Objective Measures of Prenatal Alcohol Exposure: A Systematic Review

Cheryl McQuire[a], MSc, Shantini Paranjothy[a], PhD, Lisa Hurt[a], PhD; Mala Mann[b], MIInfSc, MCLIP, Daniel Farewell[a], PhD, Alison Kemp[a], FRCPCH

Affiliations: [a] Cardiff University, Division of Population Medicine, [b] Cardiff University, Specialist Unit for Review Evidence

Address correspondence to: Cheryl McQuire, Division of Population Medicine, Cardiff University, 3rd Floor, Neuadd Meirionnydd, Heath Park, Cardiff, CF14 4YS
Tel: +44(0)29 20687169, Email: McQuireC1@cardiff.ac.uk

Short title: Objective Measures of Prenatal Alcohol Exposure

Funding source: CM is a PhD candidate, funded by Cardiff University, School of Medicine. The funders had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Financial disclosure: The authors have no financial relationships relevant to this article to disclose.

Conflict of interest: The authors have no conflicts of interest relevant to this article to disclose.

Protocol registration: PROSPERO international prospective register of systematic reviews (www.crd.york.ac.uk; record number CRD42014015420)

Abbreviations:
CDT - carbohydrate deficient transferrin
EtG - ethyl glucuronide
EtS - ethyl sulphate
FAEE - fatty acid ethyl ester
GGT - gamma glutamyltransferase
Hb-Ach - haemoglobin acetaldehyde adducts
MCV - mean corpuscular volume
PAE - prenatal alcohol exposure
PEth - phosphatidylethanol

Key words: systematic review, prenatal alcohol use, objective test, biomarker, validity, diagnostic accuracy, sensitivity, specificity, predictive value

Contributors' Statement

Cheryl McQuire developed the draft protocol and search strategy, data extraction sheet and quality coding criteria, carried out study selection, data extraction, quality assessment, evidence synthesis and prepared the initial manuscript and subsequent revisions.

Professor Shantini Paranjothy developed the initial concept for the study, contributed to development of the protocol, extracted data, verified study inclusion decisions and quality assessments and reviewed and revised the manuscript.

Dr Lisa Hurt contributed to development of the protocol, extracted data, verified study inclusion decisions and quality assessments and reviewed and revised the manuscript.

Mala Mann contributed to the design of the search strategy and reviewed and revised the manuscript.

Dr Daniel Farewell provided statistical advice and reviewed and revised the manuscript.

Professor Alison Kemp contributed to development of the protocol and reviewed and revised the manuscript.

All authors approved the final manuscript as submitted.

Context: Objective measurement of prenatal alcohol exposure (PAE) is essential for identifying children at risk of adverse outcomes, including fetal alcohol spectrum disorders. Biomarkers have been advocated for use in universal screening programs but their validity has not been comprehensively evaluated.

Objective: To systematically review the validity of objective measures of PAE.

Data Sources: Thirteen electronic databases and supplementary sources were searched for studies published between January 1990 and October 2015.

Study Selection: Eligible studies were those that evaluated the diagnostic accuracy of objective measures of PAE.

Data Extraction: Three reviewers independently verified study inclusion, quality assessments and extracted data.

Results: Twelve studies met inclusion criteria. Test performance varied widely across studies of maternal blood (four studies; sensitivity 0% to 100%, specificity 79% to 100%), maternal hair (two studies; sensitivity 19% to 87%, specificity 56% to 86%) maternal urine (two studies; sensitivity 5% to 15%, specificity 97% to 100%), and biomarker test batteries (three studies; sensitivity 22% to 50%, specificity 56% to 97%). Tests of the total concentration of four fatty acid ethyl esters (in meconium - two studies, or placenta - one study) demonstrated high sensitivity (82% to 100%); however, specificity was variable (13% to 98%).

Limitations: Risk of bias was high due to self-report reference standards and selective outcome reporting.

Conclusions: Current evidence is insufficient to support the use of objective measures of prenatal alcohol exposure in practice. Biomarkers in meconium and placenta tissue may be the most promising candidates for further large-scale population-based research.

Up to 80% of women consume some level of alcohol while pregnant.[1,2] Most drink at low levels and reduce their intake throughout pregnancy. However, up to 45% may binge drink in the first trimester and up to 6% continue to drink heavily.[1] Prenatal alcohol exposure (PAE) is associated with a range of adverse perinatal and long-term outcomes including spontaneous abortion, preterm delivery and cognitive impairment.[3-8]

Heavy fetal alcohol exposure is the most likely to lead to detrimental outcomes, including fetal alcohol spectrum disorders. The estimated prevalence of FASD is 3% to 5% within the general population in North America and Europe, and up to 26% in South Africa, making it one of the leading preventable causes of developmental disability worldwide.[9,10]

Reviews of the effects of low to moderate PAE on short- and long-term developmental outcomes are inconclusive,[11-13] and debate continues as to whether it is possible to identify a safe threshold for drinking in pregnancy.[14,15] Observed effects of low to moderate PAE on childhood cognitive and behavioural outcomes range from evidence of harm,[16-18] to null findings,[19-22] to evidence of benefit.[23-25] Studies of the effects of low to moderate PAE on birth outcomes and child growth trajectories are also inconsistent.[2,26-31] Discrepancies in findings are likely due to measurement error and residual confounding owing to the socioeconomic patterning of prenatal alcohol use.[16,32,33]

Inconsistencies in the evidence base are reflected in international guidelines for drinking in pregnancy. Most countries in North America, Europe and Australasia endorse a clear abstinence message.[34] Current UK guidelines are controversial. The National Institute for Health and Care Excellence (NICE) recommend abstinence in the first trimester followed by no more than one to two units once or twice a week.[35] However, in January 2016 the UK Chief Medical Officer issued new guidance stating that women should avoid alcohol throughout pregnancy.[36]

Self-report measures are the most common method for assessing PAE and include survey methods and standardised questionnaires.[37-39] Although widely used, self-report measures are likely to underestimate true levels of alcohol consumption for reasons including social stigma and difficulties in recalling drinking behaviour.[40-42] Objective measures are needed to help researchers ascertain the true prevalence of PAE, and to better understand the effects of low to moderate exposure. For clinicians, objective measures of PAE could support FASD diagnosis and guide efforts to prevent alcohol-related harm.[43-48]

Biomarkers have received increasing attention as objective measures of PAE.[49-54] Alcohol biomarkers can be found in a range of biological matrices including maternal blood, sweat, urine, oral fluid and hair; newborn blood, urine, hair and meconium; and in maternal-fetal matrices such as the placenta.[51] Direct biomarkers include ethanol itself or the products of ethanol metabolism. Indirect markers are those that signal alcohol-induced pathology, following chronic alcohol use.[52]

Some groups have advocated the use of biomarkers of PAE within universal screening programs.[55] For example, meconium testing features as a recommended method within the Canadian FASD National Screening Tool Kit.[56] However, the evidence for the diagnostic accuracy of biomarkers has not been comprehensively evaluated. In this context, we carried out a systematic review of the diagnostic accuracy of objective measures of PAE.

Methods

We followed the Cochrane Collaboration guidelines for systematic reviews of diagnostic test accuracy,[57] the Standards for the Reporting of Diagnostic Accuracy Studies (STARD),[58] and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (see Supplementary File 1 for PRISMA checklist).[59] The full protocol is available

from the PROSPERO international prospective register of systematic reviews (www.crd.york.ac.uk; record number CRD42014015420).

Search strategy

Figure 1 presents an overview of the search strategy including sources of literature. We searched 13 electronic databases, including sources of grey literature, from January 1990 to August 2015 for original articles using combinations of terms related to *objective measures* and *diagnostic accuracy* and *prenatal alcohol exposure*. Searches were limited to publications from January 1990 onwards to increase precision. A scoping exercise of existing non-systematic reviews revealed no relevant articles prior to 1990.[50,52,60-62] Full details of the search strategy for Medline is available in Supplementary File 2. This search string was translated for use in all other databases. Supplementary sources were searched for articles published between November 2012 and October 2015. We contacted authors to request further information about missing or conflicting data.

Inclusion/exclusion criteria

Eligible studies were randomised screening studies and diagnostic cross-sectional, cohort, and case-control studies of pregnant and/or postpartum women and/or neonates that investigated the performance of any objective measure of PAE in comparison to any reference standard.

We excluded conference abstracts, studies with missing outcome data, and animal studies. Non-English language publications were excluded due to a lack of funding for translation costs.

Study selection

After removing duplicates, CM screened the search records against pre-determined inclusion criteria and excluded ineligible studies based on the title or abstract. Full text versions were obtained to determine the inclusion of potentially relevant studies. A random selection of 10% of these studies were independently assessed for eligibility by two other reviewers (LH and SP). The level of agreement for inclusion decisions was 100%.

Data extraction

All data were extracted by CM into a standardised electronic form, designed based on the Guidelines International Network (GIN) template for diagnostic studies and STARD.[58,63] SP or LH independently repeated data extraction to ensure accuracy. We extracted information about study design, participant, index and reference test characteristics, and diagnostic accuracy outcomes. Alcohol data were classified according to the United States National Institute on Alcohol Abuse and Alcoholism criteria for the general population in which one standard drink is equivalent to 0.6 oz. or 14 g of ethanol and light drinking is equivalent to < 3 drinks per week, moderate drinking 3 to 7 drinks per week, and heavy drinking > 7 drinks per week or a binge pattern of ≥ 4 drinks per occasion.[64]

Quality assessment

We used a modified version of the Quality Assessment of Diagnostic Accuracy Studies tool (QUADAS-2) to assess the methodological quality of included studies.[65] QUADAS-2 was tailored to address specific areas of relevance for this review, based on guidance from the Cochrane Collaboration (see Supplementary File 3 for quality coding criteria).[66,67] Methodological quality was independently assessed by two members of the review team (CM and SP or LH) and referred to a third member of the team to resolve any disagreements in risk of bias decisions.

Analysis and data synthesis

A minimum of four studies per test are required for meta-analyses of diagnostic accuracy data within STATA using the recommended Rutter and Gatsonis HSROC model.[68,69] Due to the diverse nature of the data, which represented a variety of measures and assay methods across a range of matrices, none of the test categories had a sufficient number of studies to facilitate meta-analysis. Therefore, we conducted a narrative synthesis of the data. Sensitivity, specificity, predictive values, and likelihood ratios were the key diagnostic accuracy outcomes.

Diagnostic accuracy data were entered into Review Manager (RevMan).[70] Where necessary, the RevMan calculator to was used to derive diagnostic summary statistics from true positive, true negative, false positive and false negative values and vice-versa. If there were insufficient data to enable the use of the RevMan calculator, we used R software[71] to conduct an exhaustive search of all possible 2x2 tables that were consistent with the data supplied in the primary study. Confidence intervals were generated automatically using the exact binomial method[72] in RevMan for sensitivity and specificity values and with the MedCalc online calculator[73] for predictive values. Confidence intervals for likelihood ratios were generated using the method described by Koopman.[74]

Many studies reported a range of diagnostic accuracy values according to characteristics such as positivity cut-off and period of measurement. In order to aid presentation of findings, we report only the highest values of both sensitivity and specificity per study within the summary of results.

Ethics committee approval

Ethical approval was not required as the study used secondary data.

Results

Characteristics of included studies

From 4278 search records, 12 studies, including 1614 unique participants, met eligibility criteria.[75-86] Figure 1 depicts the study selection process and Table 1 presents characteristics of included studies.

Studies included data on participants from the USA (5),[75,77,82,84,85] Korea (2),[80,81] Spain (1),[86] South Africa (1),[76] and Finland (1)[83], and combined data from the USA and Jordan (1),[78] and Canada and Israel (1).[79] Diagnostic accuracy data were available for eight types of biomarker: carbohydrate deficient transferrin (CDT),[75,83,85] ethyl sulphate (EtS),[75,85] ethyl glucuronide (EtG),[75,85,86] fatty acid ethyl esters (FAEEs),[75-80,82,84] gamma glutamyltransferase (GGT),[75,83,85] haemoglobin acetaldehyde adducts (Hb-Ach),[83] mean corpuscular volume (MCV)[83] and phosphatidylethanol (PEth),[75,81,85] within six matrices: meconium,[75-80,82] placenta,[84] maternal urine,[75,85] maternal blood,[75,81,83,85] maternal hair[85,86] and infant blood.[75] Eight studies investigated tests of moderate to heavy prenatal alcohol consumption[75,76,78,79,81,83-85] and nine recruited women from high-risk settings, such as substance misuse clinics.[75-79,82-85] Eight of the eligible studies recruited pregnant women who reported abstinence from alcohol as a comparison group.[75,77-82,85] Two of these studies included an additional control group of pregnant women from cultures that promote abstinence from alcohol.[78,79] Three of the studies explored test performance for distinguishing between heavy drinkers and women with lower levels of prenatal alcohol consumption[76,81,83,84] and one study did not report characteristics of the control group.[86]

Methodological quality of included studies

Results of the QUADAS-2 quality assessment are presented in Figures 2 and 3.

All but one of the studies used a self-report reference standard and had a high risk of bias for the reference standard domain.[75-85] Self-report is known to be an imperfect reference standard for reasons previously discussed. The remaining study[86] used meconium EtG as the reference standard. This study was considered to have an unclear risk of bias as the validity of meconium EtG has not been established and there is a lack of agreement about the optimal positivity threshold for this biomarker.[87-89] Seven studies had a high risk of bias in the participant selection domain due to the use of diagnostic case-control designs,[75-80,85] which may inflate accuracy estimates.[90-92] Nine studies had a low risk of bias for uninterpretable results,[75,76,78-83,86] withdrawals,[75,76,78-81,83,85,86] differential verification,[75-77,80,81,83-86] and partial verification.[75,76,79-81,83-86] The risk of incorporation bias was high in one study as EtG was used to indicate alcohol exposure in both the reference standard and index test.[86]

Seven studies had a low risk of bias for the detection window domain.[75,77-80,83,86] Of these studies, two reported data for multiple index tests both with and without an appropriate window of detection.[75,78] In order to reduce the risk of bias, data were excluded from one study that looked at the agreement between self-reported drinking in the first trimester and meconium testing,[78] as meconium does not begin to accumulate until the second and third trimesters.[60] We also excluded data from another study that compared self-reported PAE during the second trimester with postnatal tests of maternal EtG, EtS, GGT, CDT and PEth in dried infant blood spots due to the short detection window of these biomarkers.[75] Of the remaining studies, two had an unclear risk of bias and three had a high risk of bias. Two of these studies[76,82] were deemed to have a high risk of bias as the accuracy of meconium testing was verified against alcohol use across the whole of pregnancy, including the first trimester before meconium is generated. It was not possible to exclude first trimester data from our analysis due to the way findings were reported. One study,[85] which collected maternal hair during pregnancy for EtG testing, was also considered to have a high risk of bias as the

specimen may have captured alcohol use prior to pregnancy due to the broad detection window of EtG within this matrix.

Finally, selective outcome reporting introduced a high risk of bias in four studies.[76-79] Of these studies, three measured multiple FAEEs in meconium but only reported diagnostic accuracy outcomes for a subset of these FAEEs.[76-78] Two studies[78,84] did not provide sufficient data to enable the use of the RevMan calculator to derive missing true positive, true negative, false positive and false negative values. Using the R simulation, we were able to produce data that replicated the sensitivity, sensitivity and predictive values reported in one of the studies,[84] and generated values that approximated the published data in another study.[78] For the remaining study of meconium FAEEs,[79] the positive predictive value reported in the study did not match the value suggested by the raw data (see Table 2 for further details). The study authors were unable to provide data to further explore this discrepancy. Finally, we were unable to replicate the published sensitivity and specificity values based on the true positive, true negative, false positive, and false negative values presented in one study of maternal hair testing.[86] Following correspondence with the authors the correct sensitivity and specificity values were derived based on the raw values presented in the paper (see Table 5).

Diagnostic accuracy findings

Tables 2 - 7 present diagnostic accuracy outcomes with 95% confidence intervals.

Meconium testing

Meconium testing for FAEEs was the most commonly investigated index test and featured in seven studies.[75-80,82] The diagnostic accuracy of FAEEs varied widely across studies (see Table 2). A measure of the total concentration of four FAEEs demonstrated the highest levels

of diagnostic accuracy overall, but there were a high number of false positives in one study[75] and specificity was inconsistent.[75,79]

Placenta testing

FAEEs were measured in placenta tissue in one study of premature deliveries.[84] Sensitivity and specificity values were high, although 30% to 56% of positive test results were false positives (see Table 3).

Blood testing

Four studies investigated CDT, GGT, Hb-Ach, MCV and PEth[75,81,83,85] within prenatal samples of maternal blood. Blood biomarkers generally demonstrated low sensitivity, although specificity was high (see Table 4). Likelihood ratios suggested that the accuracy of PEth testing improved as PAE increased from low to moderate to heavy. However, findings of high levels of sensitivity (100%) and specificity (96%) in one study of heavy PAE[81] were not supported by two other studies in which sensitivity was 18% to 22%.[75,85]

Hair testing

Maternal hair was tested for EtG in two studies[85,86] with contrasting findings (see Table 5). Neither of the studies demonstrated high levels of both sensitivity and specificity.

Urine testing

EtS and EtG in maternal urine had low sensitivity but high specificity in two studies[75,85] (see Table 6).

Test batteries

Three studies investigated the diagnostic accuracy of test batteries including combinations of different biomarkers across multiple matrices.[75,83,85] Sensitivity was poor, while specificity was generally good (see Table 7).

Discussion

This systematic review demonstrates that the accuracy of biomarkers of PAE varies widely across studies. Tests of the total concentration of four FAEEs demonstrated the highest levels of sensitivity across studies. Sensitivity was 100% in two studies of meconium testing,[75,79] and 82% in one study of placenta testing.[84] However, specificity was variable (13% to 98%). Positive likelihood ratios ranged from 1, suggesting little test utility, to 73. As a guide, positive likelihood ratios greater than 10 may indicate that a test is informative.[93] Due to a small number of cases in these studies confidence intervals were wide, leading to imprecise estimates of diagnostic accuracy. Placenta testing was conducted with a sample of premature newborns.[84] It is important to note that because alcohol is a risk factor for prematurity,[94] the prevalence of PAE is likely to be higher within this sample than in the general population. Accordingly, positive predictive values from this study are likely to be higher than what would be expected within routine antenatal care.[95]

There is no consensus as to what level of diagnostic accuracy is acceptable for objective measures of PAE. Many screening initiatives prioritise sensitivity over specificity and thus permit a high number of false positive results in order to maximise early detection of asymptomatic conditions.[96] As early diagnosis and intervention are associated with improved outcomes for children with FASD[43] it could be argued that a test with high sensitivity could be favoured over specificity in this context, as it may facilitate appropriate monitoring and follow-up among children with positive PAE screens. However, PAE is a highly emotive topic and false positive errors may lead to stigmatisation, unnecessary burden on healthcare

resources, and may even be used in legal proceedings against mothers.[55,96-99] Conversely, false negative errors represent a missed opportunity to provide support to families affected by PAE.[55] Many screening programmes are conducted in a tiered fashion with high sensitivity favoured over specificity in the initial phase. Second-tier screening using methods such as detailed maternal interviews and behavioural assessment of children with suspected PAE have been suggested as a strategy to reduce false positive results.[55] However, these methods also have limitations. Information from maternal interview may be inaccurate, and the cognitive-behavioural profile associated with PAE is heterogeneous and may not be detectable until later in childhood, thus precluding the opportunity for early intervention. Given the implications of both types of diagnostic error, various authors have suggested that both high sensitivity and specificity are a prerequisite for the introduction of PAE screening.[55,96,99]

Limitations of the evidence

The methodological quality of studies included in this review was generally poor. Risk of bias was high due to the use of imperfect self-report reference standards, case-control diagnostic designs, and selective outcome reporting. Therefore, it is unclear whether the low diagnostic accuracy values of biomarkers explored in this review result from a true lack of test validity, or are simply an artefact of comparison with an imperfect reference standard. As self-report reference standards are known to underestimate true PAE it is possible that the biomarker index tests explored in this review correctly detected true PAE while the reference standard did not. This lack of agreement in classification would result in false positives and reduced specificity. However, it is also possible that the observed false positives were genuine. Incidental exposure to ethanol can occur through an individual's diet, medications, mouthwash and hand sanitizer, although the extent to which incidental exposure influences

biomarker tests for PAE has not been fully established.[100-104] It is unlikely that false negative results are due to the use of an imperfect reference standard as mothers are unlikely to report that they drank alcohol while pregnancy when they had not.[105] Therefore, the low sensitivity values demonstrated by many studies in this review may be considered the most persuasive evidence against the validity of current objective measures of PAE. Some authors have proposed that apparent false negative errors may occur because self-report measures are better able to detect low levels of alcohol use than many of the objective measures, which typically detect moderate to heavy consumption. This raises the possibility that pregnant women who report drinking modest amounts of alcohol could be detected by the self-report reference standard but not the index test.[106] This explanation is not likely to account for findings in the present review, however, as the majority of studies investigated moderate to heavy self-reported PAE.

Case-control diagnostic designs may produce overestimates of diagnostic accuracy.[90,92] However, this form of bias is not likely to influence the conclusions of this review as most studies did not show high levels of accuracy despite the use of case-control diagnostic designs. It is, however, important to note that PAE prevalence is fixed by design in many of the studies included in this review due to the use of case-control methods and it is not possible to generalise predictive values from individual studies to settings with a different prevalence of PAE.[93,95,107]

Finally, selective outcome reporting was common among studies and introduced a further risk of bias. As recommended by the STARD guidelines,[58] we would urge authors to report reference and index test results in the form of absolute values to enable independent verification of findings.

Limitations of the review

This study is the first systematic review of its kind and provides a rigorous evaluation of the evidence relating to the diagnostic accuracy of a range of objective measures of PAE. Our findings are broadly consistent with findings from existing non-systematic reviews, which suggest the need for improved objective measures of PAE.[49,52,60] However, due to the diverse nature of the data and a limited number of studies per index test it was not possible to address some of the objectives listed in our original protocol. For example, we were not able to conduct the intended meta-analyses to answer questions about the relative impact of study characteristics on test accuracy. Our search strategy was comprehensive and covered a range of published and unpublished sources. It is possible, however, that some studies were missed as a result of excluding non-English language publications.

As previously noted, participants were mainly recruited from high risk settings, such as substance misuse clinics. Therefore, findings have limited applicability to general population samples. Population-based studies of biomarkers of PAE are needed to inform universal screening strategies and to clarify the epidemiology of PAE and its developmental consequences in the short- and long-term.

Implications

Prenatal alcohol use is a challenging and emotive issue. Consequently, tests to detect prenatal alcohol exposure must be accurate, feasible, and acceptable to the population. These criteria are emphasised in the World Health Organisation and UK National Screening Committee guidelines for screening procedures.[108,109] Our review demonstrates that the evidence base for the accuracy of current objective measures of PAE is not yet robust enough to support their use in routine care. More research is required to establish which method is most feasible and acceptable to stakeholders including clinicians, policy makers and families.[96,97,110,111]

Assay methods for the biomarkers included in this review were highly variable. This is likely to account for some of the observed heterogeneity in findings. Future work that aims to standardise procedures may provide a clearer picture of the performance of different biomarkers for PAE. Furthermore, positivity thresholds must be validated with general population samples. The 600 ng/g (2 nmol/g) cut-off for total concentration of four FAEEs was derived from a study comparing abstainers to women with alcoholism and, therefore, may not be suitable for determining PAE in the general population.[79]

With the exception of hair testing, objective measures of PAE have a limited detection period, which does not span the whole of pregnancy (see Supplementary File 3). For many women, patterns of alcohol consumption change throughout pregnancy. Women are most likely to drink in the first trimester and then reduce their intake or abstain in later trimesters.[1] Risk of harm to the developing fetus is highest if a mother drinks heavily throughout pregnancy,[112] however first trimester exposure poses a particular risk of physical abnormalities including dysmorphic facial features.[113] Meconium testing only captures PAE late in the second and third trimesters of pregnancy and, therefore, may fail to identify a large proportion of babies at risk of alcohol-related harm.[60] In addition, currently available biomarkers have insufficient sensitivity to detect low levels of PAE, which is the most prevalent pattern of consumption among pregnant women.[1] An objective test which measures alcohol itself in breath, urine, or blood could detect low level use. However, because alcohol is only present in these matrices for a matter of hours, this form of test is difficult to implement in research or in practice.[49,52] Due to the limitations of current biomarkers, authors have emphasised the need for novel biomarkers that can detect even low levels of alcohol use across the duration of pregnancy.[49]

Given the absence of a gold standard test, research attempting to validate objective measures of PAE may benefit from abandoning the classic diagnostic accuracy paradigm, in which validity is determined by agreement between the index test and reference standard. Instead, future research may benefit from using a clinical validation approach,[114] in which a convergent body of evidence is used to increase confidence in the validity of a measure. Some studies have adopted this method to demonstrate the predictive validity of meconium testing. Prospective studies have reported a significant inverse relationship between levels of FAEEs in meconium and cognitive outcomes up to age 15.[115,116] Such evidence lends support to the validity of FAEEs as markers of PAE, but require replication. Animal models may also be useful for the development of novel testing procedures for PAE.[49,117,118] However, translation from animal studies to human populations is complicated by differences in alcohol exposure methods, gestation and alcohol metabolism. In summary, validation will rely on an ongoing body of research that produces convergent evidence to suggest that objective measures are meaningfully associated with PAE.[114]

Conclusions

Tests of the total concentration of FAEEs in meconium and placental tissue offer some promise as objective measures of PAE but findings are inconsistent, studies are small-scale and require replication. Therefore, we conclude that current evidence is insufficient to support the use of objective measures of PAE in practice. The poor performance of many of the measures evaluated in this review could be due to a true lack of diagnostic validity or a result of bias introduced by sub-optimal study design, most notably the absence of a gold standard for PAE. Further research that investigates test validity, acceptability, and feasibility within a large population-based sample is required to inform strategies for population-based screening and epidemiological research.

Acknowledgments: We would like to thank Dr Helena Kemp, Consultant Chemical Pathologist, for her feedback on previous versions of this manuscript. We would also like to thank the authors of the primary studies included in this review for their helpful responses to our queries during the review process, and the reviewers for their valuable comments.

Table/Figure Legends:

Table 1. Characteristics of included studies of objective measures of prenatal alcohol exposure

Table 2. Diagnostic accuracy outcomes for studies of meconium testing for prenatal alcohol exposure

Table 3. Diagnostic accuracy outcomes for one study of placenta testing for prenatal alcohol exposure

Table 4. Diagnostic accuracy outcomes for studies of blood testing for prenatal alcohol exposure

Table 5. Diagnostic accuracy outcomes for studies of hair testing for prenatal alcohol exposure

Table 6. Diagnostic accuracy outcomes for studies of maternal urine testing for prenatal alcohol exposure

Table 7. Diagnostic accuracy outcomes for studies of objective test batteries of maternal biomarkers of prenatal alcohol exposure

Figure 1. Flow diagram depicting study selection process

Figure 2. Methodological quality summary for each included study. Ratings indicate risk of bias for each quality domain.

Figure 3: Methodological quality ratings for each domain represented as percentages across all included studies.

Supplementary File 1. PRISMA Checklist

Supplementary File 2. Search string for Medline

Supplementary File 3. Coding criteria for methodological quality assessment

Supplementary File 4. Reviewer comments table (1^{st} revision)

Supplementary File 5. Reviewer comments table (2^{nd} revision)

REFERENCES

1. O'Keeffe LM, Kearney PM, McCarthy FP, et al. Prevalence and predictors of alcohol use during pregnancy: findings from international multicentre cohort studies. *BMJ* Open. 2015;5:e006323. doi:006310.001136/bmjopen-002014-006323.
2. Nykjaer C, Alwan NA, Greenwood DC, et al. Maternal alcohol intake prior to and during pregnancy and risk of adverse birth outcomes: evidence from a British cohort. *J Epidemiol and Commun Health*. 2014;68:542–549.
3. Abel EL. Consumption of alcohol during pregnancy: a review of effects on growth and development of offspring. *Hum Biol*. 1982;54:421—453.
4. Harlap S, Shiono P. Alcohol, smoking, and incidence of spontaneous abortions in the first and second trimester. *Lancet*. 1980;316:173—176.
5. Mattson SN, Riley EP. A review of the neurobehavioral deficits in children with fetal alcohol syndrome or prenatal exposure to alcohol. *Alcohol Clin Exp Res*. 1998;22:279—294.
6. Sokol RJ, Janisse JJ, Louis JM, et al. Extreme prematurity: an alcohol-related birth effect. *Alcohol Clin Exp Res*. 2007;31:1031—1037.
7. Sayal K, Heron J, Golding J, et al. Binge pattern of alcohol consumption during pregnancy and childhood mental health outcomes: longitudinal population-based study. *Pediatrics*. 2009;123:e289—e296.
8. O'Leary CM, Taylor C, Zubrick SR, Kurinczuk JJ, Bower C. Prenatal alcohol exposure and educational achievement in children aged 8–9 years. *Pediatrics.* 2013;132:e468—e475.
9. Roozen S, Peters G-JY, Kok G, Townend D, Nijhuis J, Curfs L. Worldwide prevalence of fetal alcohol spectrum disorders: a systematic literature review including meta-analysis. *Alcohol Clin Exp Res.* 2016;40:18—32.
10. May PA, de Vries MM, Marais A-S, et al. The continuum of fetal alcohol spectrum disorders in four rural communities in south africa: prevalence and characteristics. *Drug Alcohol Depend 2016;* doi: 10.1016/j.drugalcdep.2015.12.023.
11. Henderson J, Gray R, Brocklehurst P. Systematic review of effects of low–moderate prenatal alcohol exposure on pregnancy outcome. *Brit J Obstet Gynaec.* 2007;114:243—252.
12. O'Leary CM, Bower, C. Guidelines for pregnancy: what's an acceptable risk, and how is the evidence (finally) shaping up? *Drug Alcohol Rev.* 2012;31:170—183.
13. Flak AL, Su S, Bertrand J, Denny CH, Kesmodel US, Cogswell ME. The association of mild, moderate, and binge prenatal alcohol exposure and child neuropsychological outcomes: a meta-analysis. *Alcohol Clin Exp Res.* 2014;38:214—226.
14. Mukherjee RAS, Hollins S, Abou-Saleh MT, Turk J. Low level alcohol consumption and the fetus. *BMJ.* 2005;330:375—376.
15. Mather M, Wiles K, O'Brien P. Should women abstain from alcohol throughout pregnancy? *BMJ.* 2015;351: h5232.
16. Scholder S, Wehby GL, Lewis S, Zuccolo L. Alcohol exposure in utero and child academic achievement. *Econ J.* 2014;124:634—667.
17. Lewis SJ, Zuccolo L, Davey Smith G, et al. Fetal alcohol exposure and IQ at age 8: evidence from a population-based birth-cohort study. *PLoS One.* 2012;7:e49407.
18. Sood B, Delaney-Black V, Covington C, et al. Prenatal alcohol exposure and childhood behavior at age 6 to 7 years: I. dose-response effect. *Pediatrics.* 2001;108:e34—e34.

19. Falgreen Eriksen HL, Mortensen EL, Kilburn T, et al. The effects of low to moderate prenatal alcohol exposure in early pregnancy on IQ in 5-year-old children. Brit J Obstet Gynaec. 2012;119:1191—1200.
20. Kesmodel US, Bertrand J, Støvring H, et al. The effect of different alcohol drinking patterns in early to mid pregnancy on the child's intelligence, attention, and executive function. *Brit J Obstet Gynaec.* 2012;119:1180—1190.
21. Skogerbø Å, Kesmodel US, Wimberley T, et al. The effects of low to moderate alcohol consumption and binge drinking in early pregnancy on executive function in 5-year-old children. *Brit J Obstet Gynaec.* 2012;119:1201—1210.
22. Underbjerg M, Kesmodel US, Landrø NI, et al. The effects of low to moderate alcohol consumption and binge drinking in early pregnancy on selective and sustained attention in 5-year-old children. *Brit J Obstet Gynaec.* 2012;119:1211—1221.
23. Kelly YJ, Sacker A, Gray R, et al. Light drinking during pregnancy: still no increased risk for socioemotional difficulties or cognitive deficits at 5 years of age? *J Epidemiol and Commun Health.* 2010:doi:10.1136/jech.2009.103002.
24. Kelly Y, Iacovou M, Quigley MA, et al. Light drinking versus abstinence in pregnancy – behavioural and cognitive outcomes in 7-year-old children: a longitudinal cohort study. *Brit J Obstet Gynaec.* 2013;120:1340—1347.
25. Kelly Y, Sacker A, Gray R, Kelly J, Wolke D, Quigley MA. Light drinking in pregnancy, a risk for behavioural problems and cognitive deficits at 3 years of age? *Int J Epidemiol.* 2009;38:129—140.
26. Day NL, Leech SL, Richardson GA, Cornelius MD, Robles N, Larkby C. Prenatal alcohol exposure predicts continued deficits in offspring size at 14 years of age. *Alcohol Clin Exp Res.* 2002;26:1584—1591.
27. Day NL, Richardson G, Robles N, et al. Effect of prenatal alcohol exposure on growth and morphology of offspring at 8 months of age. *Pediatrics.* 1990;85:748—752.
28. Day NL, Zuo Y, Richardson GA, Goldschmidt L, Larkby CA, Cornelius MD. Prenatal alcohol use and offspring size at 10 years of age. *Alcohol Clin Exp Res.* 1999;23:863—869.
29. O'Callaghan FV, O'Callaghan M, Najman JM, Williams GM, Bor W. Maternal alcohol consumption during pregnancy and physical outcomes up to 5 years of age: a longitudinal study. *Early Hum Dev.* 2003;71:137—148.
30. Barr HM, Streissguth AP, Martin DC, Herman CS. Infant size at 8 months of age: relationship to maternal use of alcohol, nicotine, and caffeine during pregnancy. *Pediatrics.* 1984;74:336—341.
31. O'Keeffe LM, Kearney PM, Greene RA, et al. Maternal alcohol use during pregnancy and offspring trajectories of height and weight: A prospective cohort study. *Drug Alcohol Depend.* 2015;153:323—329.
32. Gray R. Low-to-moderate alcohol consumption during pregnancy and child development – moving beyond observational studies. *Brit J Obstet Gynaec.* 2013;120:1039—1041.
33. Astley S, Grant T. Another perspective on 'The effect of different alcohol drinking patterns in early to mid pregnancy on the child's intelligence, attention, and executive function.' *Brit J Obstet Gynaec.* 2012;119:1672—1672.
34. International Center for Alcohol Studies (ICAP). International guidelines on drinking and pregnancy. 2009: http://www.icap.org/table/InternationalDrinkingGuidelines.html. Accessed April 5, 2016.

35. National Institute for Health and Care Excellence (NICE). *Antenatal care for uncomplicated pregnancies.* NICE Clinical Guideline 62;2008. https://www.nice.org.uk/guidance/cg62. Accessed April 5, 2016.
36. Department of Health. *How to keep health risks from drinking alcohol to a low level: public consultation on proposed new guidelines.* 2016. https://www.gov.uk/government/consultations/health-risks-from-alcohol-new-guidelines. Accessed April 5, 2016.
37. 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—1439.
38. Jones TB, Bailey BA, Sokol RJ. Alcohol use in pregnancy: insights in screening and intervention for the clinician. *Clin Obstet Gynecol*. 2013;56: 114—123.
39. Burns E, Gray R, Smith LA. Brief screening questionnaires to identify problem drinking during pregnancy: a systematic review. *Addiction.* 2010;105:601—614.
40. Muggli E, Cook, B., O'Leary, C., Forster, D., Halliday, J. Increasing accurate self-report in surveys of pregnancy alcohol use. *Midwifery*. 2015;31:e23—e28.
41. Ernhart CB, Morrow-Tlucak M, Sokol RJ, Martier S. Underreporting of alcohol use in pregnancy. *Alcohol Clin Exp Res* 1988; 12:506—511.
42. Feunekes GI, van 't Veer P, van Staveren WA, Kok FJ. Alcohol intake assessment: the sober facts. *Am J Epidemiol* 1999; 150:105—112.
43. Streissguth AP, Bookstein FL, Barr HM, Sampson PD, O'Malley K, Young JK. Risk factors for adverse life outcomes in fetal alcohol syndrome and fetal alcohol effects. *J Dev Behav Pediatr*. 2004;25:228—238.
44. World Health Organization. Guidelines for the identification and management of substance use and substance use disorders in pregnancy. 2014. http://www.who.int/substance_abuse/publications/pregnancy_guidelines/en/. Accessed April 29, 2015.
45. Astley SJ. Twenty years of patient surveys confirm a FASD 4-Digit-Code interdisciplinary diagnosis afforded substantial access to interventions that met patients' needs. *J Popul Ther Clin Pharmacol*. 2014;21:e81—e105.
46. Reid N, Dawe S, Shelton D, et al. Systematic review of fetal alcohol spectrum disorder interventions across the life span. *Alcohol Clin Exp Res*. 2015;39:2283–2295.
47. Astley SJ, Bailey D, Talbot C, Clarren SK. Fetal alcohol syndrome (FAS) primary prevention through FAS diagnosis: I. Identification of high-risk birth mothers through the diagnosis of their children. *Alcohol Alcohol.* 2000;35:499—508.
48. Astley SJ, Bailey D, Talbot C, Clarren SK. Fetal alcohol syndrome (FAS) primary prevention through FAS diagnosis: II. A comprehensive profile of 80 birth mothers of children with FAS. *Alcohol Alcohol.* 2000;35:509—519.
49. Bakhireva LN, Savage DD. Focus on: biomarkers of fetal alcohol exposure and fetal alcohol effects. *Alcohol Res Health*. 2011;34:56—63.
50. Bearer CF, Stoler JM, Cook JD, Carpenter S. Biomarkers of alcohol use in pregnancy. *Alcohol Res Health*. 2004;28:38—43.
51. Concheiro-Guisan A, Concheiro M. Bioanalysis during pregnancy: recent advances and novel sampling strategies. *Bioanalysis*. 2014;6:3133—3153.
52. Joya X, Friguls B, Ortigosa S, et al. Determination of maternal-fetal biomarkers of prenatal exposure to ethanol: a review. *J Pharm Biomed Anal*. 2012;69:209—222.
53. Littner Y, Bearer CF. Detection of alcohol consumption during pregnancy - current and future biomarkers. *Neurosci Biobehav Rev*. 2007;31:261—269.

54. Memo L, Gnoato E, Caminiti S, Pichini S, Tarani L. Fetal alcohol spectrum disorders and fetal alcohol syndrome: the state of the art and new diagnostic tools. *Early Hum Dev.* 2013;89:S40—S43.
55. Gifford AE, Bearer CF. Universal screening programs for gestational exposures. *J Pediatr.* 2015;166:522—524.
56. Canadian Association of Paediatric Health Centres. Fetal alcohol spectrum disorder national screening tool kit 2014. http://ken.caphc.org/xwiki/bin/view/FASDScreeningToolkit/National+Screening+Tool+Kit+for+Children+and+Youth+Identified+and+Potentially+Affected+by+FASD. Accessed July 16, 2015.
57. Deeks JJ, Bossuyt PM, Gatsonis C. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0*. 2009. http://srdta.cochrane.org/. Accessed October 16, 2014.
58. Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med.* 2003;138:W1—12.
59. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 2009;6:e1000097.
60. Burd L, Hofer R. Biomarkers for detection of prenatal alcohol exposure: a critical review of fatty acid ethyl esters in meconium. *Birth Defects Res A Clin Mol Teratol* 2008;82:487—493.
61. Chabenne A, Moon C, Ojo C, Khogali A, Nepal B, Sharma S. Biomarkers in fetal alcohol syndrome. *BGM.* 2014;6:12—22.
62. Wurst FM, Alling C, Aradottir S, et al. Emerging biomarkers: new directions and clinical applications. *Alcohol Clin Exp Res.* 2005;29:465—473.
63. Guidelines International Network (GIN). Templates of the evidence tables working group. 2014. http://www.g-i-n.net/working-groups/etwg/progresses-of-the-etwg. Accessed October 16, 2014.
64. United States National Institute on Alcohol Abuse and Alcoholism. Alcohol use and alcohol use disorders in the United States, a 3-year follow-up: main findings from the 2004–2005 Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). U.S. Alcohol Epidemiologic Data Reference Manual Vol. 8, No. 2. Bethesda, MD: National Institutes of Health. 2010. http://pubs.niaaa.nih.gov/publications/datasys.htm. Accessed November 15, 2015.
65. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155:529—536.
66. Reitsma JB RA, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Chapter 9: assessing methodological quality. In: Deeks JJ BP, Gatsonis C, eds. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0*: The Cochrane Collaboration. 2009. http://srdta.cochrane.org/. Accessed October 20, 2014.
67. Higgins J, Altman DG, Sterne JAC. Chapter 8: assessing risk of bias in included studies. In: Deeks JJ BP, Gatsonis C, eds. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 5.1.0:* The Cochrane Collaboration. 2011. http://srdta.cochrane.org/. Accessed November 6, 2014.
68. *Stata Statistical Software: Release 13* [computer program]. Texas: College Station, StataCorp LP; 2013.
69. Macaskill P, Gatsonis, C., Deeks, J., Harbord R., Takwoingi, Y. Chapter 10: analysing and presenting results. In: Deeks JJ BP, Gatsonis C., eds. *Cochrane*

Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0: Cochrane Collaboration; 2010. Accessed November 6, 2014.
70. *Review Manager (RevMan)* [computer program]. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration; 2008.
71. *R: A language and environment for statistical computing* [computer program]. Vienna, Austria; 2015.
72. Clopper C, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika.* 1934:404—413.
73. *MedCalc*: *version 16.1*. 2016; https://www.medcalc.org/calc/diagnostic_test.php. Accessed January 12, 2016.
74. Koopman PAR. Confidence intervals for the ratio of two binomial proportions. *Biometrics.* 1984;40:513—517.
75. Bakhireva LN, Leeman L, Savich RD, et al. The validity of phosphatidylethanol in dried blood spots of newborns for the identification of prenatal alcohol exposure. *Alcohol Clin Exp Res*. 2014;38:1078—1085.
76. Bearer CF, Jacobson JL, Jacobson SW, et al. Validation of a new biomarker of fetal exposure to alcohol. *J Pediatr*. 2003;143:463—469.
77. Bearer CF, Lee S, Salvator AE, et al. Ethyl linoleate in meconium: a biomarker for prenatal ethanol exposure. *Alcohol Clin Exp Res*. 1999;23:487—493.
78. Bearer CF, Santiago LM, O'Riordan MA, Buck K, Lee SC, Singer LT. Fatty acid ethyl esters: quantitative biomarkers for maternal alcohol consumption. *J Pediatr.* 2005;146:824—830.
79. Chan D, Bar-Oz B, Pellerin B, et al. Population baseline of meconium fatty acid ethyl esters among infants of nondrinking women in Jerusalem and Toronto. *Ther Drug Monit.* 2003;25:271—278.
80. Kwak HS, Han JY, Choi JS, et al. Dose-response and time-response analysis of total fatty acid ethyl esters in meconium as a biomarker of prenatal alcohol exposure. *Prenat Diag*. 2014a;34:831—838.
81. Kwak HS, Han JY, Choi JS, et al. Characterization of phosphatidylethanol blood concentrations for screening alcohol consumption in early pregnancy. *Clin Toxicol.* 2014b;52:25—31.
82. Ostrea EM, Hernandez JD, Bielawski DM, et al. Fatty acid ethyl esters in meconium: are they biomarkers of fetal alcohol exposure and effect? *Alcohol Clin Exp Res*. 2006;30:1152—1159.
83. Sarkola T, Eriksson CJ, Niemela O, Sillanaukee P, Halmesmaki E. Mean cell volume and gamma-glutamyl transferase are superior to carbohydrate-deficient transferrin and hemoglobin-acetaldehyde adducts in the follow-up of pregnant women with alcohol abuse. *Acta Obstet Gynecol Scand*. 2000;79:359—366.
84. Gauthier TW, Mohan SS, Gross TS, Harris FL, Guidot DM, Brown LAS. Placental fatty acid ethyl esters are elevated with maternal alcohol use in pregnancies complicated by prematurity. *PLoS ONE*. 2015;10:e0126552.
85. Gutierrez HL, Hund L, Shrestha S, et al. Ethyl glucuronide in maternal hair as a biomarker of prenatal alcohol exposure. *Alcohol.* 2015;49:617—623.
86. Joya X, Marchei E, Salat-Batlle J, et al. Fetal exposure to ethanol: relationship between ethyl glucuronide in maternal hair during pregnancy and ethyl glucuronide in neonatal meconium. *Clin Chem Lab Med*. 2015. doi:10.1515/cclm-2015-0516.
87. Morini L, Marchei E, Tarani L, et al. Testing ethyl glucuronide in maternal hair and nails for the assessment of fetal exposure to alcohol: comparison with meconium testing. *Ther Drug Monit.* 2013;35:402—407.

88. Pichini S, Morini L, Pacifici R, et al. Development of a new immunoassay for the detection of ethyl glucuronide (EtG) in meconium: validation with authentic specimens analyzed using LC-MS/MS. Preliminary results. *Clin Chem Lab Med*. 2014;52:1179—1185.
89. Morini L, Groppi A, Marchei E, et al. Population baseline of meconium ethyl glucuronide and ethyl sulfate concentrations in newborns of nondrinking women in 2 mediterranean cohorts. *Ther Drug Monit*. 2010;32:359—363.
90. Rutjes AW, Reitsma JB, Di Nisio M, Smidt N, Van Rijn JC, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. *CMAJ*. 2006;174:469—476.
91. Rutjes AW, Reitsma JB, Vandenbroucke JP, Glas AS, Bossuyt PM. Case–control and two-gate designs in diagnostic accuracy studies. *Clin Chem*. 2005;51:1335—1341.
92. Whiting PF, Rutjes AW, Westwood ME, Mallett S. A systematic review classifies sources of bias and variation in diagnostic test accuracy studies. *J Clinical Epidemiol*. 2013;66:1093—1104.
93. Bossuyt P DC, Deeks J, Hyde C, Leeflang M, Scholten R. . Chapter 11: Interpreting results and drawing conclusions. In: Deeks JJ BP, Gatsonis C., ed. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.9.* The Cochrane Collaboration; 2013.
94. Patra J, Bakker R, Irving H, Jaddoe VW, Malini S, Rehm J. Dose–response relationship between alcohol consumption before and during pregnancy and the risks of low birthweight, preterm birth and small for gestational age (SGA)—a systematic review and meta-analyses. *Brit J Obstet Gynaec*. 2011;118:1411—1421.
95. Altman DG, Bland JM. Statistics notes: diagnostic tests 2: predictive values. *BMJ*. 1994;309:102.
96. Zizzo N, Di Pietro N, Green C, Reynolds J, Bell E, Racine E. Comments and reflections on ethics in screening for biomarkers of prenatal alcohol exposure. *Alcohol Clin Exp Res*. 2013;37:1451—1455.
97. Marcellus L. Is meconium screening appropriate for universal use? Science and ethics say no. *Adv Neonatal Care*. 2007;7:207—214.
98. Nicholls SG, Wilson BJ, Etchegary H, et al. Benefits and burdens of newborn screening: public understanding and decision-making. *Per Med*. 2014;11:593—607.
99. Yan A, Bell E, Racine E. Ethical and social challenges in newborn screening for prenatal alcohol exposure. *Can J Neurol Sci*. 2014;41:115—118.
100. Goldberger BA, Cone EJ, Kadehjian L. Unsuspected ethanol ingestion through soft drinks and flavored beverages. *J Anal Toxicol*. 1996;20:332—333.
101. Reisfield GM, Goldberger BA, Crews BO, et al. Ethyl glucuronide, ethyl sulfate, and ethanol in urine after sustained exposure to an ethanol-based hand sanitizer. *J Anal Toxicol*. 2011;35:85—91.
102. Reisfield GM, Goldberger BA, Pesce AJ, et al. Ethyl glucuronide, ethyl sulfate, and ethanol in urine after intensive exposure to high ethanol content mouthwash. *J Anal Toxicol*. 2011;35:264—268.
103. Musshoff F, Albermann E, Madea B. Ethyl glucuronide and ethyl sulfate in urine after consumption of various beverages and foods-misleading results? *Int J Legal Med*. 2010;124:623—630.
104. Logan BK, Distefano S. Ethanol content of various foods and soft drinks and their potential for interference with a breath-alcohol test. *J Anal Toxicol*. 1998;22:181—183.
105. Magnusson Å, Göransson M, Heilig M. Unexpectedly high prevalence of alcohol use among pregnant Swedish women: failed detection by antenatal care and simple tools that improve detection. *J Stud Alcohol Drugs*. 2005;66:157—164.

106. Lange S, Shield K, Koren G, Rehm J, Popova S. A comparison of the prevalence of prenatal alcohol exposure obtained via maternal self-reports versus meconium testing: a systematic literature review and meta-analysis. *BMC Pregnancy Childbirth*. 2014;14: doi: 10.1186/1471-2393-1114-1127
107. Zhou X-H, McClish DK, Obuchowski NA. *Statistical methods in diagnostic medicine.* Vol 569. 2nd ed: John Wiley & Sons; 2011.
108. Wilson J, Jungner G. Principles and practice of screening for disease. Geneva: World Health Organization, 1968. *Public health papers.* 2011;34.
109. UK National Screening Committee (NSC). Criteria for appraising the viability, effectiveness and appropriateness of a screening programme. 2013; http://www.screening.nhs.uk/criteria. Accessed November 7, 2014.
110. Bakhireva LN, Savich RD, Raisch DW, et al. The feasibility and cost of neonatal screening for prenatal alcohol exposure by measuring phosphatidylethanol in dried blood spots. *Alcohol Clin Exp Res*. 2013;37:1008—1015.
111. Zelner I, Shor S, Lynn H, et al. Neonatal screening for prenatal alcohol exposure: Assessment of voluntary maternal participation in an open meconium screening program. *Alcohol.* 2012;46:269—276.
112. May PA, Blankenship J, Marais A-S, et al. Maternal alcohol consumption producing fetal alcohol spectrum disorders (FASD): quantity, frequency, and timing of drinking. *Drug Alcohol Depend*. 2013;133:502—512.
113. Sawada Feldman H, Lyons Jones K, Lindsay S, et al. Prenatal alcohol exposure patterns and alcohol-related birth defects and growth deficiencies: a prospective study. *Alcohol Clin Exp Res*. 2012;36:670—676.
114. Rutjes A, Reitsma J, Coomarasamy A, Khan K, Bossuyt P. Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technology Assessment.* 2007;11(50).
115. Peterson J, Kirchner HL, Xue W, Minnes S, Singer LT, Bearer CF. Fatty acid ethyl esters in meconium are associated with poorer neurodevelopmental outcomes to two years of age. *J Pediatr*. 2008;152:788—792.
116. Min MO, Singer LT, Minnes S, Wu M, Bearer CF. Association of fatty acid ethyl esters in meconium and cognitive development during childhood and adolescence. *J Pediatr*. 2015;166:1042—1047.
117. Caprara DL, Brien JF, Iqbal U, Reynolds JN, Klein J, Koren G. A guinea pig model for the identification of in utero alcohol exposure using fatty acid ethyl esters in neonatal hair. *Pediatr Res*. 2005;58:1158—1163.
118. Zelner I, Kenna K, Brien JF, et al. Meconium fatty acid ethyl esters as biomarkers of late gestational ethanol exposure and indicator of ethanol-induced multi-organ injury in fetal sheep. PLoS One. 2013;8: e59168.