. 2019 September ; 114(9): 1683–1693. doi:10.1111/add.14651.

Accuracy of five self-report screening instruments for substance use in pregnancy

Steven J. Ondersma¹, Grace Chang², Tiffany Blake-Lamb³, Kathryn Gilstad-Hayden⁴, John Orav⁵, Jessica R. Beatty¹, Gregory L. Goyert⁶, Kimberly A. Yonkers^{4,7}

¹Merrill Palmer Skillman Institute and Department of Psychiatry and Behavioral Neurosciences, Wayne State University, Detroit, MI, USA,

²Department of Psychiatry, Harvard Medical School, Boston, MA, USA,

³Department of Obstetrics and Gynecology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA,

⁴Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA,

⁵Department of Medicine (Biostatistics) and Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA,

⁶Department of Obstetrics and Gynecology, Henry Ford Health System, Detroit, MI, USA

⁷Department of Obstetrics, Gynecology, and Reproductive Sciences, Yale University School of Medicine, New Haven, CT, USA

Abstract

Background and aims—The accuracy of current screening instruments for identification of substance use in pregnancy is unclear, particularly given methodological shortcomings in existing research. This diagnostic accuracy study compared five existing instruments for ability to identify illicit drug, opioid and alcohol use, under privacy expectations consistent with applied practice and using a gold standard incorporating toxicological analysis.

Design—Prospective cross-sectional screening accuracy study.

Setting—Three sites encompassing four prenatal care clinics in the United States.

Participants—Convenience sample of 1220 racially, ethnically and socio-economically diverse pregnant women aged 18 years and over.

Measurements—In Phase I, participants completed the five screening instruments in counterbalanced order. Instruments included the Substance Use Risk Profile—Pregnancy (SURP-P), CRAFFT (acronym for five-item screener with items related to car, relax, alone, forget,

Correspondence to: Steven J. Ondersma, Merrill-Palmer Skillman Institute, Wayne State University, 71 East Ferry Avenue, Detroit, MI 48202, USA. s.ondersma@wayne.edu.

Declaration of interests

S.J.O. discloses that he is part owner of Interva, Inc., which markets e-intervention authoring software (not used for the present study). K.A.Y. and G.C. disclose royalties from Up-To-Date. Other authors report no conflicts of interest.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

friends and trouble), 5Ps (parents, peers, partner, pregnancy, past), Wayne Indirect Drug Use Screener (WIDUS) and the National Institute on Drug Abuse (NIDA) Quick Screen. In Phase II, participants provided a urine sample and completed a calendar recall-based interview regarding substance use. These screeners were tested, using receiver operating characteristic (ROC) analysis and accuracy statistics, against a reference standard consisting of substance use in three classes (illicit drugs, opioids and alcohol), considered positive if use was evident via 30-day calendar recall or urine analysis.

Findings—Three hundred and fifteen of 1220 participants (26.3%) met reference standard criteria for positivity. The single-item screening questions from the NIDA Quick Screen showed high specificity (0.99) for all substances, but very poor sensitivity (0.10–0.27). The 5Ps showed high sensitivity (0.80–0.88) but low specificity (0.35–0.37). The CRAFFT, SURP-P and 5Ps had the highest area under the curve (AUC) for alcohol (0.67, 0.66 and 0.62, respectively), and the WIDUS had the highest AUC for illicit drugs and opioids (0.70 and 0.69, respectively). Performance of all instruments varied significantly with race, site and economic status.

Conclusions—Of five screening instruments for substance use in pregnancy tested (Substance Use Risk Profile—Pregnancy (SURP-P), CRAFFT, 5Ps, Wayne Indirect Drug Use Screener (WIDUS) and the National Institute on Drug Abuse (Quick Screen), none showed both high sensitivity and high specificity, and area under the curve was low for nearly all measures.

Keywords

Alcohol; illicit drugs; opioids; pregnancy; risk identification; screening; substance use

INTRODUCTION

Data from the National Survey on Drug Use and Health indicate that approximately 285 000 pregnant women (12.6%) used alcohol or illicit drugs in the past 30 days [1]. Under-reporting is common, however, meaning that the actual numbers are almost certainly higher [2–5]. Although effects are inconsistent and often subtle, substance use during pregnancy can lead to adverse infant and child outcomes [6–10] as well as negative effects for women themselves [11–13]. Early and accurate identification of substance use in pregnancy is critical.

Existing research into screening pregnant women for substance use has several limitations. First, few studies evaluate detection of illicit or prescription drug use during pregnancy [14–16]. Secondly, most studies use structured diagnostic interviews as a reference standard rather than substance use itself. Substance use is the primary interest during pregnancy because of its adverse fetal impacts. Additional shortcomings include a lack of direct comparisons between instruments, infrequent inclusion of biological measures of substance use and use of single-site or homogeneous cohorts. Finally, previous studies have neglected subtle but important differences in expectations of privacy between research and clinical settings.

These shortcomings highlight the clear need for stronger data to inform health-care-based screening practices with pregnant women, particularly given recent increases in opioid use

during pregnancy [17]. The present study sought to address this need via a prospective multi-site comparison of multiple screening instruments' ability to detect illicit drug and opioid use, as well as alcohol use, among pregnant women.

METHODS

Sites and participants

Participants were a prospective convenience series of pregnant women seeking prenatal care at one of three sites (Yale New Haven Health System in New Haven, CT; Massachusetts General Hospital in Boston, MA; and the Henry Ford Health System in Detroit, MI). Recruitment took place between February 2016 and April 2017. To be included, participants had to be pregnant, 18 years or older and able to understand English. Women were excluded if they showed cognitive impairment, were currently hospitalized, were considering either termination of the pregnancy or adoption or did not provide consent.

Procedure

This study used a previously described approach [5,18] to maximize generalizability by administering measures under the same privacy expectations typically found in clinical practice, and without foreknowledge of a pending urine sample request; and to also maximize participant protection from research-related harm. Achieving these goals required initially implying to participants that their responses would be shared with their doctor, when in fact all data were anonymous; and asking for voluntary provision of a urine sample only after completion of the screening instruments. This approach was implemented at the Detroit and New Haven sites, with approval from the Institutional Review Boards at Yale, Wayne State and the Henry Ford Health System. At the Boston site, participants were told during the consent process that the first phase was completely anonymous. This approach was approved by the Institutional Review Board of Partners Healthcare (Massachusetts General Hospital).

Pregnant women were approached during a regularly scheduled prenatal care visit, either in the waiting area or after being told of the study by clinic staff. Research assistants described the study as having two phases, noting that the second phase would be described in detail after conclusion of the first, and that the participant would be free to decline at that time. Participants were told that Phase 1 involved providing answers regarding substance use using an iPad and a gift card worth \$10. Participants in Boston were told that all participation was anonymous. Participants in Detroit and New Haven were initially told that their answers would be shared with clinic staff but would be kept confidential. However, after completion of Phase 1, participants in Detroit and New Haven were then told that their responses would not be shared with staff, that no identifying information was being retained and that this was conducted to provide as much protection for them as possible while also administering the measures in a realistic way. Participants were debriefed by asking if they had any concerns or negative reactions. As with past studies using this approach [5,19], no participants expressed concern. Screening instruments were administered in counterbalanced order (all possible orders equally represented) at all sites.

At all sites, participants who completed Phase 1 were then asked to complete calendar-based recall of substance use via the iPad and to provide a urine sample. Those who provided consent for and completed Phase 2 received a \$30 gift card. Use of an external laboratory and direct self-report via iPads effectively prevented investigator bias during collection of either the index tests or reference standards. As results were not known to study staff at this point, there was no obligation to report substance use in pregnancy to child protective services. All participants were given a list of free or reduced-cost services in their community.

Screening instrument selection

This study included all known non-proprietary screeners with published evidence for use in identifying substance use (not just alcohol) in pregnancy. We found three measures meeting these criteria: the Substance Use Risk Profile–Pregnancy (SURP-P) [20], the Wayne Indirect Drug Use Screener (WIDUS) [5] and the CRAFFT (acronym for five-item screener with items related to car, relax, alone, forget, friends and trouble) [21,22]. We also included the 5Ps (parents, peers, partner, pregnancy, past) screener [23,24], because it is widely used with pregnant women, making rigorous evaluation necessary. All four measures were used with their original time-frame and cut score.

All these measures include one or more items evaluating substance use consequences or correlates of substance use rather than focusing exclusively on use itself. Further, each of these measures uses a life-time/before pregnancy time-frame rather than asking specifically about use during pregnancy. Both these approaches are intended to increase sensitivity, given the prevalence of under-reporting in pregnancy. Further, in the case of the SURP-P and WIDUS, items with these characteristics were selected empirically from among a larger set of items. This similarity among all four measures suggests the need for a comparator asking directly about substance use during pregnancy. We thus included the National Institute on Drug Abuse (NIDA) Quick Screen [25], a parsimonious screener that directly measures substance use. Further, to focus specifically on use during pregnancy, the NIDA Quick Screen response options were modified to refer to the past month rather than the past year. Each of these five measures is reviewed below and in Table 1 (Figure 1).

SURP-P—The three-item SURP-P (Table 1) was developed in a training sample of 1610 pregnant women and cross-validated in a separate validation sample of 1704 pregnant women. On cross-validation, the SURP-P (using the high-risk cut-off) identified alcohol use with a sensitivity of 48% and specificity of 85%, and identified marijuana use with a sensitivity of 68% and specificity of 86%; it also outperformed the original measures from which its items were drawn [20].

WIDUS—The WIDUS is an indirect screening instrument that seeks to identify risk in the perinatal period by asking about correlates of drug use [5]. On cross-validation, the WIDUS showed sensitivity of 76%, specificity of 68% and area under the receiver operating curve (AUC) of 0.74. The WIDUS also outperformed the 10-item version of the Drug Abuse Screening Test (DAST-10), and WIDUS scores showed a strong linear association with likelihood of positive toxicology [5].

CRAFFT questionnaire—The CRAFFT (a mnemonic representing each item in this six-item instrument: car, relax, alone, forget, friends and trouble) was originally designed to detect adolescent substance use risk [21,26,27]. Items were selected from existing measures and reduced using a self-report criterion [21]. In a pilot study ($n = 30$), the CRAFFT showed utility in identifying substance use among pregnant young adults, using both calendar based-recall [positive predictive value (PPV) of 90% and negative predictive value (NPV) of 80%] and diagnostic interview (PPV of 58% and NPV of 83%) reference standards [22].

The 5Ps—The 5Ps Prenatal Substance Abuse Screen for Alcohol and Drugs [23,24] (with ‘5Ps’ being a mnemonic representing each question in this five-item measure: parents, peers, partner, pregnancy, past) is an adaptation of an earlier measure (the 4Ps) designed for use in pregnancy. Although not separately validated, the 5Ps overlaps significantly with closely related measures that have reported utility in identifying substance use in pregnancy [28,29], and is in wide use in several states for identification of substance use in pregnancy. Data on the validity of the 5Ps are clearly needed.

NIDA quick screen—The NIDA Quick Screen [25] consists of four questions regarding the frequency with which respondents have used illegal drugs, prescription drugs for non-medical reasons, tobacco or multiple drinks containing alcohol (four or more drinks in a day). The alcohol and drug use items have been validated as single-item measures [30–32]. As noted, although the Quick Screen has not been validated among pregnant women, its parsimony, strong performance in other samples and direct measurement of in-pregnancy substance use make it an important comparator.

Study protocol—The full study protocol can be accessed by contacting K.A. Y. at kimberly.yonkers@yale.edu.

Reference standard

Consistent with recommendations for measurement of drug use [33,34], we used an a priori reference standard in which participants were considered positive if either urine drug screening [35] (enzyme immunoassay; Redwood Toxicology Laboratory, Santa Rosa, CA, USA) or 30-day calendar-based self-report recall [36–38] were positive. Agreement between self-report and urine drug screen results was high, at 90.9% for illicit drugs, 88.1% for alcohol and 97.0% for opioids. The illicit drug use reference standard combined self-report of illicit drug use in the past 30 days (marijuana, cocaine, heroin, amphetamines, barbiturates and hallucinogens) with testing of urine samples for amphetamines, barbiturates, cocaine and marijuana. The opioid reference standard combined self-report of prescription opioid misuse in the past 30 days with any evidence of opioid use from the toxicology screen. Alcohol was considered positive if either calendar-based recall for the past 30 days or urine screening (direct testing for ethanol) was positive. For participants missing either urine screening ($n = 4$) or self-report data ($n = 1$), the reference standard was based upon the available measure.

Statistical analyses

Following a priori power analysis, the study was designed to recruit 400 women with drug or alcohol use so that a 7.5% difference in sensitivity between two screeners could be detected with 91% power. Each measure's association with the reference standard in each category (illicit drugs, alcohol and opioids) was first evaluated using logistic regression models, which were used to calculate AUC. Sensitivity, specificity, PPV, NPV, raw accuracy (percentage of correct classifications) and 95% confidence intervals (CI) were then calculated using published cut scores (an a priori decision). Individual items from the NIDA Quick Screen were tested for their ability to detect their respective reference standard (e.g. the drug use item was tested for associations with drug use).

Direct comparisons between screening instruments were calculated separately for each of the three reference standards and focused on AUC as the best overall measure of classification performance, particularly given its stability across cut score and base rate. Differences in AUC between measures were compared by calculating Z -scores

$$\left(\frac{AUC_1 - AUC_2}{\sqrt{(SE_1)^2 + (SE_2)^2}} \right)$$
 for each pairwise comparison. Comparisons of sensitivity, specificity and

accuracy between screening instruments, as well as pre-planned analysis of moderation of differences in screener performance by race/ethnicity, site, trimester and current receipt of any form of public assistance, were tested with generalized estimating equation (GEE) models specifying a binomial distribution, logit link and exchangeable correlation structure. Two-sided P -values of < 0.05 were considered significant. Analyses were conducted with SAS software version 9.4 (SAS Institute Inc., Cary, NC, USA).

Sample derivation and missing data

We approached 3429 women for screening; 1101 did not meet eligibility criteria, 941 declined or did not provide consent, 149 did not complete the study and 18 were excluded for other reasons (e.g. technical problems with tablet, ineligibility discovered), leaving a final sample of 1220. Of the 149 incompletes, 57 were unable to complete Phase 1 and 13 were unable to complete Phase 2 due to time constraints. After completing Phase 1, 55 women declined to complete Phase 2 and 24 were unable to provide a urine sample. When calculating performance measures, participants with incomplete data for a given screener were excluded from analyses for that screener. The numbers excluded ranged from four for 5Ps to 21 for NIDA Quick Screen-illegal drugs (Fig. 2). GEE models used all available screening information from each participant, even if individual screen results were missing. However, women with missing covariates ($n = 33$) were dropped from GEE analyses.

RESULTS

As seen in Table 2, the final sample included 1220 pregnant women, most of whom were either non-Hispanic African American ($n = 480$, 40.1%) or non-Hispanic white ($n = 444$, 37.1%). Fewer than half the participants ($n = 539$, 44.7%) received some form of public assistance. A total of 315 (26.3%) was positive for alcohol, opioids or other drugs, by either self-report or urine drug screen, a rate that driven in part by high rates of marijuana use at the New Haven and Detroit sites (182 participants in total were positive for marijuana). We

did not identify any adverse results (significant distress, socio-legal consequences or breach of confidentiality) as a consequence of participation.

Identification of illicit drug use

Each screener's ability to identify illicit drug use and comparisons between screeners are presented in Table 3. For illicit drug use, significant differences between measures in sensitivity (type 3 $\chi^2 = 102.5$, $P < 0.001$) and specificity (type 3 $\chi^2 = 636.3$, $P < 0.001$) were detected, adjusting for trimester, site, race/ethnicity and receipt of public assistance. The single item from the NIDA Quick Screen showed the highest overall specificity (0.99, 95% CI = 0.98, 1.00), but the lowest sensitivity (0.27, 95% CI = 0.21, 0.34). Conversely, the 5Ps had the highest sensitivity (0.80, 95% CI = 0.73, 0.85), but the lowest specificity (0.37, 95% CI = 0.34, 0.40). No screening instruments were both highly sensitive and specific. WIDUS had the highest AUC (0.70, 95% CI = 0.67, 0.74).

Identification of alcohol use

Performance measures for identification of alcohol use and comparisons among measures are also presented in Table 3. Sensitivity and specificity differed significantly between measures (type 3 χ^2 Quick Screen showed the highest specificity (0.99, 95% CI = 0.98, 0.99) but lowest sensitivity (0.10, 95% CI = 0.05, 0.16), and 5Ps had the highest sensitivity (0.88, 95% CI = 0.81, 0.92) but lowest specificity (0.37, 95% CI = 0.34, 0.40). AUC was highest for CRAFFT (0.67, 95% CI = 0.62, 0.71), SURP-P (0.66, 95% CI = 0.62, 0.70), and 5Ps (0.62, 95% CI = 0.59, 0.65) and no instrument had both high sensitivity and high specificity.

Identification of opioid use—Each screener's ability to identify opioid use is also presented and compared in Table 3. Results followed the same pattern as the other illicit drugs outcome, with significant differences in sensitivity (type 3 $\chi^2 = 28.7$, $P < 0.001$) and specificity (type 3 $\chi^2 = 745.3$, $P < 0.001$) between questionnaires. The NIDA Quick Screen again showed the highest specificity (0.99, 95% CI = 0.98, 0.99) but lowest sensitivity (0.16, 95% CI = 0.07, 0.31), and the 5Ps showed the highest sensitivity (0.81, 95% CI = 0.67, 0.92) but lowest specificity (0.35, 95% CI = 0.32, 0.38). WIDUS had the highest AUC (0.69, 95% CI = 0.62, 0.77).

Moderators of screening instrument performance

Table 4 presents performance of each measure in predicting illicit drug use by race/ethnicity, site, trimester and public assistance, as well as the P -value for the GEE interaction term testing if these factors moderated performance differences between screeners. Two-way interactions were present for at least one measure of merit (overall accuracy, sensitivity or specificity) for the potential moderators examined. For example, as seen in Table 4, differences in accuracy varied between screeners according to race/ethnicity, with accuracy for white women being higher with the WIDUS and NIDA Quick Screen and lower for the SURP-P, CRAFFT and 5Ps compared to black and Hispanic women. Similarly, specificity was lowest among white participants for the SURP-P, CRAFFT, and 5Ps, but not for WIDUS and NIDA; and the SURP-P showed greater specificity in New Haven than in

Boston. Additionally, specificity was highest in the third trimester SURPP, but it was lowest in the third trimester for NIDA.

DISCUSSION

In this pragmatic multi-site study addressing shortcomings from prior research, none of the tested measures exceeded an AUC of 0.70 (often interpreted as the cut-off between ‘poor’ and ‘fair’ accuracy) for any outcome. Using published cut scores, some screening instruments were highly specific and others were highly sensitive, but none were both. This level of performance is lower than reported in previous studies [30], probably because of differences in samples (pregnant women rather than general adults) and reference standards (substance use rather than presence of a substance use disorder). As noted above, most studies evaluating screening tools use structured interview-based reference standards. Women who are willing to disclose substance use or related consequences on a short questionnaire are likely to also do so on a longer questionnaire, and those denying it on a short questionnaire are likely to do so on a longer measure. This can result in superficially elevated accuracy statistics that do not capture respondents who consistently fail to disclose substance use or consequences. Although there is certainly value in identification of substance use disorders in pregnancy, women who do not meet criteria for a use disorder could nevertheless engage in risky use, and many women meeting criteria for disordered use nevertheless cut down or quit during pregnancy (diagnostic criteria refer to symptoms at any time in the past 12 months).

Further, measures evaluated in this study showed significant variations in performance as a function of race, trimester, receipt of public assistance and site. This variability could lead to significant differences in positive outcomes such as treatment provision, as well as in negative consequences such as involvement with child protective services. Implicit assumptions of the invariance of screener accuracy across such factors appear unfounded. The origins of this variability are unclear, but may include heterogeneity in rates and types of substance use; contextual differences, especially those related to the socio-legal consequences of substance use in pregnancy; and cultural influences. Future measure development should consider consistency across such factors when evaluating potential items.

Future research should also seek to improve the performance of self-report screening instruments for use in pregnancy. The high sensitivity of some measures in this analysis, and the high specificity of others, suggests that items in existing measures should be reviewed for possible combination into a single tool. Inclusion of demographic characteristics such as age—which is strongly associated with drug use in pregnancy [1]—or use of scoring algorithms rather than a simple sum may also be helpful. As a further challenge, screening approaches must also address obstacles to integrating evidence-based screening into clinical settings; in a 2010 survey, only 10.6% of obstetrician–gynecologists reported using a validated screening tool to assess alcohol risk [39]. Mere identification of a superior screening tool is insufficient; valid tools must also be embedded in delivery mechanisms and administrative environments that support their use.

Several limitations must be highlighted. First, although the multi-site design resulted in a large and diverse sample, the sample was non-representative and excluded non-English speakers. Secondly, because of variability in Institutional Review Board (IRB) requirements, participants at the Boston site knew their responses were anonymous when completing screening. Although sample differences between sites preclude analysis of whether this variation affected outcomes, it should be noted that—in a sensitivity analysis—AUCs for only the Detroit and New Haven sites were nearly identical to AUCs for all three sites combined (Supporting information, Table S1). Thirdly, although our use of urine drug screening is a significant strength, it provided only a short window of detection. For example, ethanol is only detectable in urine for a few hours after ingestion; marijuana is typically detectable for up to 2 weeks, and most other drugs are detectable for only 24–48 hours [40]. Overall rates of use, and thus screening instrument performance, may have been affected by our inability to detect drug use beyond a brief window. Fourthly, we were unable to distinguish between urine samples that were positive because of illicit opioid use, prescription opioid misuse or opioid use under a physician’s guidance (e.g. buprenorphine). Information from the participant’s medical records would have increased our ability to make these distinctions. However, because only 20 of 1220 urine samples (1.6%) were positive for opioids of any kind, this limitation is unlikely to have significantly affected results. (This value is consistent with findings from the 2016 National Survey on Drug Use and Health, in which 1.2% of pregnant women reported use of heroin or misuse of opioid pain relievers [41].)

Although the performance of some instruments on some measures of merit was strong, none showed either adequate or consistent performance, and none can currently be recommended for applied practice with pregnant women. The challenge is particularly marked with opioid use, which is more difficult to identify accurately given its low prevalence in most areas. Until screeners are identified that can improve on direct, face-valid questions, the NIDA Quick Screen—which does nothing more than quickly query frequency of self-reported use—appears to be the best approach to take. Future research should consider a broader range of candidate items in developing screening tools; should consider traditional as well as algorithm scoring; and should take implementation challenges into account.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors wish to acknowledge the invaluable contributions of Ms Edith L. Combs (Henry Ford Health System) and Ms Erica Hvizdos (Wayne State University). Funding for this study was provided by the Centers for Disease Control and Prevention, award #DP006082 to K.A.Y. S.J.O. also received support from the Helene Lycacki/Joe Young Sr funds from the State of Michigan. Neither funding source had any role in study design, manuscript writing, publication decisions or data collection, analysis or interpretation.

References

1. Center for Behavioral Health Statistics and Quality National Survey on Drug Use and Health: Detailed Tables. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2016, p. 2017.

2. Lendoiro E, Gonzalez-Colmenero E, Concheiro-Guisan A, de Castro A, Cruz A, Lopez-Rivadulla Met al. Maternal hair analysis for the detection of illicit drugs, medicines, and alcohol exposure during pregnancy. *Ther Drug Monit* 2013; 35: 296–304. [PubMed: 23666580]
3. Bessa MA, Mitsuhiro SS, Chalem E, Barros MM, Guinsburg R, Laranjeira R Underreporting of use of cocaine and marijuana during the third trimester of gestation among pregnant adolescents. *Addict Behav* 2010; 35: 266–9. [PubMed: 19896774]
4. Young-Wolff KC, Tucker LY, Alexeeff S, Armstrong MA, Conway A, Weisner Cet al. Trends in self-reported and biochemically tested marijuana use among pregnant females in California from 2009–2016. *JAMA* 2017; 318: 2490–1. [PubMed: 29279917]
5. Ondersma SJ, Svikis DS, LeBreton JM, Streiner DL, Grekin ER, Lam PK et al. Development and preliminary validation of an indirect screener for drug use in the perinatal period. *Addiction* 2012; 107: 2099–106. [PubMed: 22882721]
6. Donald KA, Roos A, Fouche JP, Koen N, Howells FM, Woods R Pet al. A study of the effects of prenatal alcohol exposure on white matter microstructural integrity at birth. *Acta Neuropsychiatr* 2015; 27: 197–205. [PubMed: 26022619]
7. Khoury JE, Milligan K, Girard TA Executive functioning in children and adolescents prenatally exposed to alcohol: a meta-analytic review. *Neuropsychol Rev* 2015; 25: 149–70. [PubMed: 26037669]
8. El Marroun H, Schmidt MN, Franken IH, Jaddoe VW, Hofman A, van der Lugt A et al. Prenatal tobacco exposure and brain morphology: a prospective study in young children. *Neuropsychopharmacology* 2014; 39: 792–800. [PubMed: 24096296]
9. Frank DA, Kuranz S, Appugliese D, Cabral H, Chen C, Crooks Det al. Problematic substance use in urban adolescents: role of intrauterine exposures to cocaine and marijuana and post-natal environment. *Drug Alcohol Depend* 2014; 142: 181–90. [PubMed: 24999059]
10. Diaz SD, Smith LM, LaGasse LL, Derauf C, Newman E, Shah Ret al. Effects of prenatal methamphetamine exposure on behavioral and cognitive findings at 7.5 years of age. *J Pediatr* 2014; 164: 1333–8. [PubMed: 24630350]
11. Bauer CR, Shankaran S, Bada HS, Lester B, Wright LL, Krause-Steinrauf Het al. The maternal lifestyle study: drug exposure during pregnancy and short-term maternal outcomes. *Am J Obstet Gynecol* 2002; 186: 487–95. [PubMed: 11904612]
12. Sordo L, Indave BI, Degenhardt L, Barrio G, Kaye S, Ruiz-Perez I et al. A systematic review of evidence on the association between cocaine use and seizures. *Drug Alcohol Depend* 2013; 133: 795–804. [PubMed: 24051062]
13. Pinto SM, Dodd S, Walkinshaw SA, Siney C, Kakkar P, Mousa HA Substance abuse during pregnancy: effect on pregnancy outcomes. *Eur J Obstet Gynecol Reprod Biol* 2010; 150: 137–41. [PubMed: 20227162]
14. Sarkar M, Einarson T, Koren G Comparing the effectiveness of TWEAK and T-ACE in determining problem drinkers in pregnancy. *Alcohol Alcohol* 2010; 45: 356–60. [PubMed: 20497951]
15. Bradley KA, Boyd-Wickizer J, Powell SH, Burman ML Alcohol screening questionnaires in women: a critical review. *JAMA* 1998; 280: 166–71. [PubMed: 9669791]
16. Russell M, Martier SS, Sokol RJ, Mudar P, Jacobson S, Jacobson J Detecting risk drinking during pregnancy: a comparison of four screening questionnaires. *Am J Public Health* 1996; 86: 1435–9. [PubMed: 8876514]
17. Patrick SW, Davis MM, Lehmann CU, Cooper WO Increasing incidence and geographic distribution of neonatal abstinence syndrome: United States 2009 to 2012. *J Perinatol* 2015; 35: 650–5. [PubMed: 25927272]
18. Grekin ER, Svikis DS, Lam P, Connors V, Lebreton JM, Streiner D Let al. Drug use during pregnancy: validating the drug abuse screening test against physiological measures. *Psychol Addict Behav* 2010; 24: 719–23. [PubMed: 21198230]
19. Beatty JR, Chase SK, Ondersma SJ A randomized study of the effect of anonymity, quasi-anonymity, and certificates of confidentiality on postpartum women's disclosure of sensitive information. *Drug Alcohol Depend* 2014; 134: 280–4. [PubMed: 24246900]

20. 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. *Obstet Gynecol* 2010; 116: 827–33. [PubMed: 20859145]
21. Knight JR, Shrier LA, Bravender TD, Farrell M, Vander Bilt J, Shaffer HJA new brief screen for adolescent substance abuse. *Arch Pediatr Adolesc Med* 1999; 153: 591–6. [PubMed: 10357299]
22. Chang G, Orav EJ, Jones JA, Buynitsky T, Gonzalez S, Wilkins-Haug L Self-reported alcohol and drug use in pregnant young women: a pilot study of associated factors and identification. *J Addict Med* 2011; 5: 221–6. [PubMed: 21844837]
23. Kennedy C, Finkelstein N, Hutchins E, Mahoney J Improving screening for alcohol use during pregnancy: the Massachusetts ASAP program. *Matern Child Health J* 2004; 8: 137–47. [PubMed: 15503394]
24. Watson E, Barnes H, Brown E, Kennedy C, Finkelstein N Alcohol Screening Assessment in Pregnancy: the ASAP Curriculum. Cambridge, MA: Institute for Health and Recovery; 2003.
25. National Institute on Drug Abuse (NIDA). Resource Guide: Screening for Drug Use in General Medical Settings 2012 [National Institute on Drug Abuse website]. Available at: <https://www.drugabuse.gov/publications/resource-guide-screening-drug-use-in-general-medical-settings/nida-quick-screen> (accessed 24 March 2019) (Archived at <http://www.webcitation.org/78Vt0b8bi> on 20 May 2019).
26. Knight JR, Sherritt L, Harris SK, Gates EC, Chang G Validity of brief alcohol screening tests among adolescents: a comparison of the AUDIT, POSIT, CAGE, and CRAFFT. *Alcohol Clin Exp Res* 2003; 27: 67–73. [PubMed: 12544008]
27. Knight JR, Sherritt L, Shrier LA, Harris SK, Chang G Validity of the CRAFFT substance abuse screening test among adolescent clinic patients. *Arch Pediatr Adolesc Med* 2002; 156: 607–14. [PubMed: 12038895]
28. Chasnoff IJ, Wells AM, McGourty RF, Bailey LK Validation of the 4P's plus screen for substance use in pregnancy validation of the 4P's plus. *J Perinatol* 2007; 27: 744–8. [PubMed: 17805340]
29. Chasnoff IJ, McGourty RF, Bailey GW, Hutchins E, Lightfoot SO, Pawson L Let al. The 4P's plus screen for substance use in pregnancy: clinical application and outcomes. *J Perinatol* 2005; 25: 368–74. [PubMed: 15703775]
30. Saitz R, Cheng DM, Allensworth-Davies D, Winter MR, Smith PC The ability of single screening questions for unhealthy alcohol and other drug use to identify substance dependence in primary care. *J Stud Alcohol Drugs* 2014; 75: 153–7. [PubMed: 24411807]
31. Smith PC, Schmidt SM, Allensworth-Davies D, Saitz RA single-question screening test for drug use in primary care. *Arch Intern Med* 2010; 170: 1155–60. [PubMed: 20625025]
32. Smith PC, Schmidt SM, Allensworth-Davies D, Saitz R Primary care validation of a single-question alcohol screening test. *J Gen Intern Med* 2009; 24: 783–8. [PubMed: 19247718]
33. Hoeke H, Roeder S, Bertsche T, Lehmann I, Borte M, von Bergen Met al. Monitoring of drug intake during pregnancy by questionnaires and LC-MS/MS drug urine screening: evaluation of both monitoring methods. *Drug Test Anal* 2015; 7: 695–702. [PubMed: 25545167]
34. Donovan DM, Bigelow GE, Brigham GS, Carroll KM, Cohen AJ, Gardin J Get al. Primary outcome indices in illicit drug dependence treatment research: systematic approach to selection and measurement of drug use end-points in clinical trials. *Addiction* 2012; 107: 694–708. [PubMed: 21781202]
35. Wolff K, Farrell M, Marsden J, Monteiro MG, Ali R, Welch Set al. A review of biological indicators of illicit drug use, practical considerations and clinical usefulness. *Addiction* 1999; 94: 1279–98. [PubMed: 10615715]
36. Fals-Stewart W, O'Farrell TJ, Freitas TT, McFarlin SK, Rutigliano P The timeline followback reports of psychoactive substance use by drug-abusing patients: psychometric properties. *J Consult Clin Psychol* 2000; 68: 134–44. [PubMed: 10710848]
37. Brown RA, Burgess ES, Sales SD, Whiteley JA, Evans DM, Miller IW Reliability and validity of a smoking timeline follow-back interview. *Psychol Addict Behav* 1998; 12: 101–12.
38. Sobell L, Brown J, Leo G, Sobell M The reliability of the alcohol timeline follow back when administered by telephone and by computer. *Drug Alcohol Depend* 1996; 42: 49–54. [PubMed: 8889403]

39. Anderson BL, Dang EP, Floyd RL, Sokol R, Mahoney J, Schulkin J Knowledge, opinions, and practice patterns of obstetrician-gynecologists regarding their patients' use of alcohol. *J Addict Med* 2010; 4: 114–21. [PubMed: 21769028]
40. Vandevenne M, Vandenbussche H, Verstraete A Detection time of drugs of abuse in urine. *Acta Clin Belg* 2000; 55: 323–33. [PubMed: 11484423]
41. Substance Abuse and Mental Health Services Administration. Key Substance Use And Mental Health Indicators in the United States: Results from the 2016 National Survey on Drug Use and Health (HHS Publication no. SMA 17–5044, NSDUH Series H-52) Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration; 2017. Available at: <http://datafiles.samhsa.gov/> (accessed 4 March 2019) (Archived at <http://www.webcitation.org/78VvTLYix> on 20 May 2019).

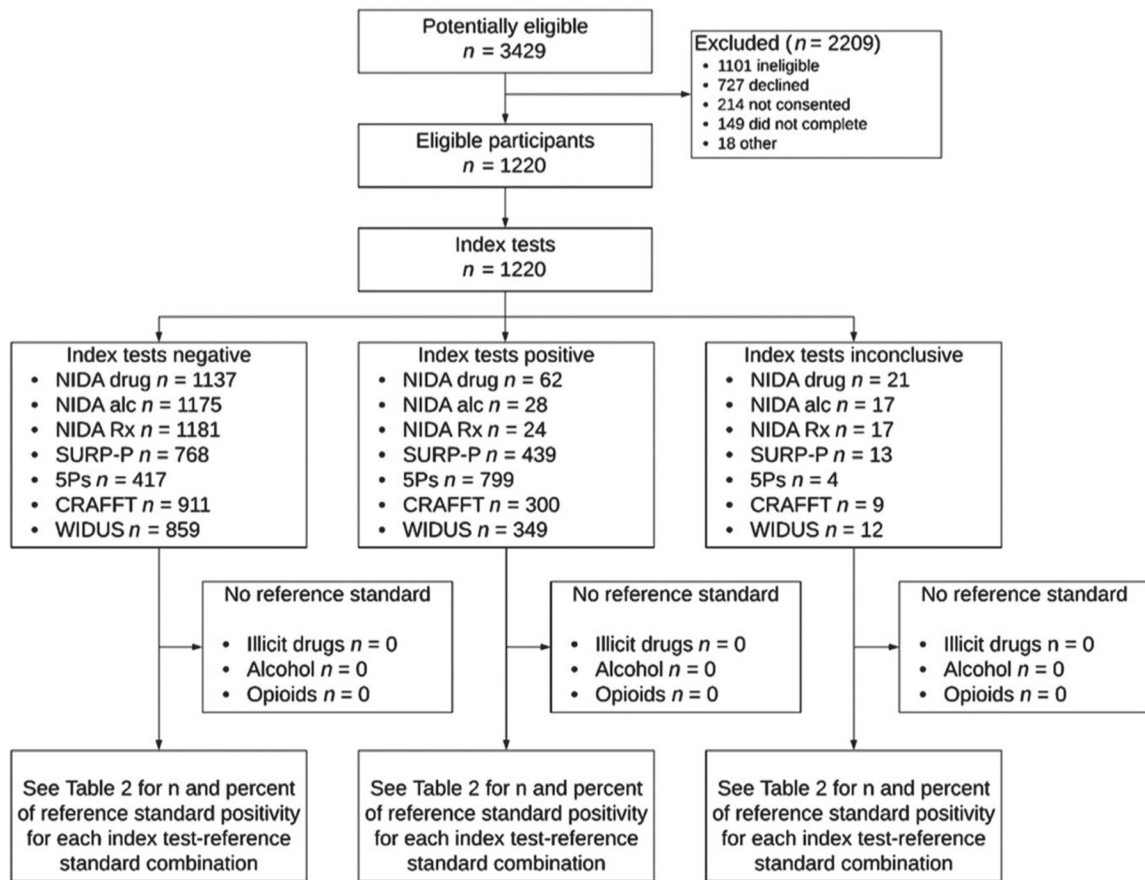


Figure 1.

Flow diagram of study comparing self-report screening instruments for detection of substance use among pregnant women. National Institute on Drug Abuse (NIDA) drug = single drug use screening question from NIDA Quick Screen; NIDA alc = single alcohol use screening question from NIDA Quick Screen; NIDA Rx = single question regarding prescription drug use for non-medical reasons, from NIDA Quick Screen; SURP-Substance Use Risk Profile-Pregnancy; CRAFFT = acronym for five-item screener with items related to car, relax, alone, forget, friends and trouble; WIDUS = Wayne Indirect Drug Use Screener; 5Ps = parents, peers, partner, pregnancy, past

Table 1

Screening instruments evaluated in the present study.

| Screening instrument | Original reference standard | Cut score (published) | Performance in original study (sensitivity/specificity) |
|---|----------------------------------|-----------------------|---|
| Substance Use Risk Profile–Pregnancy (SURP-P) [20] | Past 30-day use, per self-report | 2 | 0.57/0.88 |
| 10. Have you ever smoked marijuana? | | | |
| 20. In the month before you knew you were pregnant, how many beers, how much wine, or how much liquor did you drink? | | | |
| 30. Have you ever felt that you needed to cut down on your drug or alcohol use? | | | |
| Wayne Indirect Drug Use Screener (WIDUS) [5] | Urine and hair (dichotomous) | 3 | 0.68/0.69 ^a |
| 10. I am currently married | | | |
| 20. In the past year, I have been bothered by pain in my teeth or mouth | | | |
| 30. I have smoked at least 100 cigarettes in my entire life | | | |
| 40. Most of my friends smoke cigarettes | | | |
| 50. There have been times in my life, for at least 2 weeks straight, where I felt like everything was an effort | | | |
| 60. I get mad easily and feel a need to blow off some steam | | | |
| CRAFT questionnaire [21,22] | Calendar recall | 2 | NA ^b |
| 10. Have you ever ridden in a car driven by someone (including yourself) who was 'high' or had been using alcohol or drugs? | | | |
| 20. Do you ever use alcohol or drugs to relax, feel better about yourself, or fit in? | | | |
| 30. Do you ever use alcohol/drugs while you are by yourself, alone? | | | |
| 40. Do you ever forget things you did while using alcohol or drugs? | | | |
| 50. Do your family or friends ever tell you that you should cut down on your drinking or drug use? | | | |
| 60. Have you ever gotten into trouble while you were using alcohol or drugs? | | | |

5Ps questionnaire [23]

| | | | |
|--|-------------------------------|---|---|
| 10. Did any of your parents have a problem with using alcohol or drugs? | | | |
| 20. Do any of your friends have a problem with drug or alcohol use? | | | |
| 30. Does your partner have a problem with drug or alcohol use? | | | |
| 40. Before you knew you were pregnant, how often did you drink beer, wine, wine coolers or liquor? | | | |
| 50. In the past month, how often did you drink beer, wine, wine coolers or liquor? | Dependence (CID) ^c | 1 | 1 |
| NIDA Quick Screen [25–32] | | | |

| Screening instrument | Original reference standard | Cut score (published) | Performance in original study (sensitivity/specificity) |
|---|-----------------------------|-----------------------|---|
| 10. In the past year, ^d how often have you had four or more drinks a day? | | | 0.88/0.84 |
| 20. In the past year, ^d how often have you used tobacco products? | | | 0.97/0.79 |
| 3. In the past year, ^d how often have you used prescription drugs for non-medical reasons? | | | |
| 4. In the past year, ^d how often have you used illegal drugs? | | | |

^aFrom cross-validation sample.

^bSensitivity and specificity for CRAFFT (acronym for five-item screener with items related to car, relax, alone, forget, friends and trouble) with pregnant women not available; 5Ps = parents, peers, partner, pregnancy, past, positive predictive value (PPV) = 0.90, negative predictive value (NPV) = 0.80, area under the curve (AUC) = 0.90.

^cAccuracy statistics are available for the alcohol and drug questions of the NIDA Quick Screen but not the tobacco or prescription drug misuse items. CIDI = Composite International Diagnostic Interview.

^dThe NIDA Quick Screen uses a 'past-year' time-frame. Because women tend to reduce substance use during pregnancy, this study used a past-month time-frame.

Table 2

Participant demographic and drug use characteristics.

| Characteristic | Total sample (N = 1220) | New Haven (n = 554) | Detroit (n = 333) | Boston (n = 333) | P-value* |
|---|-------------------------|---------------------|-------------------|------------------|----------|
| Age, mean (SD) | 29.0 (5.9) | 28.7 (6.1) | 26.1 (5.1) | 32.6 (4.3) | <0.001 |
| Education, n (%) | | | | | <0.001 |
| Less than high school | 124 (10.3) | 64 (11.7) | 57 (17.5) | 3 (0.9) | |
| High school degree | 339 (28.2) | 207 (37.9) | 115 (35.3) | 17 (5.1) | |
| Some college | 320 (26.6) | 140 (25.6) | 130 (39.9) | 50 (15.1) | |
| College degree or higher | 421 (35.0) | 135 (24.7) | 24 (7.4) | 262 (78.9) | |
| Trimester, n (%) | | | | | <0.001 |
| First | 430 (35.3) | 183 (33.0) | 159 (47.8) | 88 (26.4) | |
| Second | 394 (32.3) | 187 (33.8) | 90 (27.0) | 117 (35.1) | |
| Third | 396 (32.5) | 184 (33.2) | 84 (25.2) | 128 (38.4) | |
| Public assistance, n (%) | 539 (44.7) | 261 (47.8) | 236 (72.0) | 42 (12.6) | <0.001 |
| Race/ethnicity, n (%) | | | | | <0.001 |
| Non-Hispanic black | 480 (40.1) | 174 (32.2) | 283 (86.8) | 23 (7.0) | |
| Non-Hispanic white | 444 (37.1) | 190 (35.2) | 16 (4.9) | 238 (71.9) | |
| Hispanic | 190 (15.9) | 141 (26.1) | 12 (3.7) | 37 (11.2) | |
| Non-Hispanic other/mixed race | 83 (6.9) | 35 (6.5) | 15 (4.6) | 33 (10.0) | |
| Substance positive ^a , n (%) | | | | | |
| Alcohol | 145 (11.9) | 38 (6.9) | 63 (18.9) | 44 (13.2) | <0.001 |
| Opioids ^b | 43 (3.5) | 18 (3.3) | 16 (4.8) | 9 (2.7) | 0.303 |
| Other illicit drugs ^c | 199 (16.3) | 87 (15.7) | 100 (30.0) | 12 (3.6) | <0.001 |
| Tobacco | 201 (16.5) | 98 (17.7) | 90 (27.0) | 13 (3.9) | <0.001 |
| Any substance | 397 (32.5) | 175 (31.6) | 160 (48.1) | 62 (18.6) | <0.001 |

* P-values are based upon analysis of variance (ANOVA) for continuous variables and χ^2 tests for categorical variables. SD = standard deviation.

^a Defined as positive if either the calendar recall interview (with a past 30-days window) or the urine drug screen was positive.

^b Including prescription opioid misuse or illicit opioid use.

^c Including marijuana, cocaine, heroin, amphetamines, barbiturates and hallucinogens.

Table 3

Screening instrument^d performance in identifying illicit drug use, alcohol use, or opioids.

| Illicit drugs | n (% positive) | | Sensitivity ^d (95% CI) | Specificity ^d (95% CI) | Accuracy ^d (95% CI) | PPV (95% CI) | NPV (95% CI) |
|-------------------------------|----------------|--------------------------------------|--------------------------------------|-----------------------------------|--------------------------------|--------------------------|--------------------------|
| | 199 (16.3) | AUC ^c (95% CI) | | | | | |
| SURP-P | 439 (36.4) | 0.58 ^x (0.54, 0.62) | 0.49 (0.42, 0.57) | 0.66 (0.63, 0.69) | 0.63 (0.61, 0.66) | 0.22 (0.18, 0.26) | 0.87 (0.85, 0.89) |
| WIDUS | 349 (69.8) | 0.70 (0.67, 0.74) | 0.63 (0.55, 0.69) | 0.78 ^x (0.75, 0.80) | 0.75 (0.72, 0.78) | 0.36 (0.31, 0.41) | 0.91 (0.89, 0.93) |
| CRAEFT | 300 (28.9) | 0.56 ^x (0.52, 0.59) | 0.34 ^x (0.27, 0.41) | 0.77 ^x (0.74, 0.80) | 0.70 (0.69, 0.73) | 0.22 (0.18, 0.27) | 0.86 (0.83, 0.88) |
| 5Ps | 799 (65.7) | 0.58 ^x (0.55, 0.61) | 0.80 (0.73, 0.85) | 0.37 (0.34, 0.40) | 0.44 (0.41, 0.47) | 0.20 (0.17, 0.22) | 0.90 (0.87, 0.93) |
| NIDA ^b | 62 (5.2) | 0.63 (0.60, 0.66) | 0.27 ^x (0.21, 0.34) | 0.99 (0.98, 1.00) | 0.87 (0.85, 0.89) | 0.84 (0.72, 0.92) | 0.88 (0.86, 0.90) |
| Alcohol use 145 (11.9) | | | | | | | |
| SURP-P | 439 (36.4) | 0.66 ^x (0.62, 0.70) | 0.65 (0.56, 0.73) | 0.68 (0.65, 0.70) | 0.67 ^x (0.64, 0.70) | 0.21 (0.18, 0.26) | 0.93 (0.91, 0.95) |
| WIDUS | 349 (69.8) | 0.54 ^y (0.50, 0.58) | 0.36 (0.28, 0.45) | 0.72 (0.69, 0.75) | 0.68 ^x (0.65, 0.70) | 0.15 (0.11, 0.19) | 0.89 (0.87, 0.91) |
| CRAEFT | 300 (28.9) | 0.67^x (0.62, 0.71) | 0.54 (0.46, 0.62) | 0.79 (0.77, 0.82) | 0.76 (0.74, 0.79) | 0.26 (0.21, 0.31) | 0.93 (0.91, 0.94) |
| 5Ps | 799 (65.7) | 0.62 ^x (0.59, 0.65) | 0.88 (0.81, 0.92) | 0.37 (0.34, 0.40) | 0.43 (0.40, 0.46) | 0.16 (0.13, 0.18) | 0.96 (0.93, 0.97) |
| NIDA ^b | 28 (2.3) | 0.54 ^y (0.52, 0.57) | 0.10 (0.05, 0.16) | 0.99 (0.98, 0.99) | 0.88 (0.86, 0.90) | 0.50 (0.31, 0.69) | 0.89 (0.87, 0.91) |
| Opioids 43 (3.5) | | | | | | | |
| SURP-P | 439 (36.4) | 0.54 ^x (0.46, 0.62) | 0.44 ^{x,y} (0.29, 0.60) | 0.64 (0.61, 0.67) | 0.63 (0.60, 0.66) | 0.04 (0.03, 0.07) | 0.97 (0.95, 0.98) |
| WIDUS | 349 (69.8) | 0.69 (0.62, 0.77) | 0.66 ^{x,z} (0.49, 0.80) | 0.72 ^x (0.70, 0.75) | 0.72 ^x (0.70, 0.75) | 0.08 (0.05, 0.11) | 0.98 (0.97, 0.99) |
| CRAEFT | 300 (28.9) | 0.56 ^x (0.49, 0.64) | 0.37 ^y (0.23, 0.53) | 0.76 ^x (0.73, 0.78) | 0.74 ^x (0.72, 0.77) | 0.05 (0.03, 0.09) | 0.97 (0.96, 0.98) |
| 5Ps | 799 (65.7) | 0.58 ^x (0.52, 0.64) | 0.81^z (0.67, 0.92) | 0.35 (0.32, 0.38) | 0.37 (0.34, 0.39) | 0.04 (0.03, 0.06) | 0.98 (0.96, 0.99) |
| NIDA ^b | 24 (2.0) | 0.57 ^x (0.52, 0.63) | 0.16 (0.07, 0.31) | 0.99 (0.98, 0.99) | 0.96 (0.94, 0.97) | 0.29 (0.13, 0.51) | 0.97 (0.96, 0.98) |

^a All measures were evaluated using their recommended published score for each separate outcome.

^b The single drug use question was used for identifying illicit drug use; the single alcohol use question was used for identifying alcohol use and the single question regarding prescription drug misuse was used for identifying any opioid use (whether prescription misuse or heroin).

^c *Post-hoc* pairwise comparisons of questionnaires for substance use outcomes from Z-score tests are shown with superscripts. Performance measures not statistically significantly different at $P < 0.05$ share a letter; those that are statistically significantly different do not share a letter.

^d *Post-hoc* pairwise comparisons of questionnaires for substance use outcomes from generalized estimating equation models, adjusted for race/ethnicity, site, trimester and public assistance, are shown with superscripts. Performance measures not statistically significantly different do not share a letter.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

SURP-P = Substance Use Risk Profile–Pregnancy; WIDUS = Wayne Indirect Drug Use Screener; CRAFFT = acronym for five-item screener with items related to car, relax, alone, forget, friends and trouble; 5Ps parents, peers, partner, pregnancy, past; NIDA = National Institute on Drug Abuse; AUC = area under the receiver operating characteristic curve; PPV = positive predictive value; NPV = negative predictive value; CI = confidence interval; Accuracy = proportion of correct classifications. Values in bold type represent the top score for that performance measure within each substance outcome (but are not necessarily significantly different from the other scores).

Table 4
Screening instrument performance in identifying illicit drug use by race, site, trimester and receipt of public assistance.

| | SURP-P | | | WIDUS | | | CRAPPT | | | 5Ps | | | NIDA Quick Screen | | | Interaction significance (P) | | |
|-----------------------------|--------|-------|-------|-------|-------|-------|--------|-------|-------|------|-------|-------|-------------------|-------|-------|------------------------------|-------|--------|
| | Acc. | Sens. | Spec. | Acc. | Sens. | Spec. | Acc. | Sens. | Spec. | Acc. | Sens. | Spec. | Acc. | Sens. | Spec. | Acc. | Sens. | Spec. |
| Race/ethnicity | | | | | | | | | | | | | | | | | | |
| Black (<i>n</i> = 480) | 0.67 | 0.51 | 0.73 | 0.68 | 0.63 | 0.70 | 0.72 | 0.34 | 0.86 | 0.52 | 0.80 | 0.42 | 0.81 | 0.31 | 0.99 | < 0.01 | 0.57 | < 0.01 |
| White (<i>n</i> = 444) | 0.55 | 0.61 | 0.55 | 0.86 | 0.76 | 0.87 | 0.62 | 0.45 | 0.64 | 0.31 | 0.91 | 0.26 | 0.94 | 0.22 | 0.99 | | | |
| Hispanic, (<i>n</i> = 190) | 0.67 | 0.29 | 0.74 | 0.63 | 0.39 | 0.68 | 0.78 | 0.21 | 0.88 | 0.46 | 0.64 | 0.43 | 0.87 | 0.19 | 0.98 | | | |
| Site | | | | | | | | | | | | | | | | | | |
| New Haven (<i>n</i> = 554) | 0.67 | 0.47 | 0.71 | 0.70 | 0.63 | 0.71 | 0.75 | 0.30 | 0.84 | 0.49 | 0.75 | 0.44 | 0.87 | 0.20 | 0.99 | < 0.01 | 0.38 | < 0.01 |
| Detroit (<i>n</i> = 333) | 0.65 | 0.51 | 0.72 | 0.67 | 0.61 | 0.70 | 0.69 | 0.35 | 0.84 | 0.50 | 0.82 | 0.37 | 0.79 | 0.34 | 0.99 | | | |
| Boston (<i>n</i> = 333) | 0.55 | 0.58 | 0.55 | 0.92 | 0.75 | 0.93 | 0.62 | 0.58 | 0.62 | 0.29 | 0.92 | 0.26 | 0.97 | 0.17 | 10.00 | | | |
| Trimester | | | | | | | | | | | | | | | | | | |
| First (<i>n</i> = 430) | 0.65 | 0.44 | 0.70 | 0.78 | 0.59 | 0.83 | 0.68 | 0.27 | 0.78 | 0.45 | 0.70 | 0.39 | 0.86 | 0.29 | 10.00 | 0.13 | 0.03 | NA |
| Second (<i>n</i> = 394) | 0.63 | 0.51 | 0.66 | 0.74 | 0.66 | 0.76 | 0.69 | 0.41 | 0.76 | 0.44 | 0.87 | 0.34 | 0.87 | 0.34 | 0.99 | | | |
| Third (<i>n</i> = 396) | 0.62 | 0.58 | 0.62 | 0.73 | 0.64 | 0.74 | 0.73 | 0.37 | 0.77 | 0.42 | 0.87 | 0.38 | 0.89 | 0.11 | 0.98 | | | |
| Public assistance | | | | | | | | | | | | | | | | | | |
| Yes (<i>n</i> = 539) | 0.69 | 0.52 | 0.75 | 0.63 | 0.65 | 0.62 | 0.73 | 0.34 | 0.87 | 0.53 | 0.79 | 0.44 | 0.80 | 0.27 | 0.98 | < 0.01 | 0.60 | < 0.01 |
| No (<i>n</i> = 668) | 0.59 | 0.45 | 0.60 | 0.86 | 0.53 | 0.88 | 0.67 | 0.35 | 0.70 | 0.36 | 0.80 | 0.32 | 0.94 | 0.28 | 1.00 | | | |

SURP-P = Substance Use Risk Profile–Pregnancy; WIDUS = Wayne Indirect Drug Use Screener; CRAPPT = acronym for five-item screener with items related to car, relax, alone, forget, friends and trouble; 5Ps = parents, peers, partner, pregnancy, past; Acc. = accuracy, defined as overall rate of agreement (true positives + true negatives/full sample); sens. = sensitivity; spec. = specificity; NA = not applicable. Accuracy, sensitivity and specificity were calculated using published cut scores.