Objective Measures of Prenatal Alcohol Exposure: A Systematic Review

Cheryl McQuire, MSc,³ Shantini Paranjothy, PhD,³ Lisa Hurt, PhD,³ Mala Mann, MILInfSc, MCLIP, Daniel Farewell, PhD,³ Alison Kemp, FRCPCH³

abstract

CONTEXT: Objective measurement of prenatal alcohol exposure (PAE) is essential for identifying children at risk for 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 (4 studies; sensitivity 0%–100%, specificity 79%–100%), maternal hair (2 studies; sensitivity 19%–87%, specificity 56%–86%) maternal urine (2 studies; sensitivity 5%–15%, specificity 97%–100%), and biomarker test batteries (3 studies; sensitivity 22%–50%, specificity 56%–97%). Tests of the total concentration of 4 fatty acid ethyl esters (in meconium: 2 studies; in placenta: 1 study) demonstrated high sensitivity (82%–100%); however, specificity was variable (13%–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.

³Division of Population Medicine and ²Specialist Unit for Review Evidence, Cardiff University, Cardiff, United Kingdom

Ms McQuire developed the draft protocol and search strategy, data extraction sheet, and quality coding criteria; carried out study selection, data extraction, quality assessment, and evidence synthesis; and prepared the initial manuscript and subsequent revisions; Dr 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 Hurt contributed to development of the protocol, extracted data, verified study inclusion decisions and quality assessments, and reviewed and revised the manuscript; Ms Mann contributed to the design of the search strategy and reviewed and revised the manuscript; Dr Farewell provided statistical advice and reviewed and revised the manuscript; Dr Kemp contributed to development of the protocol and reviewed and revised the manuscript; and all authors approved the final manuscript as submitted.

This review has been registered at the PROSPERO international prospective register of systematic reviews (www.crd.york.ac.uk) (identifier CRD42014015420).

DOI: 10.1542/peds.2016-0517
Accepted for publication Jun 15, 2016


Downloaded from www.aappublications.org/news by guest on July 8, 2021
Up to 82% of women consume alcohol while pregnant.¾ 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 (FASDs). 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 behavioral 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.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 contradictory. The National Institute for Health and Care Excellence recommend abstinence in the first trimester followed by no more than 1 to 2 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 standardized 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 behavior.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

---

**FIGURE 1**
Flow diagram depicting study selection process.
<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Cases</th>
<th>Controls</th>
<th>High-Risk Setting</th>
<th>PAE Level</th>
<th>Study Prevalence of PAE, %</th>
<th>Matrix</th>
<th>Analyte(s)</th>
<th>Period of Sample Collection</th>
<th>Reference Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bakhireva 2014</td>
<td>USA</td>
<td>28</td>
<td>32</td>
<td>Substance misuse clinic</td>
<td>Moderate-heavy</td>
<td>47</td>
<td>Meconium</td>
<td>FAE</td>
<td>Postnatal</td>
<td>TLFB and AUDIT</td>
</tr>
<tr>
<td>Bearer 1999</td>
<td>USA</td>
<td>56</td>
<td>88</td>
<td>Substance misuse clinic</td>
<td>Any</td>
<td>39</td>
<td>Meconium</td>
<td>FAE</td>
<td>Postnatal</td>
<td>Material postnatal interview</td>
</tr>
<tr>
<td>Bearer 2003</td>
<td>South Africa</td>
<td>19</td>
<td>6</td>
<td>Dop system region</td>
<td>Heavy</td>
<td>76</td>
<td>Meconium</td>
<td>FAE</td>
<td>Postnatal</td>
<td>TLFB</td>
</tr>
<tr>
<td>Bearer 2005</td>
<td>Jordan and USA</td>
<td>15(^a)</td>
<td>21(^b)</td>
<td>Substance misuse clinic</td>
<td>Heavy</td>
<td>6(^c)</td>
<td>Meconium</td>
<td>FAE</td>
<td>Postnatal</td>
<td>Material postnatal interview</td>
</tr>
<tr>
<td>Chan 2003</td>
<td>Canada and Israel</td>
<td>6</td>
<td>73</td>
<td>Women with alcoholism</td>
<td>Heavy</td>
<td>8</td>
<td>Meconium</td>
<td>FAE</td>
<td>Postnatal</td>
<td>Cases: unspecified self-report</td>
</tr>
<tr>
<td>Gauthier 2015</td>
<td>USA</td>
<td>11</td>
<td>69</td>
<td>Premature newborns</td>
<td>Heavy</td>
<td>14</td>
<td>Placenta</td>
<td>FAE</td>
<td>Postnatal</td>
<td>Self-report based on AUDIT</td>
</tr>
<tr>
<td>Gutierrez 2015</td>
<td>USA</td>
<td>42(^d)</td>
<td>43(^e)</td>
<td>Substance misuse clinic</td>
<td>Moderate-heavy</td>
<td>49</td>
<td>Maternal blood</td>
<td>CDT, GGT, PEth</td>
<td>Prenatal</td>
<td>TLFB and AUDIT</td>
</tr>
<tr>
<td>Joya 2019</td>
<td>Spain</td>
<td>30</td>
<td>50</td>
<td>NR</td>
<td>Any</td>
<td>38</td>
<td>Meconium</td>
<td>EtG, ESG</td>
<td>Postnatal</td>
<td>Meconium EtG</td>
</tr>
<tr>
<td>Kwak 2014a</td>
<td>Korea</td>
<td>54</td>
<td>182</td>
<td>NR</td>
<td>Low-moderate</td>
<td>23</td>
<td>Meconium</td>
<td>FAE</td>
<td>Postnatal</td>
<td>Unspecified self-report</td>
</tr>
<tr>
<td>Kwak 2014b</td>
<td>Korea</td>
<td>117</td>
<td>188</td>
<td>NR</td>
<td>Low</td>
<td>38</td>
<td>Maternal blood</td>
<td>PEth</td>
<td>Prenatal</td>
<td>Unspecified self-report</td>
</tr>
<tr>
<td>Ostrea 2006</td>
<td>USA</td>
<td>93</td>
<td>31</td>
<td>Substance misuse clinic</td>
<td>Any</td>
<td>75</td>
<td>Meconium</td>
<td>FAE</td>
<td>Postnatal</td>
<td>Unspecified self-report, MAST, CAGE and TACE</td>
</tr>
<tr>
<td>Sarkola 2000</td>
<td>Finland</td>
<td>13(^f)</td>
<td>31(^g)</td>
<td>Substance misuse clinic</td>
<td>Heavy</td>
<td>30</td>
<td>Maternal blood</td>
<td>CDT, GGT, Hb-Ach, MCV</td>
<td>Prenatal</td>
<td>Unspecified self-report</td>
</tr>
</tbody>
</table>

AUDIT, Alcohol Use Disorder Identification Test; CAGE, Cut down, Annoyed, Guilt, Eye-opener; MAST, Michigan Alcohol Screening Test; NR, not reported; TACE, Tolerance, Annoyed, Cut-down, Eye-opener; TLFB, timeline follow-back procedure.

\(^a\) Number of participants included in analysis. Cases are participants with the defined level of PAE within the study, and controls are those without PAE as defined by the study.

\(^b\) We excluded data from this study that compared the results of postnatal maternal EtG, ESS, GGT, CDT, and PEth in dried infant blood spots with prenatal self-report because of the short detection window of these biomarkers.

\(^c\) The number of cases and controls are not reported in this study. Therefore, the figures presented in the table are those that provided the closest match to the study outcome data based on a simulation of all possible values in R software.

\(^d\) Cases and controls for analysis of hair EtG. Because of missing data, there were 41 cases and 42 controls in the analysis of urine EtG and ESS, 40 cases and 43 controls in the analysis of PEth, and 40 cases and 42 controls in the analysis of GGT and CDT.

\(^e\) Because of missing data, there were 15 cases and 28 controls for GGT analyses due to exclusion of 3 participants with hepatitis C and elevated alanine aminotransferase.
meconium; and in maternal-fetal matrices such as the placenta. Direct biomarkers include ethanol itself or the products of ethanol metabolism. Indirect markers are those that signal alcohol-induced pathology after chronic alcohol use.

Some groups have advocated the use of biomarkers of PAE within universal screening programs. For example, meconium testing features as a recommended method within the Canadian FASD National Screening Tool Kit. 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, the Standards for the Reporting of Diagnostic Accuracy Studies (STARD), and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. The full protocol is available from the PROSPERO international prospective register of systematic reviews (www.crd.york.ac.uk/PROSPERO; 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 gray 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 onward to increase precision. A scoping exercise of existing nonsystematic reviews revealed no relevant articles before 1990. Full details of the search strategy for Medline is available in Supplemental Table 8. 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 additional information when there were missing or conflicting data.

### Inclusion–Exclusion Criteria

Eligible studies were randomized 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 with any reference standard.

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

### Study Selection

After removing duplicates, C.M. screened the search records against...
predetermined 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 was independently assessed for eligibility by 2 other reviewers (L.H. and S.P.). The level of agreement for inclusion decisions was 100%.

Data Extraction

All data were extracted by C.M. into a standardized electronic form, designed based on the Guidelines International Network template for diagnostic studies and STARD.63 S.P. or L.H. independently repeated data extraction for each study 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 US National Institute on Alcohol Abuse and Alcoholism criteria for the general population in which 1 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 Supplemental Table 9 for quality coding criteria).66,67 Methodological quality was independently assessed by 2 members of the review team (C.M. and S.P. or L.H.) 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 4 studies per test are required for meta-analyses of diagnostic accuracy data within Stata using the recommended Rutter and Gatsonis HSROC model.68,69 Because of 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, Cochrane, London, United Kingdom). 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 (Gnu Project, Boston, MA) to conduct an exhaustive search of all possible 2 × 2 tables that were consistent with the data supplied in the primary study. Confidence intervals were generated automatically using the exact binomial method in RevMan for sensitivity and specificity values and with the MedCalc online calculator (MedCalc Software, Ostend, Belgium) for predictive values. Confidence intervals for likelihood ratios were generated by using the method described by Koopman.71

Many studies reported a range of diagnostic accuracy values according to characteristics such as positivity cutoff and period of measurement. 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 because the study used secondary data.

RESULTS

Characteristics of Included Studies

From 4278 search records, 12 studies, including 1614 unique
TABLE 2 Diagnostic Accuracy Outcomes for Studies of Meconium Testing for PAE

<table>
<thead>
<tr>
<th>Test Analyte(s)</th>
<th>Study</th>
<th>Trimester of PAE</th>
<th>Assay Positivity Threshold</th>
<th>Sens % (95% CI)</th>
<th>Spec % (95% CI)</th>
<th>PPV % (95% CI)</th>
<th>NPV % (95% CI)</th>
<th>LR+ (95% CI)</th>
<th>LR– (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethyl arachidonate</td>
<td>Bearer 2005a, b</td>
<td>3rd</td>
<td>GC/FID</td>
<td>306 ng/g</td>
<td>88 (55–98)</td>
<td>63 (56–70)</td>
<td>9 (6–23)</td>
<td>99 (95–100)</td>
<td>2.3 (1.7–3.1)</td>
</tr>
<tr>
<td></td>
<td>Ostrea 2006</td>
<td>Any</td>
<td>GC/MS</td>
<td>902 ng/g</td>
<td>18 (11–28)</td>
<td>97 (83–100)</td>
<td>94 (73–100)</td>
<td>28 (20–38)</td>
<td>5.7 (0.8–40.9)</td>
</tr>
<tr>
<td>Ethyl docosahexanoate</td>
<td>Ostrea 2006</td>
<td>Any</td>
<td>GC/MS</td>
<td>1000 ng/g</td>
<td>4 (1–11)</td>
<td>100 (89–100)</td>
<td>100 (40–100)</td>
<td>∞ (0.1–∞)</td>
<td>1.0 (0.9–1.0)</td>
</tr>
<tr>
<td>Ethyl laurate</td>
<td>Ostrea 2006</td>
<td>Any</td>
<td>GC/MS</td>
<td>50 ng/g</td>
<td>19 (12–29)</td>
<td>81 (63–93)</td>
<td>75 (53–90)</td>
<td>25 (18–35)</td>
<td>1.0 (0.4–2.3)</td>
</tr>
<tr>
<td>Ethyl linoleate</td>
<td>Bearer 2005a, b</td>
<td>2nd</td>
<td>GC/FID</td>
<td>383 ng/g</td>
<td>89 (64–100)</td>
<td>58 (51–65)</td>
<td>9 (6–20)</td>
<td>99 (95–100)</td>
<td>2.2 (1.8–2.7)</td>
</tr>
<tr>
<td></td>
<td>Bearer 2003</td>
<td>Any</td>
<td>GC/MS/MS</td>
<td>1 pmol/g</td>
<td>68 (54–80)</td>
<td>51 (40–62)</td>
<td>47 (36–58)</td>
<td>71 (60–82)</td>
<td>1.4 (1.1–1.8)</td>
</tr>
<tr>
<td>Ethyl linolenate</td>
<td>Ostrea 2006</td>
<td>Any</td>
<td>GC/MS</td>
<td>100 ng/g</td>
<td>3 (1–9)</td>
<td>100 (89–100)</td>
<td>100 (29–100)</td>
<td>∞ (0.1–∞)</td>
<td>1.0 (0.9–1.0)</td>
</tr>
<tr>
<td>Ethyl myristate</td>
<td>Ostrea 2006</td>
<td>Any</td>
<td>GC/MS</td>
<td>50 ng/g</td>
<td>68 (57–77)</td>
<td>29 (14–48)</td>
<td>74 (63–85)</td>
<td>23 (18–34)</td>
<td>8.3 (1.2–59.0)</td>
</tr>
<tr>
<td>Ethyl oleate</td>
<td>Bearer 2005a, b</td>
<td>3rd</td>
<td>GC/FID</td>
<td>445 ng/g</td>
<td>80 (46–95)</td>
<td>58 (51–65)</td>
<td>5 (5–18)</td>
<td>99 (93–100)</td>
<td>1.8 (1.3–2.6)</td>
</tr>
<tr>
<td></td>
<td>Bearer 2003</td>
<td>Any</td>
<td>GC/MS/MS</td>
<td>32 ng/g</td>
<td>84 (60–97)</td>
<td>83 (36–100)</td>
<td>94 (71–100)</td>
<td>67 (54–91)</td>
<td>5.1 (0.8–3.9)</td>
</tr>
<tr>
<td>Bakhireva 2014</td>
<td>Any</td>
<td>GC/MS</td>
<td>600 ng/g</td>
<td>100 (88–100)</td>
<td>13 (4–29)</td>
<td>50 (36–64)</td>
<td>100 (40–100)</td>
<td>1.1 (0.3–4.2)</td>
<td>1.1 (0.4–2.1)</td>
</tr>
<tr>
<td>Total concentration of 4 FAEEs</td>
<td>Bakhireva 2014</td>
<td>Any</td>
<td>GC/MS</td>
<td>10 000 ng/g</td>
<td>86 (67–96)</td>
<td>21 (14–30)</td>
<td>74 (63–85)</td>
<td>23 (18–34)</td>
<td>8.3 (1.2–59.0)</td>
</tr>
<tr>
<td>Total concentration of 6 FAEEs</td>
<td>Bakhireva 2014</td>
<td>Any</td>
<td>GC/MS</td>
<td>100 000 ng/g</td>
<td>29 (13–49)</td>
<td>81 (63–93)</td>
<td>57 (29–82)</td>
<td>57 (41–71)</td>
<td>1.5 (0.5–4.1)</td>
</tr>
<tr>
<td>Total concentration of 9 FAEEs</td>
<td>Bakhireva 2014</td>
<td>Any</td>
<td>GC/MS</td>
<td>10 000 ng/g</td>
<td>58 (47–68)</td>
<td>42 (25–61)</td>
<td>75 (63–84)</td>
<td>25 (18–36)</td>
<td>1.0 (0.7–1.4)</td>
</tr>
<tr>
<td>Kwak 2014a</td>
<td>2nd or 3rd</td>
<td>LC/MS/MS</td>
<td>20 nmol/g</td>
<td>4 (0–13)</td>
<td>98 (95–100)</td>
<td>41 (6–85)</td>
<td>78 (72–89)</td>
<td>2.2 (0.3–13.1)</td>
<td>1.0 (0.9–1.0)</td>
</tr>
</tbody>
</table>

CI, confidence interval; GC/FID, gas chromatography with flame ionization detection; GC/MS, gas chromatography mass spectrometry; GC/MS/MS, gas chromatography tandem mass spectrometry; LR+, positive likelihood ratio; LR–, negative likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; sens, sensitivity; spec, specificity.

a Confidence intervals are an approximation, due to missing data, based on a simulation of all possible 2 × 2 tables in R software assuming N = 224 and 13 individuals with the target condition (cases). The summary statistics reported in this table are those from the original study. The corresponding values suggested by the R simulation were sensitivity 85%, specificity 95%, PPV 12%, and NPV 99% for ethyl arachidonate; sensitivity 92%, specificity 58%, PPV 12%, and NPV 99% for ethyl linoleate; and sensitivity 77%, specificity 58%, PPV 10%, and NPV 98% for ethyl oleate.

b Bearer 1999 and Bearer 2005 include the same participants.

c Ethyl palmitate, ethyl stearate, ethyl oleate, and ethyl linoleate.

d For Chan 2003, the PPV presented in this table (63%) is the value reported in the primary study. However, our calculations produce a PPV of 87% based on the reported number of participants included in the sensitivity and specificity analyses (79 total, including 6 cases) and reported values of 100% specificity and 98% sensitivity. The authors were unable to provide the original data to further explore this discrepancy.

e Ethyl palmitate, ethyl palmitoleate, ethyl stearate, ethyl oleate, ethyl linoleate, and ethyl arachidonate.

f Ethyl palmitate, ethyl palmitoleate, ethyl stearate, ethyl oleate, ethyl linoleate, ethyl linolenate, ethyl arachidonate, ethyl laurate, and ethyl myristate.
TABLE 3 Diagnostic Accuracy Outcomes for 1 Study of Placenta Testing for PAE

<table>
<thead>
<tr>
<th>Index Test Analyte(s)</th>
<th>Study</th>
<th>Trimester of Assay</th>
<th>Assay</th>
<th>Positivity Threshold</th>
<th>Sens % (95% CI)</th>
<th>Spec % (95% CI)</th>
<th>PPV % (95% CI)</th>
<th>NPV % (95% CI)</th>
<th>LR+ (95% CI)</th>
<th>LR– (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethyl stearate</td>
<td>Gauthier 2015</td>
<td>Any GC/MS</td>
<td>NR</td>
<td>82 (48–98)</td>
<td>87 (77–94)</td>
<td>50 (28–74)</td>
<td>97 (88–100)</td>
<td>6.3 (3.2–12.3)</td>
<td>0.2 (0.1–0.7)</td>
<td></td>
</tr>
<tr>
<td>Ethyl linoleate</td>
<td>Gauthier 2015</td>
<td>Any GC/MS</td>
<td>NR</td>
<td>82 (48–98)</td>
<td>83 (72–91)</td>
<td>44 (22–66)</td>
<td>97 (88–100)</td>
<td>4.7 (2.6–8.4)</td>
<td>0.2 (0.1–0.8)</td>
<td></td>
</tr>
<tr>
<td>Total concentration of 3 FAEEs</td>
<td>Gauthier 2015</td>
<td>Any GC/MS</td>
<td>NR</td>
<td>82 (48–98)</td>
<td>94 (89–98)</td>
<td>70 (39–91)</td>
<td>97 (90–100)</td>
<td>14.1 (5.2–38.0)</td>
<td>0.2 (0.1–0.7)</td>
<td></td>
</tr>
<tr>
<td>Ethyl oleate, ethyl linoleate, and ethyl linolenate.</td>
<td>Gauthier 2015</td>
<td>Any GC/MS</td>
<td>NR</td>
<td>82 (48–98)</td>
<td>93 (84–98)</td>
<td>64 (55–87)</td>
<td>97 (89–100)</td>
<td>11.3 (4.6–27.5)</td>
<td>0.2 (0.1–0.7)</td>
<td></td>
</tr>
<tr>
<td>Total concentration of 4 FAEEs</td>
<td>Gauthier 2015</td>
<td>Any GC/MS</td>
<td>NR</td>
<td>82 (48–98)</td>
<td>93 (84–98)</td>
<td>64 (55–87)</td>
<td>97 (89–100)</td>
<td>11.3 (4.6–27.5)</td>
<td>0.2 (0.1–0.7)</td>
<td></td>
</tr>
</tbody>
</table>

CI, confidence interval; GC/MS, gas chromatography mass spectrometry; LR+, positive likelihood ratio; LR–, negative likelihood ratio; NPV, negative predictive value; NR, not reported; PPV, positive predictive value; sens, sensitivity; spec, specificity.

*Ethyl oleate, ethyl linoleate, and ethyl linolenate.

**Ethyl oleate, ethyl linoleate, ethyl linolenate, and ethyl stearate.
deemed to have a high risk of bias because 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 because of the way findings were reported. One study, which collected maternal hair during pregnancy for EtG testing, was also considered to have a high risk of bias because the specimen may have captured alcohol use before pregnancy due to the broad detection window of EtG within this matrix.

Finally, selective outcome reporting introduced a high risk of bias in 4 studies. Of these studies, 3 measured multiple FAEEs in meconium, but only reported diagnostic accuracy outcomes for a subset of these FAEEs. Two studies 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, specificity, and predictive values reported in 1 of the studies and generated values that approximated the published data in another study. For the remaining study, we were unable to produce data that replicated the published data because the positive predictive value reported in the study did not match the value suggested by the raw data (Table 2). For the remaining study, we were unable to produce data that replicated the published data because the positive predictive value reported in the study did not match the value suggested by the raw data (Table 2).

<table>
<thead>
<tr>
<th>Index Test Analyte(s)</th>
<th>Study</th>
<th>Trimester of PAE</th>
<th>Assay</th>
<th>Positivity Threshold</th>
<th>Sens % (95% CI)</th>
<th>Spec % (95% CI)</th>
<th>PPV % (95% CI)</th>
<th>NPV % (95% CI)</th>
<th>LR+ (95% CI)</th>
<th>LR– (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>%CDT</td>
<td>Bakhireva 2014</td>
<td>Any</td>
<td>HPLC</td>
<td>2%</td>
<td>4 (0–18)</td>
<td>100 (89–100)</td>
<td>100 (3–100)</td>
<td>54 (41–67)</td>
<td>∞ (0.1–∞)</td>
<td>1.0 (0.9–1.0)</td>
</tr>
<tr>
<td></td>
<td>Gutierrez 2015</td>
<td>Any</td>
<td>HPLC</td>
<td>2%</td>
<td>5 (1–17)</td>
<td>100 (82–100)</td>
<td>100 (16–100)</td>
<td>53 (41–64)</td>
<td>∞ (0.2–∞)</td>
<td>1.0 (0.9–1.0)</td>
</tr>
<tr>
<td>Total CDT</td>
<td>Sarkola 2000</td>
<td>Any</td>
<td>Enzymatic rate method</td>
<td>40 U/L</td>
<td>15 (4–33)</td>
<td>100 (89–100)</td>
<td>100 (40–100)</td>
<td>57 (43–70)</td>
<td>∞ (0.5–∞)</td>
<td>0.9 (0.7–1.0)</td>
</tr>
<tr>
<td>GGT</td>
<td>Bakhireva 2014</td>
<td>Any</td>
<td>Enzymatic rate method</td>
<td>40 U/L</td>
<td>20 (9–36)</td>
<td>98 (87–100)</td>
<td>89 (52–100)</td>
<td>56 (44–68)</td>
<td>8.4 (1.1–64.2)</td>
<td>0.8 (0.7–1.0)</td>
</tr>
<tr>
<td></td>
<td>Gutierrez 2015</td>
<td>Any</td>
<td>Immunoassay</td>
<td>NR</td>
<td>31 (9–61)</td>
<td>79 (53–92)</td>
<td>40 (12–74)</td>
<td>71 (52–86)</td>
<td>1.4 (0.5–4.2)</td>
<td>0.9 (0.6–1.3)</td>
</tr>
<tr>
<td></td>
<td>Sarkola 2000</td>
<td>Any</td>
<td>Immuno-enzymatic method</td>
<td>NR</td>
<td>0 (0–25)</td>
<td>97 (85–100)</td>
<td>0 (0–98)</td>
<td>70 (54–83)</td>
<td>0.0 (0.0–33.4)</td>
<td>1.0 (1.0–1.1)</td>
</tr>
<tr>
<td>MCV</td>
<td>Sarkola 2000</td>
<td>Any</td>
<td>Coulter C7 cell counter</td>
<td>NR</td>
<td>15 (2–45)</td>
<td>100 (89–100)</td>
<td>100 (16–100)</td>
<td>74 (58–86)</td>
<td>∞ (0.5–∞)</td>
<td>0.8 (0.7–1.1)</td>
</tr>
<tr>
<td>PEth</td>
<td>Bakhireva 2014</td>
<td>Any</td>
<td>LC/MS/MS</td>
<td>8 ng/mL</td>
<td>22 (8–41)</td>
<td>100 (83–100)</td>
<td>100 (54–100)</td>
<td>60 (45–72)</td>
<td>∞ (0.8–∞)</td>
<td>0.8 (0.9–1.0)</td>
</tr>
<tr>
<td></td>
<td>Gutierrez 2015</td>
<td>Any</td>
<td>LC/MS/MS</td>
<td>8 ng/mL</td>
<td>18 (7–33)</td>
<td>100 (82–100)</td>
<td>100 (53–100)</td>
<td>57 (45–68)</td>
<td>∞ (0.9–∞)</td>
<td>0.8 (0.7–1.0)</td>
</tr>
<tr>
<td></td>
<td>Kwak 2014b 1st</td>
<td>LC/MS/MS</td>
<td>4.2 nmol/L</td>
<td>100 (63–100)</td>
<td>98 (83–98)</td>
<td>43 (20–67)</td>
<td>100 (99–100)</td>
<td>27 (15.1–48.2)</td>
<td>0.0 (0.0–0.9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kwak 2014b 2nd</td>
<td>LC/MS/MS</td>
<td>3.8 nmol/L</td>
<td>67 (47–83)</td>
<td>96 (83–98)</td>
<td>67 (47–83)</td>
<td>96 (83–98)</td>
<td>18.3 (9.5–35.4)</td>
<td>0.3 (0.2–0.8)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kwak 2014b 3rd</td>
<td>LC/MS/MS</td>
<td>1.2 nmol/L</td>
<td>41 (32–50)</td>
<td>95 (91–98)</td>
<td>85 (72–93)</td>
<td>72 (66–87)</td>
<td>8.6 (4.4–16.8)</td>
<td>0.6 (0.5–0.7)</td>
<td></td>
</tr>
</tbody>
</table>

CI, confidence interval; HPLC, high-performance liquid chromatography; LC/MS/MS, liquid chromatography tandem mass spectrometry; LR+, positive likelihood ratio; LR–, negative likelihood ratio; NPV, negative predictive value; NR, not reported; PPV, positive predictive value; sens, sensitivity; spec, specificity.

Percentage of CDT in total transferrin.

A threshold of 4.2 nmol/L was specified for the detection of heavy PAE, 3.8 nmol/L for moderate PAE, and 1.2 nmol/L for low PAE.
Diagnostic Accuracy Findings

Tables 2 through 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 7 studies. The diagnostic accuracy of FAEEs varied widely across studies (see Table 2). A measure of the total concentration of 4 FAEEs demonstrated the highest levels of diagnostic accuracy overall, but there were a high number of false-positives in 1 study, and specificity was inconsistent.

Placenta Testing

FAEEs were measured in placenta tissue in 1 study of premature deliveries. Sensitivity and specificity values were high, although 30% to 56% of positive test results were false-positives (Table 3).

Blood Testing

Four studies investigated CDT, GGT, Hb-Ach, MCV, and PEth within prenatal samples of maternal blood. Blood biomarkers generally demonstrated low sensitivity, although specificity values were high, and positive likelihood ratios ranged from 1, suggesting little test utility, to 73. As a guide, positive likelihood ratios >10 may indicate that a test is informative. Because of the 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. It is important to note that because alcohol is a risk factor for prematurity, 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.

Hair Testing

Maternal hair was tested for EtG in 2 studies with contrasting findings (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 2 studies (Table 6).

Test Batteries

Three studies investigated the diagnostic accuracy of test batteries including combinations of different biomarkers across multiple matrices. Sensitivity was poor, whereas specificity was generally good (Table 7).

DISCUSSION

This systematic review demonstrates that the accuracy of biomarkers of PAE varies widely across studies. Tests of the total concentration of 4 FAEEs demonstrated the highest levels of sensitivity across studies. Sensitivity was 100% in 2 studies of meconium testing and 82% in 1 study of placenta testing. However, specificity was variable (13%–98%). Positive likelihood ratios ranged from 1, suggesting little test utility, to 73. As a guide, positive likelihood ratios >10 may indicate that a test is informative. Because of the 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. It is important to note that because alcohol is a risk factor for prematurity, 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.
false-positive results to maximize early detection of asymptomatic conditions. Early diagnosis and intervention are associated with improved outcomes for children with FASD, and thus it could be argued that a test with high sensitivity could be favored over specificity in this context because 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 stigmatization and unnecessary burden on health care resources, and may even be used in legal proceedings against mothers. Conversely, false-negative errors represent a missed opportunity to provide support to families affected by PAE. Many screening programs are conducted in a tiered fashion with high sensitivity favored over specificity in the initial phase. Second-tier screening using methods such as detailed maternal interviews and behavioral assessment of children with suspected PAE have been suggested as a strategy to reduce false-positive results. However, these methods also have limitations. Information from maternal interview may be inaccurate, and the cognitive-behavioral 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.

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 of biomarker tests for PAE observed in this review reflects a true lack of test validity or is simply an artifact of comparison with an imperfect reference standard. Because 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, whereas 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 sanitizers, although the extent to which incidental exposure influences biomarker tests for PAE has not been fully established. Conversely, false-negative results are unlikely 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 of reporting by pregnant women who report drinking modest amounts of alcohol detected by the self-report reference standard but not the index test. This explanation is unlikely because the majority of studies investigated moderate to heavy self-reported PAE.

### TABLE 6 Diagnostic Accuracy Outcomes for Studies of Maternal Urine Testing for PAE

<table>
<thead>
<tr>
<th>Index Test Analyte(s)</th>
<th>Study</th>
<th>Trimester of PAE Assay</th>
<th>Sens % (95% CI)</th>
<th>Spec % (95% CI)</th>
<th>PPV % (95% CI)</th>
<th>NPV % (95% CI)</th>
<th>LR+ (95% CI)</th>
<th>LR– (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EtG</td>
<td>Bakhireva 2014</td>
<td>Any</td>
<td>LC/MS/MS</td>
<td>25 ng/mL</td>
<td>15 (4–33)</td>
<td>97 (84–100)</td>
<td>81 (28–98)</td>
<td>57 (42–70)</td>
</tr>
<tr>
<td>EtG</td>
<td>Gutierrez 2015</td>
<td>Any</td>
<td>LC/MS/MS</td>
<td>33 ng/mL</td>
<td>6 (1–17)</td>
<td>98 (87–100)</td>
<td>67 (9–99)</td>
<td>51 (40–63)</td>
</tr>
<tr>
<td>EtS</td>
<td>Bakhireva 2014</td>
<td>Any</td>
<td>LC/MS/MS</td>
<td>7 ng/mL</td>
<td>15 (4–33)</td>
<td>100 (89–100)</td>
<td>100 (40–100)</td>
<td>57 (43–70)</td>
</tr>
<tr>
<td>EtS</td>
<td>Gutierrez 2015</td>
<td>Any</td>
<td>LC/MS/MS</td>
<td>7 ng/mL</td>
<td>6 (2–20)</td>
<td>98 (87–100)</td>
<td>75 (19–99)</td>
<td>52 (40–63)</td>
</tr>
</tbody>
</table>

CI, confidence interval; LC/MS/MS, liquid chromatography tandem mass spectrometry; LR+, positive likelihood ratio; LR–, negative likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; sens, sensitivity; spec, specificity.
<table>
<thead>
<tr>
<th>Index Test Analyte(s)</th>
<th>Study</th>
<th>Trimester of PAE</th>
<th>Assay</th>
<th>Positivity Threshold</th>
<th>Sens % (95% CI)</th>
<th>Spec % (95% CI)</th>
<th>PPV % (95% CI)</th>
<th>NPV % (95% CI)</th>
<th>LR+ (95% CI)</th>
<th>LR– (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood GGT, %CDT and PEth, urine EtG, and EtS</td>
<td>Bakhireva 2014</td>
<td>Any</td>
<td>Enzymatic rate method, HPLC, LC/MS/MS</td>
<td>GGT 40 U/L</td>
<td>32 (16–52)</td>
<td>97 (84–100)</td>
<td>90 (56–100)</td>
<td>62 (47–75)</td>
<td>10.3 (1.4–76.2)</td>
<td>0.7 (0.5–0.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CDT 2%</td>
<td>PEth 8 ng/mL</td>
<td>EtG 25 ng/mL</td>
<td>EtS 7 ng/mL</td>
<td>EtG 8 pg/mg</td>
<td>97 (84–100)</td>
<td>97 (84–100)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PEth 8 ng/mL</td>
<td>EtG 25 ng/mL</td>
<td>EtS 7 ng/mL</td>
<td>EtS 7 ng/mL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hair EtG and urine EtG</td>
<td>Gutierrez 2015</td>
<td>Any</td>
<td>LC/MS/MS</td>
<td>GGT 38.7 ng/mL</td>
<td>22 (11–38)</td>
<td>83 (69–93)</td>
<td>56 (30–80)</td>
<td>52 (40–65)</td>
<td>1.3 (0.5–3.2)</td>
<td>0.9 (0.8–1.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EtG 8 pg/mg</td>
<td>24 (12–40)</td>
<td>83 (69–93)</td>
<td>59 (33–82)</td>
<td>53 (40–65)</td>
<td>1.5 (0.6–3.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EtS 7.2 ng/mL</td>
<td>EtG 8 pg/mg</td>
<td>33 (19–49)</td>
<td>83 (69–93)</td>
<td>65 (41–85)</td>
</tr>
<tr>
<td>Hair EtG and blood GGT</td>
<td>Gutierrez 2015</td>
<td>Any</td>
<td>LC/MS/MS, enzymatic rate method</td>
<td>GGT 40 U/L</td>
<td>23 (11–38)</td>
<td>86 (71–95)</td>
<td>60 (32–84)</td>
<td>54 (41–66)</td>
<td>1.6 (0.6–4.0)</td>
<td>0.9 (0.7–1.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EtG 8 pg/mg</td>
<td>28 (15–44)</td>
<td>86 (72–95)</td>
<td>65 (38–80)</td>
<td>56 (43–68)</td>
<td>2.0 (0.8–4.8)</td>
</tr>
<tr>
<td>Hair EtG and blood PEth</td>
<td>Gutierrez 2015</td>
<td>Any</td>
<td>LC/MS/MS</td>
<td>PEth 8 pg/mg</td>
<td>50 (34–66)</td>
<td>56 (41–72)</td>
<td>53 (36–69)</td>
<td>54 (39–70)</td>
<td>1.2 (0.7–1.9)</td>
<td>0.9 (0.6–1.3)</td>
</tr>
<tr>
<td>Hair EtG and blood CDT and GGT</td>
<td>Gutierrez 2015</td>
<td>Any</td>
<td>LC/MS/MS, HPLC, enzymatic rate method</td>
<td>GGT 40 U/L</td>
<td>38 (23–54)</td>
<td>83 (69–93)</td>
<td>68 (45–88)</td>
<td>58 (45–71)</td>
<td>2.3 (1.0–4.9)</td>
<td>0.8 (0.6–1.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EtG 8 pg/mg</td>
<td>43 (27–59)</td>
<td>60 (43–74)</td>
<td>50 (32–68)</td>
<td>52 (37–67)</td>
<td>1.1 (0.6–1.8)</td>
</tr>
<tr>
<td>Hair EtG and blood GGT and PEth</td>
<td>Gutierrez 2015</td>
<td>Any</td>
<td>LC/MS/MS, enzymatic rate method</td>
<td>GGT 40 U/L</td>
<td>38 (14–68)</td>
<td>79 (59–92)</td>
<td>45 (17–77)</td>
<td>73 (54–88)</td>
<td>1.8 (0.7–4.8)</td>
<td>0.8 (0.5–1.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EtG 8 pg/mg</td>
<td>43 (27–59)</td>
<td>60 (43–74)</td>
<td>50 (32–68)</td>
<td>52 (37–67)</td>
<td>1.1 (0.6–1.8)</td>
</tr>
</tbody>
</table>

CI, confidence interval; HPLC, high-performance liquid chromatography; LC/MS/MS, liquid chromatography tandem mass spectrometry; LR+, positive likelihood ratio; LR–, negative likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; sens, sensitivity; spec, specificity.

a A test battery was considered positive if at least 1 of the results for an individual biomarker exceeded the designated positivity threshold.
Case-control diagnostic designs may produce overestimates of diagnostic accuracy. However, this form of bias is not likely to influence the conclusions of this review because 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 generalize predictive values from individual studies to settings with a different prevalence of PAE.

Finally, selective outcome reporting was common among studies and introduced a further risk of bias. As recommended by the STARD guidelines, 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 nonsystematic reviews, which suggest the need for improved objective measures of PAE. However, because of 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 PAE must be accurate, feasible, and acceptable to the population. These criteria are emphasized in the World Health Organization and UK National Screening Committee guidelines for screening procedures.

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.

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 standardize 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) cutoff for total concentration of 4 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.

With the exception of hair testing, objective measures of PAE have a limited detection period, which does not span the whole of pregnancy (see Supplemental Table 9). 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. Risk of harm to the developing fetus is highest if a mother drinks heavily throughout pregnancy; however, first trimester exposure poses a particular risk of physical abnormalities, including dysmorphic facial features. Meconium testing only captures PAE late in the second and third trimesters of pregnancy and therefore may fail to identify a large proportion of infants at risk for alcohol-related harm. 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. An objective test that 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. Because of the limitations of current biomarkers, authors have emphasized the need for novel biomarkers that can detect even low levels of alcohol use across the duration of pregnancy.

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 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.\textsuperscript{112,113} 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.\textsuperscript{49,114,115} 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.\textsuperscript{111}

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, and 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 suboptimal 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 epidemiologic research.

ACKNOWLEDGMENTS

We thank Dr Helena Kemp, consultant chemical pathologist, for her feedback on previous versions of this article. We also 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.

ABBREVIATIONS

CDT: carbohydrate deficient transferrin  
EtG: ethyl glucuronide  
EtS: ethyl sulfate  
FAEE: fatty acid ethyl ester  
FASD: fetal alcohol spectrum disorder  
GGT: γ glutamyltransferase  
Hb-Ach: hemoglobin acetaldehyde adducts  
MCV: mean corpuscular volume  
PAE: prenatal alcohol exposure  
PEth: phosphatidylethanol  
QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies tool  
STARD: Standards for the Reporting of Diagnostic Accuracy Studies tool

Address correspondence to Cheryl McQuire, MSc, Division of Population Medicine, Cardiff University, 3rd Floor, Neuadd Meirionnydd, Heath Park, Cardiff, CF14 4YS, United Kingdom. E-mail: mcquirec1@cardiff.ac.uk

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).  
Copyright © 2016 by the American Academy of Pediatrics

FINANCIAL DISCLOSURE: The authors have indicated they have no financial relationships relevant to this article to disclose.

FUNDING: Ms McQuire 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.

POTENTIAL CONFLICT OF INTEREST: The authors have indicated they have no potential conflicts of interest to disclose.

REFERENCES


Objective Measures of Prenatal Alcohol Exposure: A Systematic Review
Cheryl McQuire, Shantini Paranjothy, Lisa Hurt, Mala Mann, Daniel Farewell and Alison Kemp
PEDIATRICS 2016;138;
DOI: 10.1542/peds.2016-0517 originally published online August 30, 2016;
Objective Measures of Prenatal Alcohol Exposure: A Systematic Review
Cheryl McQuire, Shantini Paranjothy, Lisa Hurt, Mala Mann, Daniel Farewell and Alison Kemp
Pediatrics 2016;138;
DOI: 10.1542/peds.2016-0517 originally published online August 30, 2016;

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://pediatrics.aappublications.org/content/138/3/e20160517

Data Supplement at:
http://pediatrics.aappublications.org/content/suppl/2016/08/25/peds.2016-0517.DCSupplemental