Prenatal Alcohol Exposure, FASD, and Child Behavior: A Meta-analysis

Tracey W. Tsang, PhD,[a,b] Barbara R. Lucas, MPH,[a,b,c,d] Heather Carmichael Olson, PhD,[e] Rafael Z. Pinto, PhD,[f] Elizabeth J. Elliott, MD[a,b,g,h]

abstract

CONTEXT: Fetal alcohol spectrum disorders (FASD) and prenatal alcohol exposure (PAE) are associated with behavioral difficulties, although there are no published systematic reviews that summarize and critique the literature.

OBJECTIVE: To describe the behavioral characteristics of children with PAE and/or FASD, assessed using the Achenbach System of Empirically Based Assessments (ASEBA) for school-aged children with parent, teacher, and youth (self-report) forms.

DATA SOURCES: Electronic literature databases, reference lists, hand-searches.

STUDY SELECTION: peer-reviewed observational studies.

DATA EXTRACTION: Study appraisal and data extraction were undertaken by 2 independent assessors. Meta-analyses were performed for parent-rated Internalizing, Externalizing, and Total problems scales. All other ASEBA scales were summarized qualitatively.

RESULTS: Included were 23 articles; 16 were used in meta-analyses. Pooled results showed higher Total (mean difference 12.1, 95% confidence interval [95% CI] 7.7–16.5), Internalizing (6.3, 95% CI 3.1–9.5), and Externalizing problems scores (12.5, 95% CI 7.9–17.0) in FASD than No FASD; and greater odds of scoring in the "Clinical" range in FASD. Pooled results demonstrated higher problem scores in children with PAE ($P > .05$). Qualitative summaries of other scales from parents, teachers, and self-report show poorer behavior ratings in children with FASD and PAE on composite Problem and Competence scores and many Syndrome subscales.

LIMITATIONS: Findings were restricted to behaviors assessed using the ASEBA. The published literature was limited, often with only 1 study reporting on a particular scale.

CONCLUSIONS: Meta-analysis reveals that FASD and PAE are associated with problematic behavior in many, but not all domains. This clearly affects families, and should be considered in clinical practice by providers.

[a]Discipline of Paediatrics and Child Health, Sydney Medical School, The University of Sydney, The Children's Hospital at Westmead, Westmead, New South Wales, Australia; [b]The George Institute for Global Health, and [c]Poche Centre for Indigenous Health, Sydney Medical School, The University of Sydney, Sydney, New South Wales, Australia; [d]Physiotherapy Department, Royal North Shore Hospital, St Leonards, New South Wales, Australia; [e]Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle Children's Research Institute, Seattle, Washington; [f]Departamento de Fisioterapia, Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, Presidente Prudente, Sao Paulo, Brazil; [g]The Sydney Children's Hospital Networks (Westmead), Westmead, New South Wales, Australia; [h]The Australian Paediatric Surveillance Unit, Kids' Research Institute, Westmead, New South Wales, Australia.

Dr Tsang conceptualized and designed the study, undertook the electronic and manual searches of the literature, data extraction, analysis, and appraisal of the articles, designed and created the data collection forms, coordinated the data collection, and drafted the initial manuscript; Ms Lucas contributed to the study design and to the design of the data collection forms, was the second assessor for this review, which involved independently checking the selection of articles for the review, data extracted, and study appraisal, and reviewed and revised, the final manuscript; Prof Carmichael Olson assisted in original design of data review documents, reviewed and critically revised the manuscript; Dr Pinto provided guidance on meta-analysis procedures, assisted in discussions regarding eligibility of

To cite: Tsang TW, Lucas BR, Carmichael Olson H, et al. Prenatal Alcohol Exposure, FASD, and Child Behavior: A Meta-analysis. *Pediatrics*. 2016;137(3):e20152542

Behavior impairments (predominantly in attention, disruptive behavior and conduct disorders, academic performance, and social judgment) have been documented in children with fetal alcohol spectrum disorders (FASDs) and significant prenatal alcohol exposure (PAE).[1,2] Secondary conditions found with high frequency among individuals with FASD and PAE include mental health problems, psychiatric illness,[3,4] school difficulties, trouble with the law, placement in confined settings (eg, psychiatric hospitals), inappropriate sexual behavior, and substance abuse.[5]

The importance of screening for and assessing behavior problems in the field of FASD is well-recognized[6,7] and used to guide clinical practice. The psychometrically sound Achenbach System of Empirically Based Assessment (ASEBA)[8] has been the most commonly used system of behavior assessment in children and adolescents with FASD since the mid-1990s. ASEBA School-Age Forms include the parent/caregiver-rated Child Behavior Checklist (CBCL), the Teacher Report Form (TRF), and the Youth Self-Report (YSR).[8] The School-Age forms assess a broad range of behaviors, including competencies, problems, and adaptive function. They also provide raw and, for clinical interpretation, standardized scores. We aimed to conduct a systematic review of the literature, with meta-analysis if appropriate, on behavioral ratings from multiple informants (parents/caregivers, teachers, self-report) in children with FASD and/or PAE, who were assessed by using the ASEBA School-Age Forms. To our knowledge, this is the first review of this type. We hypothesized that behavior ratings would be poorer in children with PAE and/or FASD than without.

TABLE 1 Eligibility Criteria

Inclusion criteria:
Design:
• English language
• Observational studies (cohort, case-control, cross-sectional)
Participants:
• Humans
• Aged ≤18 y and assessed by using age-appropriate school-age forms valid for ages 4 to 18 y depending on the ASEBA version used
Exposure:
• Diagnosis of a FASD (based on internationally recognized criteria); or
• Specified PAE (where PAE results could be separated; and excluding other exposures unless controlled for). PAE was not assumed in cohorts of children of alcoholic parents where PAE was not specifically measured.
Outcomes:
• Primary: T scores, raw scores, and/or proportion of participants within defined clinical ranges (*n*, %, or ORs) for scales: Total competence, Total problems, Internalizing problems, and/or Externalizing problems, assessed by using the ASEBA School-Age Forms (ie, Child Behavior Checklist, Teacher Report Form, and/or Youth Self-Report). Percentile scores will not be analyzed but will be described.
• Secondary: Raw scores, and/or proportion within defined clinical ranges (*n*, %, or ORs) for other ("syndrome," narrowband, DSM-oriented, and 2007) scales, assessed by using the ASEBA School-Age Forms. T scores or percentile scores will be described but not analyzed. Proportion of critical items endorsed will also be recorded if reported.
• Quantitative size effect including either an SE, SD, or 95% CI had to be reported.
Exclusion criteria:
Design:
• Articles not available in the English language
• Case studies, review articles, book chapters, abstracts, dissertations, news articles, clinical trials/intervention studies
Exposure:
• Studies that looked at the effect of exposures other than alcohol as a primary exposure of interest, where PAE results could not be separated.
• Studies that included conditions comorbid with diagnosis of a FASD, which may affect the primary outcomes of interest (unless they were controlled for).
Outcomes:
• Studies that reported data other than T scores, raw scores, or proportion (*n*, %, or ORs) within defined clinical ranges for individual ASEBA scales or forms.
• Articles that only reported data graphically
• Articles that combined scores from different ASEBA forms (eg, combined scores from the parent- and teacher-rated forms).
• Scores obtained from ASEBA forms other than the school-age forms (eg, preschool forms).

DSM, *Diagnostic and Statistical Manual of Mental Disorders*.

METHODS

Protocol Registration

This review was registered with PROSPERO (registration: CRD42014007040; Web link: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014007040#.VOJ7A8lI_MA), and conducted according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses recommendations.[9]

Eligibility Criteria

Selection criteria are summarized in Table 1. If a research team reported the same outcomes from the same cohort in multiple articles, only the article with the larger sample size was included to prevent duplicate inclusion of results.

Key Outcomes

The primary variables of interest were the scores for Internalizing problems (comprising the syndrome subscales: Anxious/depressed, Withdrawn/depressed, and Somatic complaints), Externalizing problems (comprising Rule-breaking and Aggressive behavior syndrome subscales), Total problems (sum

TABLE 2 Critical Appraisal Criteria

Criteria	Description
1. Having a defined sample	Defined eligibility and exclusion criteria for their sample; and time period (dates) and location(s) of recruitment and assessment. If details were reported in a previous article, those articles were retrieved and inspected for this information.
2. Having a representative sample	Articles were considered representative and given a score of 1 ("Yes") if they included cohorts recruited from the general population, or from multisite studies and/or large databases of consecutively recruited samples. Single-site clinical studies were given a score of 0 ("No").
3. Outcome rater blinding	Study staff were considered blinded if they were unaware of a FASD diagnosis, previous ASEBA scores, and/or presence/absence of PAE during administration of the ASEBA (whichever was applicable). Articles were scored as 0 if blinding was not mentioned and if it was unclear whether the ASEBA was self-administered.
4. Reporting of relevant ASEBA data	T scores, raw scores, and/or proportion (*n*, %, or OR) within defined clinical ranges reported for the primary outcome measures (Total competence, Total problems, Internalizing, and/or Externalizing scores); and/or for the narrowband measures, DSM-oriented scales, and 2007 scales.
5. Adequate sample size	Power calculation provided.
6. Statistical adjustment for prenatal exposure to other substances	Statistical adjustment for prenatal exposure to other substances in the analyses used to generate the ASEBA results extracted for this review. Articles were also scored 1 if they excluded participants with prenatal exposure to other substances of abuse (in their eligibility/exclusion criteria).
7. Statistical adjustment for other behavioral diagnoses	Statistical adjustment for presence of other psychiatric/behavioral diagnoses (in the child) in analyses used to generate ASEBA results extracted for this review, including by comparison with contrast groups with no teratogenic causes. Articles were also scored 1 if they excluded participants with other psychiatric/behavioral diagnoses (in eligibility/exclusion criteria).

Each of the above criteria were given a score of 0 ("no" or "unsure") or 1 ("yes"), and the scores for each of the 7 domains summed to give a total score out of 7 for each article.
DSM, *Diagnostic and Statistical Manual of Mental Disorders*.

of Internalizing and Externalizing problems scales; and syndrome subscales: Social, Thought, Attention, and Other problems), and Total competence (sum of Activities, Social, and School competence subscales).[8] All other scale scores and Critical items were considered secondary outcomes. Data extracted were restricted to T scores, raw scores, and/or proportions scoring within defined clinical ranges (Table 1).

Information Sources

The search was conducted by using electronic databases, and reference lists from eligible articles were inspected and hand-searched. Electronic databases used were Medline (Ovid; 1946 to February week 3, 2015), PsycINFO (1806 to February week 3, 2015), Maternity and Infant Care (1971 to January 2015), Embase (1974 to March 10, 2014), AMED (1985 to February 2015), CINAHL Plus (EBSCO Host), and EBM Reviews—Cochrane Database of Systematic Reviews (2005 to January 2015). The electronic literature search was conducted on March 11, 2014 (updated on February 17, 2015) and manual searches concluded in March 2015.

Search Strategy

Our search strategy was formulated with the assistance of a medical librarian. Search terms used for all databases were "fetal alcohol," "prenatal alcohol exposure," "prenatal alcohol," "ASEBA," "child behavior checklist," "Achenbach," "teacher report form," "youth self-report," "child behavior," and "psychosocial." Due to changes in FASD diagnostic criteria over time, we selected broad search terms to maximize our yield. No limits were assigned for the electronic searches. As an example, our Medline (OVID) search strategy is provided in Supplemental Appendix 1.

Study Selection

Screening and study selection was a 2-stage process. Inspection of article titles, abstracts, and article type was conducted during the first stage (by TWT). The second stage involved review of full-text articles identified as potentially eligible (Table 1) by 2 independent reviewers (TWT and BRL), with disagreements resolved with EJE.

Data Extraction Process

Data extraction/interpretation was undertaken by TWT, independently checked by BRL, and disagreements resolved by EJE and RZP. Data were collated into an electronic database. Corresponding authors were contacted via e-mail for additional information when necessary.

Risk of Bias in Individual Studies

A 7-point critical appraisal tool was created for use in this systematic review, based on Sanderson et al.[10] Eligible articles were rated according to criteria described in Table 2.

Data Synthesis and Analysis

Types of data extracted included T scores, raw scores, proportions

(%), or odds ratios (OR) for scoring within defined clinical ranges. These different types of data could not be combined in subsequent meta-analyses or forest plots, so separate analyses were undertaken for each data type (for primary outcomes) if reported by at least 2 separate studies.[11] T or raw score data were used to calculate the difference in means between "cases" and controls and 95% confidence intervals (95% CIs). Proportions within clinical ranges were used to calculate the OR and 95% CI where applicable. Because this review only included the School-Age Forms, no standardization of age and gender-normed T-score data were undertaken. T scores are more directly related to scores obtained from ASEBA forms than other standardized scores (eg, *z* scores). Also, T scores are typically used in clinical practice. Scores in the "Clinical" or "Borderline" range on ASEBA scales, defined by T scores, generally indicate need for therapeutic intervention.

Separate analyses were conducted based on the presence or absence of (1) an FASD diagnosis (including "mixed" groups containing children with PAE in which 100% of them had an FASD diagnosis), and (2) PAE; according to how results were reported in the articles. Data were grouped according to age category (mean age: <5 years, 5–10 years, and >10 years), outcome (ASEBA score), and PAE level where applicable. Effects of different PAE levels on primary outcomes were explored after attempted categorization of PAE levels, using definitions suggested by O'Leary et al,[12] and also using the classification of PAE specified in each study (low/moderate/high).

Where data were stratified and presented in multiple groups, data from PAE/FASD and control groups were extracted according to (1) consistency with groups of interest for this review, (2) largest total sample size, and (3) longest duration of PAE. If individual group sample sizes were not reported or could not be calculated from results provided, information was sought from corresponding author(s). In accordance with Cochrane guidelines, if a study had no control/comparison group, data from another study's comparison group (with comparable sample characteristics) was borrowed and the comparison group sample size divided by the number of studies by using the same comparison group.[11]

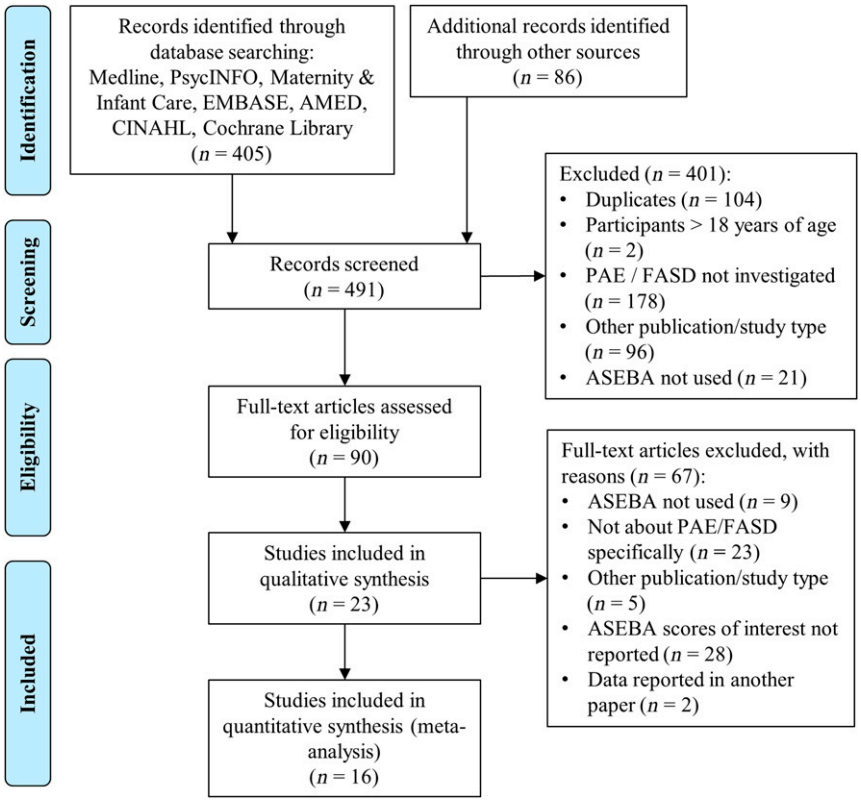

FIGURE 1
Literature flowchart.

Comprehensive Meta Analysis (version 2.2.064; Biostat, Inc, Englewood, NJ) software was used for meta-analyses and forest plots, by using random effects models. A $P < .05$ denoted statistical significance. Heterogeneity of studies in meta-analyses was assessed by using the I^2 statistic, where I^2 <40% suggested homogeneity within pooled data.[11] Funnel plot asymmetry was examined if at least 10 studies were included in the meta-analysis.[11] Data not included in meta-analyses are presented descriptively in the Supplemental Appendices.

RESULTS

Literature Search

Database and manual searches yielded 491 articles (Fig 1). After screening and review, 23 articles were included in qualitative or quantitative analysis, with between 2 and 8 articles (total: 16) included in various meta-analyses.

Study Characteristics

Tables 3 and 4 show the characteristics of included studies, according to presence of FASD (10 studies)[13–21] or PAE (14 studies), respectively[22–34]: one study reported data for both PAE and FASD.[35] Most

were from North America (n = 15; 65.2%), followed by Australia (n = 4) and Canada (n = 2), with 1 study each from Finland and Taiwan. Four North American studies assessed predominantly African American (83%–94%)[24,30] or exclusively African American cohorts.[32,33] In addition to PAE/FASD groups, 3 studies included subgroups with attention-deficit/hyperactivity disorder (ADHD),[15,26,34] and 1 study included a subgroup with prenatal cocaine exposure (Supplemental Appendix 2).[33]

Additional data (eg, individual group sample size data) were requested for 3 articles,[17,23,29] but authors were unable to access archives (n = 2) or did not respond (n = 1). Sample sizes for Robinson et al[31] were estimated based on number and corresponding percentage of children scoring within "Clinical" ranges for the primary ASEBA outcomes.

Of 23 studies included, 12 (52.2%) were retrospective cohort studies. Other study types included prospective cohort studies (n = 5), cross-sectional studies (n = 2), follow-up analysis studies (n = 2), and historical prospective cohort studies (n = 2; Tables 3 and 4) . In 13 studies, participants were accessed from common sources/databases: 5 sources/databases provided cases for 2 studies each and a sixth provided cases for 3 studies, although there was no individual subject overlap for 2 of these 3 studies. All but 1 study used the CBCL form (95.7%), 4 studies reported TRF data, and 2 used the YSR. One reported both CBCL and YSR results and 4 reported both CBCL and TRF results. Tables 3 and 4 detail the ASEBA version(s) used in each study. Due to the small number of studies reporting consistent TRF and YSR results, meta-analysis was restricted to studies that used the CBCL and reported our primary outcomes of interest: Total problems, Internalizing problems, and Externalizing problems scores. The Total competence score from the CBCL was not used in meta-analysis, as it was reported in only 3 studies (2 FASD studies, 1 PAE study). Detailed data are reported in the Supplemental Appendices.

Results for a total of 21 different ASEBA scales were reported in studies of FASD and 44 in studies of PAE. The number of FASD studies reporting data for any 1 scale ranged from 1 to 8 (median of 3 studies per scale). In PAE studies, between 1 and 9 studies provided data for any 1 scale but 28 scales (63.6% of all scales reported in PAE studies) were each reported by only 1 study. Data on children with FASD or PAE were reported for 24 CBCL scales, 19 TRF scales, and 3 YSR scales (total of 46 scales).

Parent-Rated Behavior: Primary Outcomes (Meta-analysis)

Presence of an FASD Diagnosis: CBCL Total, Internalizing and Externalizing Problems, Total Competence

FASD diagnostic criteria used in each study are included in Table 5. For all FASD studies, diagnostic assessments were made by experienced dysmorphologists, specially trained professional interdisciplinary teams, or trained clinicians. In FASD studies, very few CBCLs were completed by the biological parent (3.0%–22.7%).

Ten studies reported T scores for Total problems, or Internalizing or Externalizing problems in children with FASD (Fig 2). One study was excluded from the Total problems component of this meta-analysis, as control group sample size was not available.[17] Figure 2 indicates significantly higher Total problems scores (pooled mean difference [95% CI]: 12.1 [7.7–16.5]; P < .0001), Internalizing problems scores (pooled mean difference [95% CI]: 6.3 [3.1–9.5]; P < .0001), and Externalizing problems scores (pooled difference [95% CI]: 12.5 [7.9–17.0]; P < .0001) in children with a FASD diagnosis compared with those without.

Four FASD studies reported the number of children scoring within "Clinical" ranges for Total, Internalizing, or Externalizing problems,[13,14,17,20] 2 of these reporting data from a comparison group.[13,20] Compared with children without FASD, those with FASD were more likely to have Total (OR 34.0, 95% CI 2.6–450.8]), Internalizing (OR 10.0, 95% CI 1.3–77.6), or Externalizing problems scores (OR 18.2, 1.8–186.6) within the "Clinical" range, and less likely to have scores within the "Normal" range (Fig 3). They were also more likely to score in the "Borderline" range on the Total problems scale. There were no significant differences between groups in the ORs for scoring within the "Borderline" range for Internalizing and Externalizing problems (P > .05). Critical items important to clinical intervention (eg, Talks suicide, Attacks others), were not reported in any studies. There were insufficient studies to compare age groups.

Only 2 studies reported Total competence outcomes according to FASD diagnoses.[17,19] Both studies reported poorer (lower) parent-rated Total competence scores in children with a FASD diagnosis compared with those without, regardless of FASD diagnostic subgroup. Compared with children with similar IQs, mean Total competence scores among children with fetal alcohol syndrome (FAS) appeared lower in boys (FAS 31.3 [7.4]; No FAS IQ comparison 45.4 [9.6]) but not in girls (FAS 40.5 [12.5]; No FAS IQ comparison 42.7 [7.3]; P not calculated).[17] Olson et al[17] reported that 5 of 9 children with FAS scored within the "Clinical" range, but did not report the finding in children without FAS (Table 5).

TABLE 3 Systematic Review: Individual Study Characteristics, FASD Studies Only ($n = 10$)

Reference	Study Design	Sample Source	ASEBA Version Used[a]	*N*; Age (Mean [SD]; or Range)	ASEBA Assessor/Informant
Ernhart et al 1995[21]	Prospective cohort	All children born 1 Jan 1981 to 31 Mar 1982, to mothers screened for alcoholism in 1978 to 1981 at Cleveland Metropolitan General Hospital (USA).	CBCL (4–16 y, 1981)	*N*: FAS: 5; No FAS: 221 Age range: 6–7 y	NR, CBCL orally administered to primary caregivers.
Olson et al 1998[17]	Retrospective cohort, matched control/ comparison groups	FAS: Cases from research patient list from the Fetal Alcohol Follow-up Study (Seattle, USA). Control/comparison: The Seattle Longitudinal Prospective Study on Alcohol and Pregnancy.	CBCL (4–18 y, 1983, 1991, 1993)	*N*: FAS: 9; Control (light/no PAE): 145; IQ comparison: 40 Age range: 4–16 y	Self-administered (primary caregiver)
Mattson and Riley 2000[35b]	Retrospective cohort, matched control group	FAS/heavy PAE: Children in an ongoing research project at the Center for Behavioral Teratology, San Diego State Univ. (USA). Control: Self-referred from community outreach and advertizing (San Diego, USA).	CBCL (4–18 y, 1991)	*N*: FAS: 35; Heavy PAE (no FAS): 20 Heavy PAE (±FAS): 33; Control: 33 Age: FAS/heavy PAE: 9.2 (3.5) y (range: 4.0–15.8 y); Control: 8.4 (3.7) y (range: 4.1–16.5 y)	Self-administered (primary caregiver)
Paley et al 2006[18]	Retrospective cohort	A study on a social skills intervention which recruited children through a large medical center in the West Coast of the USA, through community mental health clinics, private practitioners, local schools, and the community.	CBCL (6–18 y, 2001) TRF (6–18 y, 2001)	*N*: 100 Age: 8.6 (1.6) y (Range: 6.0–11.6 y)	Self-administered (primary caregiver)
Franklin et al 2008[14]	Retrospective cohort	Washington State FAS DPN clinical database (Washington, USA).	CBCL (school age forms; 1991, 2001).	*N*: 44 Age at diagnosis: 5 to 6.9 y: $n = 14$ 7 to 8.9 y: $n = 17$ 9 to 10.9 y: $n = 13$	NR, primary caregiver
Greenbaum et al 2009[15]	Retrospective cohort, control/ comparison groups	FASD: Data records of the Motherisk Clinic (FASD diagnostic service) at The Hospital for Sick Children, Toronto. Control: Advertisements in the community or the Metropolitan Toronto District School Board.	CBCL (4–18 y, 2001) TRF (4–18 y, 2001)	*N*: FASD: 33; Control: 34 Age: FASD: 9.23 (NR); Control: 8.87 (NR)	Self-administered (primary caregiver, teachers)
Astley et al 2010[19]	Retrospective cohort, comparison groups	The Washington State FAS DPN electronic clinical/research database of all residents (of all ages) evaluated for FASD in 1993–2005.	CBCL (6–18 y, 2001)	*N*: [For ASEBA data only; age 6–18 y]: FAS/PFAS: 51; Normal CNS/PAE: 25 Age: 9.9 (6.2) y (range: 7 d – 50.8 y)	Pediatrician and/or psychologist
Fagerlund et al 2011[13]	Retrospective cohort, matched control group	FASD: Children diagnosed and born in 1984–1996, from the Hospital for Children and Adolescents, Univ. of Helsinki, and from a prospective follow-up study in Helsinki (Finland). Control: Finnish national population registry (random sampling).	CBCL (6–18 y, 2001 – Finnish and Swedish versions)	*N*: FASD: 73; Control: 40 Age: FASD: 13.4 (3.8) y (range: 8–21 y); Control: NR	NR, primary caregiver
Jirikowic et al 2012[16]	Retrospective cohort	Children who were enrolled in a RCT of the Families Moving Forward Program intervention, recruited through the Univ. Washington FAS DPN clinical database (USA).	CBCL (preschool, 2000; school-age, 2001)	*N*: 52 Age: 8.53 (2.03) y (range: 5–12 y)	Trained psychometrists

TABLE 3 Continued

Reference	Study Design	Sample Source	ASEBA Version Used[a]	*N*; Age (Mean [SD]; or Range)	ASEBA Assessor/Informant
Stevens et al 2013[20]	Retrospective cohort, comparison group	FASD and PAE (no FASD): Data records of patients attending the Motherisk Clinic (FASD diagnostic service) at The Hospital for Sick Children, Toronto, between 2005 and 2009.	CBCL (4–18 y, 2001)	*N*: [For CBCL]: FASD: 98; PAE (no FASD): 50 [For TRF]: FASD: 95; PAE (no FASD): 45 Age: FASD: 10.3 (3.6); PAE (no FASD): 8.9 (3.4) [$P < .01$]	Self-administered (primary caregiver, teachers)
			TRF (4–18 y, 2001)		

Participant group data (N, age) are presented only for FASD and No FASD groups. Data for all subgroups included within each study can be found in Appendix 2.
NR, not reported; Univ., university; DPN, diagnostic and prevention network; M, male; F, female; RCT, randomized controlled trial. All rows except Ernhart et al indicate clinical studies as opposed to population-based cohort studies.
[a] ASEBA forms listed in the table reflect either what was specified by the authors of the articles and/or the reference(s) they referred to when providing details of their ASEBA assessment.
[b] Mattson and Riley reported ASEBA results for both (1) FAS versus PAE (without FAS); and (2) PAE (with and without FAS) versus Controls.

Presence of PAE: CBCL Total, Internalizing and Externalizing Problems; Total Competence

Prenatal Alcohol Exposure: T Scores

Seven studies reported raw (n = 2), percentile (n = 1), or T scores (n = 4) for primary outcomes. In these studies, the CBCL was usually completed by a biological parent (range 82.8%–100%). Of 4 studies reporting T scores, 1 did not report or provide subject numbers for PAE groups and could not be included in the forest plot.[23] Of 3 studies reporting raw or percentile scores,[30,32,33] only 2 provided SDs[30,33] and only 1 had control group data.[33] Therefore, there were inadequate raw score data for meta-analysis, so a forest plot was generated for the 3 studies reporting both T scores and subject numbers. Due to limited data available for this analysis, PAE was stratified as either present (for any duration) or absent, rather than indicated by low/moderate/high levels. There was a trend for higher problems scores (poorer outcomes) among children with PAE than without, although the forest plot reveals heterogeneity in results of the 3 studies included (Fig 4). The largest difference in problem scores between PAE and control groups was observed in 2 studies in which children in the PAE group were selected based on high levels of PAE (>4 alcoholic drinks at least once per week, or ≥13 drinks per week throughout pregnancy).[26,34] Levels of PAE were not defined or reported in Brown et al.[24] Pooled effects for the 3 summary problem scales showed no statistically significant differences between groups ($P > .05$; Fig 4). Age group comparisons were not possible.

Prenatal Alcohol Exposure: ORs

Of 23 studies included in this review, 14 stratified groups based on presence or absence of PAE or level of PAE; however, levels of PAE were inconsistent across different

TABLE 4 Systematic Review: Individual Study Characteristics, PAE Studies Only (n = 14)

Reference	Study Design	Sample Source	ASEBA Version Used[a]	N; Age (Mean (SD); or Range)	ASEBA Assessor/Informant
Brown et al 1991[24]	Prospective cohort	A cohort enrolled in a PAE study which recruited mothers and children born in 1980–1983 at a large univ. teaching hospital (Atlanta, USA). The cohort was predominantly African-American and of low socioeconomic status.	CBCL (1983) TRF (1983)	N: Never drank: 19; Stopped drinking: 15; Continued to drink: 20 Age: 5.8 (NR) (eligible range: 5–8 y)	Masters-level psychology graduate students
Mattson and Riley 2000[35b]	Retrospective cohort, matched control group	FAS/heavy PAE: Children in an ongoing research project at the Center for Behavioral Teratology, San Diego State Univ. (USA). Control: Self-referred from community outreach and advertizing (San Diego, USA).	CBCL (4–18 y, 1991)	N: FAS: 35; Heavy PAE (no FAS): 20 Heavy PAE (±FAS): 33; Control: 33 Age: FAS/heavy PAE: 9.2 (3.5) y (range: 4.0–15.8 y); Control: 8.4 (3.7) y (range: 4.1–16.5 y)	Self-administered (primary caregiver)
Sood et al 2001[32]	Historical prospective cohort study	Children of mothers screened for alcohol and drug use at the Fetal Alcohol Research Center (urban univ.-based maternity clinic, Detroit, USA), who were delivered in Sep 1989-Aug 1991 were located at 6–7 y. Data were collected for African-American women only.	CBCL (4–18 y, 1991)	N: No PAE: 117; Low PAE: 323; Moderate/heavy PAE: 66 Age: Mean: 6.9 (NR) y (eligible range: 6–7 y)	Trained research assistants
Lee et al 2004[27]	Retrospective cohort, matched control group	Children enrolled in an ongoing research project at the Center for Behavioral Teratology, San Diego State Univ (USA). Control: From community outreach/advertizing.	CBCL (4–18 y, 1991)	N: FAS/heavy PAE: 30; Control: 30 Age: FAS/heavy PAE: 11.4(2.3) y (range: 9.0–16.8 y); Control: 11.6 (2.1) y (range: 9.0–16.9 y)	Self-administered (primary caregiver)
Paley et al 2005[30]	Cross-sectional study	A major medical center and affiliated health clinics in the surrounding community using flyers (West Coast of the USA). Cohort was predominantly African-American and of low socioeconomic status.	CBCL (1991)	N: PAE (moderate-heavy): 20; PAE (abstinent-light): 22 Age: 4.8 (0.8) y	Self-administered (biological mother) or a research assistant (if the mother had difficulty reading)
Sood et al 2005[33]	Historical prospective cohort study	Children of mothers screened for alcohol and drug use in antenatal clinics at Wayne State Univ. (Detroit, USA), who were delivered in Sep 1989-Aug 1991 were located at 6–7 y. Data were collected for African-American women only.	CBCL (4–18 y, 1991)	N: Boys: No PAE+No cocaine: 49; PAE+No cocaine: 103 Girls: No PAE+No cocaine: 45; PAE+No cocaine: 91 Age: 6–7 y follow-up (mean / range NR)	Trained research assistants
Bada et al 2007[23]	Prospective cohort, control group	Maternal Lifestyle Study, which enrolled children born in 1993–1995 at hospitals: Brown Univ. (Providence, RI); Univ. Miami (Florida); Univ. Tennessee; and Wayne State Univ. (Detroit, USA).	CBCL(2–3y, 1992; 14–18 y, 1991)	N: No PAE: NR; Some PAE: NR; High PAE: NR Age: 3, 5, and 7 y follow-up (mean / range NR)	Trained research interviewer

TABLE 4 Continued

Reference	Study Design	Sample Source	ASEBA Version Used[a]	N; Age (Mean (SD); or Range)	ASEBA Assessor/Informant
O'Callaghan et al 2007[28]	Follow-up analysis (from prospective cohort study)	Study of Pregnancy which enrolled children born in 1981–1984 at the Mater Misericordiae Hospital (Brisbane, Australia).	CBCL (1991) YSR (1991)	N: PAE (early preg): Nil: 2528; >0 to <0.5 glass/d: 2420; 0.5 to <1 glass/d: 134; ≥ 1 glass/d: 57 PAE (late preg): Nil: 3258; >0 to <0.5 glass/d: 1624; 0.5 to <1 glass/d: 170; ≥ 1 glass/d: 82 Bingeing: Never: 4127; < 0.5 time: 884; ≥ 0.5 time: 128 Age: 13.9 (0.3) y (range: 12.5–15.5 y)	Self-administered (mother; adolescent)
Alati et al 2009[22]	Follow-up analysis (from prospective cohort study)	Study of Pregnancy which enrolled children born in 1981–1984 at the Mater Misericordiae Hospital (Brisbane, Australia).	YSR (4–18 y, 1991)	N: No PAE: 2530; 0–0.5 glass/d: 2536; ≥ 1 glass/d: 66 Age: 14 y (mean / range NR)	NR, adolescent
Chiu et al 2009[25]	Cross-sectional study	9 kindergartens and 3 elementary schools in 1998 (Taipei, Taiwan).	CBCL (4–18 y, 1990 – Chinese version)	N: PAE: 11; No PAE: 50–107 Age: 7.4 (1.5) y (eligible range: 4–9 y)	NR, mother
O'Leary et al 2010[29]	Prospective cohort study	RASCALS study, involving children born in 1995–1996 in Western Australia.	CBCL (2–3 y, 1991)[c]	N: PAE- first trimester: None: NR; Low: NR; Moderate: NR; Heavy: NR PAE- Late pregnancy: None: NR; Low: NR; Moderate: NR; Heavy: NR Age: 2, 5, and 8 y follow-up (mean / range NR)	NR, biological mother
Robinson et al 2010[31]	Prospective cohort study	Recruited in May 1989-Nov 1991 at public anti-natal clinics in King Edward Memorial Hospital and nearby private clinics (Perth, WA).	CBCL (2–3 y at 2y follow-up, 1987; 4–18 y at subsequent follow-ups, 1991)	N: PAE 18 wks: Occasional: 539; Light: 419 Moderate: 60; Heavy: 42 PAE 34 wks: Occasional: 427; Light: 310; Moderate: 38; Heavy: 16 Age: Eligible range: 2–14 y	NR, biological mother
Graham et al 2013[26]	Retrospective cohort, matched control/ comparison groups	CIFASD data from sites: Center for Behavioral Teratology (San Diego State Univ.); Emory Univ.; 7 Northern Plains communities; Univ. New Mexico Center on Alcoholism, Substance Abuse and Addictions; and Univ. California, Los Angeles Fetal Alcohol and Related Disorders Clinic.	CBCL (4–18 y, 1991)	N: FAS/heavy PAE (no ADHD): 35; Control: 102 Age: FAS/heavy PAE (no ADHD): 12.6 (2.6) y; Control: 12.5 (2.6) y	NR, primary caregivers

TABLE 4 Continued

Reference	Study Design	Sample Source	ASEBA Version Used[a]	N; Age (Mean (SD); or Range)	ASEBA Assessor/Informant
Ware et al 2013[34]	Retrospective cohort, control/ comparison groups	CIFASD data from sites: Center for Behavioral Teratology (San Diego State Univ.); The Fetal Alcohol and Drug Exposure Clinic (Emory Univ.); 7 communities (throughout North Dakota, South Dakota, Montana, and Montana); Center on Alcoholism, Substance Abuse and Addictions (Univ. New Mexico); and Fetal Alcohol and Related Disorders Clinic (Univ. California).	CBCL (4–18 y, 1991)	N: PAE+No ADHD: 52; Control: 133 Age: 12.28 (2.46) y	Trained examiner

Participant group data (N, age) are presented only for PAE and No PAE groups. Data for all subgroups included within each study can be found in Appendix 2.
NR, not reported; Univ., university; CIFASD, Collaborative initiative on fetal alcohol spectrum disorders; YSR, youth self-report; RASCALS study, Randomly ascertained sample of children born in Australia's largest state study. Mattson and Riley, Lee et al, Graham et al, and Ware et al are clinical studies as opposed to population-based cohort studies.
[a] ASEBA forms listed in the table reflect either what was specified by the authors of the articles and/or the reference(s) they referred to when providing details of their ASEBA assessment.
[b] Mattson and Riley reported ASEBA results for both (1) FAS versus PAE (without FAS); and (2) PAE (with and without FAS) versus Controls.
[c] For school-aged children in this study, contact with the corresponding author was attempted to ask them if the school-age form was also used. No response was received.

studies. Only 3 studies reported OR for our primary outcomes.[29,31,32] Exploratory forest plots were created, incorporating the 3 primary "problems" scores (Total, Externalizing, and Internalizing problems) to examine whether odds of having problematic behavior ratings were higher after PAE. These primary outcomes, grouped by standardized PAE categories based on O'Leary et al,[12] indicated no increased odds for poorer behavior ratings with higher PAE levels (data not shown). However, when data were grouped by PAE categories designated by study authors, the odds of poorer behavior ratings tended to fall in favor of moderate/high PAE compared with no PAE (Fig 5), although results were not statistically significant ($P > .05$). There was no significant increase in odds of behavioral problems with low/light/occasional PAE compared with no PAE (Fig 5). All studies included children aged between 5 and 10 years.

CBCL Total competence score according to PAE was reported in only 1 study.[24] Children with no PAE had significantly higher (favorable) Total competence ratings than children with PAE, even if drinking had stopped during the second trimester (T scores: No PAE 16.6 [2.8]; Stopped drinking 13.8 [2.7]; Continued drinking: 13.5 [3.4]; $P < .008$).

A detailed table displaying the primary outcomes for all subgroups in included studies is provided in Supplemental Appendix 3.

Secondary Outcomes: Other Forms: Other CBCL Scores, TRF, YSR

Secondary outcomes included (1) other ASEBA School-Age Form scales and Critical items reported in eligible studies (rated by parents/caregivers, teachers, and/or youth), and (2) primary outcome scales rated by teachers and youth. Data were summarized qualitatively and highlights reported below. Detailed summaries of secondary outcomes based on presence of FASD diagnosis or PAE can be found in Supplemental Appendices 4 (summary) and 5 (table of results).

Presence of an FASD Diagnosis

Parent and teacher ratings of Total, Externalizing, and Internalizing problems were higher among children with FASD than without, even when compared with children with PAE but not FASD. CBCL scales showed the most common problems among children with FASD were the following: Thought problems, Rule-breaking behavior/Delinquency, Aggressive behavior, Attention problems, and Social problems. Children with FASD generally received higher Externalizing behavior scores than children with ADHD. However, parent and teacher ratings were inconsistent for Internalizing and Externalizing problems when comparing groups with FASD with ADHD (Supplemental Appendices 3, 4, and 5).

Total competence (and School competence subscale) scores were poorer among children with FASD than without (only parent ratings available).[17,19] Scores on other competence subscales were mixed: in 1 study, children with FASD had better Social competence scores than those with PAE (without FASD),[35] although scores in both groups were generally in the "Clinical" range. The Activities competence score in FASD was similar to comparison peers and within the "Normal" range in most studies. Somatic complaints,[14,19,35] Anxiety/depression,[14,19,35] Withdrawn/depressed,[14,19,35] Academic performance (1 study),[18] and Adaptive functioning (1 study)[18] subscale scores were similar in FASD and comparison groups, and usually within the "Normal" range (Supplemental Appendices 4 and 5).

In 1 large study, CBCL ratings showed no differences in Total, Externalizing,

TABLE 5 Results for Parent/Caregiver-rated Total, Internalizing, and Externalizing Problems Composite Scores Used in Meta-analyses Only (*n* = 16 studies)

Reference	Alcohol Exposure / FASD Guideline Used[a]	Comparison Group	Alcohol-Exposed Group Scores, Mean (SD) or OR (CI)	Comparison Group Scores	Covariates Controlled For[b]
Brown et al 1991[24]	Stopped drinking in Trimester 2: NR drinks/week (n = 23) Continued drinking: NR drinks/week (n = 23)	No PAE: 0 drinks/week (*n* = 22)	Internalizing problems (T score): Stopped drinking: 54.86 (10.24) Continued drinking: 57.15 (8.41) Externalizing problems (T score): Stopped drinking: 57.33 (8.88) Continued drinking: 60.75 (7.65) Total problems (T score): Stopped drinking: 57.48 (9.91) Continued drinking: 61.05 (8.33)	Internalizing problems (T score): No PAE: 55.8 (9.86) [*P* not significant] Externalizing problems (T score): No PAE: 58.90 (10.32) [*P* not significant] Total problems (T score): No PAE: 59.00 (9.52) [*P* not significant]	Caregiver's current alcohol consumption
Ernhart et al 1995[21]	FAS: Criteria from the Fetal Alcohol Study Group of the Research Society on Alcoholism (*n* = 5)	No FAS (*n* = 221)	FAS (T score): Internalizing problems: 59.14 (11.35) Externalizing problems: 63.4 (9.77) Total problems: 63.28 (9.91)	No FAS (T score): Internalizing problems: 60.9 (9.1) Externalizing problems: 62.9 (9.4) Total problems: 63.3 (9.2)[*P* values NR]	Prenatal marijuana exposure; excluded those with prenatal narcotic exposure.
Olson et al 1998[17]	FAS: Gestalt method from IOM (n = 9 [boys: 3; girls: 6])		Total problems (% clinical [T score>63]): FAS (boys): 66.7% FAS (girls): 100%	NA	None
Mattson and Riley 2000[35]	PAE (±FAS): Growth retardation, CNS dysfunction, and facial dysmorphology; PAE definition NR (*n* = 33)	Control (*n* = 33)	PAE (±FAS) (T score): Internalizing problems: 58.90 (12.56) Externalizing problems: 66.50 (13.87) Total problems: 67.90 (10.61)	Control (T score): Internalizing problems: 51.10 (12.74); [*P* = .019] Externalizing problems: 53.60 (12.22); [*P* = .001] Total problems: 53.30 (12.79); [*P* < .001]	Group, gender, estimated SES, verbal IQ
Sood et al 2001[32]	Low PAE: >0 to <4.3 drinks/week (n = 323) Moderate/heavy PAE: ≥4.3 drinks/week(*n* = 66) OR calculated using PAE (present/absent)	No PAE: 0 drinks/week (*n* = 117)	PAE (Clinical range OR (95%CI)): Internalizing problems: 1.6 (0.9 to 3.1) Externalizing problems: 1.7 (1.0 to 3.2) Total problems: 1.8 (1.0 to 3.0)	NA	None
Paley et al 2006[18]	FAS: 4-digit diagnostic code; PAE defined as ≥8.51 drinks/week or ≥3.65 drinks/occasion (*n* = 100)	NA	FAS (T score): Internalizing problems: 63.47 (10.29) Externalizing problems: 68.04 (9.65)	NA	None

TABLE 5 Continued

Reference	Alcohol Exposure / FASD Guideline Used[a]	Comparison Group	Alcohol-Exposed Group Scores, Mean (SD) or OR (CI)	Comparison Group Scores	Covariates Controlled For[b]
Franklin et al 2008[14]	FASD: 4-digit diagnostic code (*n* = 44)	NA	FASD (T score): Internalizing problems: 66.90 (8.60) Normal (<60): 20.5% Borderline (60–63): 13.6% Clinical (>63): 65.9% Externalizing problems: 70.10 (9.20) Normal (<60): 15.9% Borderline (60–63): 9.1% Clinical (>63): 75.0% Total problems: 71.90 (7.50) Normal (<60): 9.1% Borderline (60–63): 4.5% Clinical (>63): 86.4%	NA	None
Greenbaum et al 2009[15]	FASD: Motherisk Program ARND diagnostic criteria checklist (*n* = 33)	Control (*n* = 34)	FASD (T score): Internalizing problems: 50.60 (22.40) Externalizing problems: 67.60 (22.98)	Control (T score): Internalizing problems: 53.4 (11.66); [p not significant] Externalizing problems: 47.30 (12.25); [P = .001]	SES, tobacco exposure, abuse, adoption/foster home history; ADHD
Astley 2010[19]	All groups had PAE confirmed at any level and were diagnosed with the 4-digit diagnostic code. FAS/PFAS: *n* = 37–51	Normal CNS/PAE: PAE at any level (*n* = 331–500)	FAS/PFAS (T score): Internalizing problems: 63.4 (10.1) Externalizing problems: 69.1 (9.9) Total problems: 71.4 (8.9)	Normal CNS/PAE (T score): Internalizing problems: 60.8 (14.1); [*P* = .014] Externalizing problems: 60.3 (13.2); [*P* < .0001] Total problems: 61.9 (12.7); [*P* < .0001]	None
O'Leary et al 2010[29]	PAE- Late preg: Low: < 5 drinks/week and ≤0.71–1.43 drinks/d (*n* = NR) Moderate: ≤ 5 drinks/week (*n* = NR) Heavy: ≥4.3–27.1 drinks/week (*n* = NR)	Abstinent: None: *n* = NR	Internalizing problems (Adjusted OR (95%CI) for T score ≥ 60): Low: 1.12 (0.8 to 1.56) Moderate: 1.34 (0.82 to 2.19) Heavy: 1.36 (0.5 to 3.71) Externalizing problems (Adjusted OR (95%CI) for T score ≥ 60): Low: 1.04 (0.77 to 1.4) Moderate: 1.27 (0.84 to 1.93) Heavy: 1.31 (0.62 to 2.79) Total problems (Adjusted OR (95%CI) for T score ≥ 60): Low: 0.97 (0.7 to 1.34) Moderate: 1.42 (0.9 to 2.23) Heavy: 1.39 (0.57 to 3.41)	NA	Maternal age, ethnicity, parity, marital status, income, smoking and illicit drug use (incl. tranquilizers, sleeping tablets) during preg, postnatal depression; anxiety, stress, family functioning, parenting scales, family tension due to alcohol misuse, age of completion of CBCL.

TABLE 5 Continued

Reference	Alcohol Exposure / FASD Guideline Used[a]	Comparison Group	Alcohol-Exposed Group Scores, Mean (SD) or OR (CI)	Comparison Group Scores	Covariates Controlled For[b]
Robinson et al 2010[31]	PAE 34 wk: Occasional: (*n* = 427) Light: (*n* = 310) Moderate: (*n* = 38) Heavy: (*n* = 16)	No PAE: 0 drinks/week: No PAE (34 wk): (*n* = 1579)	Internalizing problems (Adjusted OR (95%CI) for T score ≥ 60): Occasional: 1.06 (0.82 to 1.37) Light: 0.86 (0.61 to 1.2) Moderate: 0.49 (0.16 to 1.49) Heavy: 1.21 (0.49 to 2.95) Externalizing problems Adjusted OR (95%CI) for T score ≥ 60): Occasional: 1.22 (0.92 to 1.62) Light: 0.77 (0.56 to 1.05) Moderate: 0.53 (0.19 to 1.44) Heavy: 1.24 (0.40 to 3.86) Total problems (Adjusted OR (95%CI) for T score ≥ 60): Occasional: 1.24 (0.93 to 1.65) Light: 0.88 (0.63 to 1.23) Moderate: 0.48 (0.20 to 1.14) Heavy: 1.19 (0.30 to 4.68)	NA	Maternal age and education, family income, presence of biological father in family home, stressful events in preg, prenatal cigarette smoking, child's age at each follow-up
Fagerlund et al 2011[13]	FASD: Revised IOM diagnostic criteria (*n* = 73)	No FASD (*n* = 40)	FASD (T score): Internalizing problems: 53.92 (10.22) Normal (<60): 70.4% Borderline (60–63): 11.3% Clinical (>63): 18.3% Externalizing problems: 53.59 (10.77) Normal (<60): 73.2% Borderline (60–63): 12.7% Clinical (>63): 14.1% Total problems: 56.69 (8.89) Normal (<60): 57.7% Borderline (60–63): 19.7% Clinical (>63): 22.5%	No FASD (T score): Internalizing problems: 44.80 (8.08); [*P* < .0001] - Normal (<60): 92.5% - Borderline (60–63): 5.0% - Clinical (>63): 2.5% Externalizing problems: 43.75 (8.66); [*P* < .0001] - Normal (<60): 95.0% - Borderline (60–63): 5.0% - Clinical (>63): 0% Total problems: 42.10 (8.76); [*P* < .0001] - Normal (<60): 97.5% - Borderline (60–63): 2.5% - Clinical (>63): 0%	None
Jirikowic et al 2012[16]	FASD: 4-digit diagnostic code (*n* = 52)	NA	FASD (T score): Total problems: 70.90 (6.00)	NA	None
Graham et al 2013[26]	CIFASD Dysmorphology Core diagnostic criteria; >4 drinks/occasion ≥1/week, or >13 drinks/week: FAS/heavy PAE (no ADHD): *n* = 35	Control (*n* = 102)	FAS/heavy PAE (no ADHD) (T score): Internalizing problems: 55.91 (10.17) Externalizing problems: 55.86 (10.35)	Control (T score): Internalizing problems: 45.76 (9.31); [*P* < .05] Externalizing problems: 43.99 (9.39); [*P* < .05]	ADHD

TABLE 5 Continued

Reference	Alcohol Exposure / FASD Guideline Used[a]	Comparison Group	Alcohol-Exposed Group Scores, Mean (SD) or OR (CI)	Comparison Group Scores	Covariates Controlled For[b]
Stevens et al 2013[20]	All groups had PAE confirmed at any level. FASD: Canadian guidelines (n = 98)	PAE (no FASD): n = 50	FASD (T score): Internalizing problems: 62.02 (10.16) OR (95%CI): 2.52 (1.19 to 5.31) Externalizing problems: 71.15 (8.93) OR (95%CI): 4.58 (2.18 to 9.62) Total problems: 70.32 (7.34) OR (95%CI): 2.94 (1.42 to 6.11)	PAE (no FASD) (T score): Internalizing problems: 57.60 (10.89); [P < .05] Externalizing problems: 63.62 (11.63); [P < .05] Total problems: 63.70 (11.41); [P < .05]	None
Ware et al 2013[34]	PAE+no ADHD: Heavy PAE = >4 drinks ≥1/week or ≥13 drinks/week throughout preg (n = 52)	Control: n = 133	PAE+no ADHD (T score): Internalizing problems: 55.47 (10.65) Externalizing problems: 56.02 (9.81) Total problems: 56.59 (10.56)	Control (T score): Internalizing problems: 45.52 (9.39); [P < .001] Externalizing problems: 43.05 (9.1); [P < .001] Total problems: 41.72 (10.24); [P < .001:]	Race, ethnicity, home placement, gender.

±, with or without; ARND, alcohol related neurodevelopmental disorder; CIFASD, Collaborative initiative on fetal alcohol spectrum disorders; CNS, central nervous system; IOM, Institute of Medicine; NA, not applicable; NR, not reported; PFAS, partial fetal alcohol syndrome; preg, pregnancy; SES, socioeconomic status; yo, years old.

Where provided, P values between groups are presented in square brackets.

Data were extracted only for PAE/FASD groups and their comparison groups (without PAE/FASD), and also for groups assessed using the ASEBA school-aged forms (which excludes groups assessed using the ASEBA pre-school forms). Appendix 3 contains the detailed results for Total, Internalizing, and Externalizing problems scores for all subgroups within each study included in this review, as reported by parents/caregivers and teachers.

[a] This column reports either PAE levels (as reported in the articles) or FASD diagnostic guidelines used.

[b] Covariates controlled for in analysis of ASEBA outcomes specifically, which does not take into account whether covariates were controlled for in other aspects of the article which were not extracted for this review.

or Internalizing problems in FASD diagnostic subgroups (FAS/partial FAS; static encephalopathy, alcohol-exposed; neurobehavioral disorder, alcohol-exposed) using the 4-digit diagnostic code.[19] Critical items were not reported in any of the studies in our review.

Presence of PAE

The literature on PAE was less consistent than that on FASD in terms of behavior ratings among groups with different exposure levels, and ratings from different informant types. Generally, children with PAE had poorer parent and teacher ratings for Total competence and Total problems than nonexposed children. Most scales and subscales in the PAE literature were reported in only 1 study (Supplemental Appendices 4 and 5), and no study reported Critical items.

Risk of Bias

According to our methodological appraisal criteria (Table 2), scores for the 23 included studies ranged from 1 to 5 (mean: 2.7 [1.1]; Table 6). As per eligibility criteria, all studies reported ASEBA scores of interest. Only 1 study met criterion on sample size calculation/justification[19] (which is not standard practice in reporting of observational studies). Most (9) studies received a score of 3. In relation to statistical adjustment for other prenatal exposures or psychiatric/behavior diagnoses, studies were scored "No" (ie, 0) if adjustments were not applied to ASEBA results, even if adjustments were applied to other variables (given the current study focus on FASD/PAE effects on ASEBA scores). We were unable to examine funnel plot asymmetry, as fewer than 10 studies were included in the meta-analyses.

DISCUSSION

This systematic review and meta-analysis examining behavior in

FASD and PAE (as assessed using the popular ASEBA School-Age Forms) yields several important findings. First, of clinical concern, children with FASD had much poorer parent ratings on a range of behavioral outcomes than children without FASD. Second, there were no significant differences in behavior ratings in children with and without PAE when data for the most commonly reported parent-rated problem scales were pooled. Third, the published body of literature lacks quality and consistency of reporting. Over time, diagnostic systems have become more systematized and study quality has improved.

Our meta-analysis confirms that children with FASD have significantly higher scores for Total problems, Internalizing problems, and Externalizing problems than children without FASD. Pooled mean group differences in T scores ranged from 6 for Internalizing problems to 12 for Total and Externalizing problems. Children with FASD were more likely to have scores within the "Clinical" range for these primary measures with pooled ORs ranging from 10 (95% CI 1.3–77.6) to 34 (95% CI 2.6–450.8). Children with "Clinical" scores usually require therapeutic intervention. Although a limited number of scales were reported by both parents and teachers, there was general agreement between informants that children with FASD had poorer behavior than children without, with some inconsistencies when FASD and ADHD (without PAE) groups were compared.

Few informants for FASD groups were biological parents, whereas most informants in PAE studies were biological parents. The effect of different informants and home placements on CBCL ratings is unclear. Fagerlund et al[13] was the only author to compare CBCL ratings in different living situations:

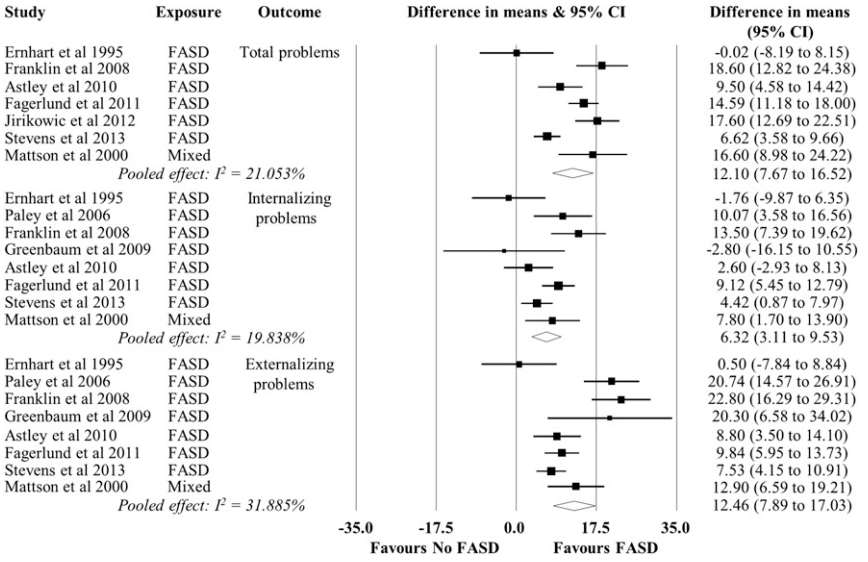

FIGURE 2
CBCL problem ratings (T scores) in FASD versus No FASD.

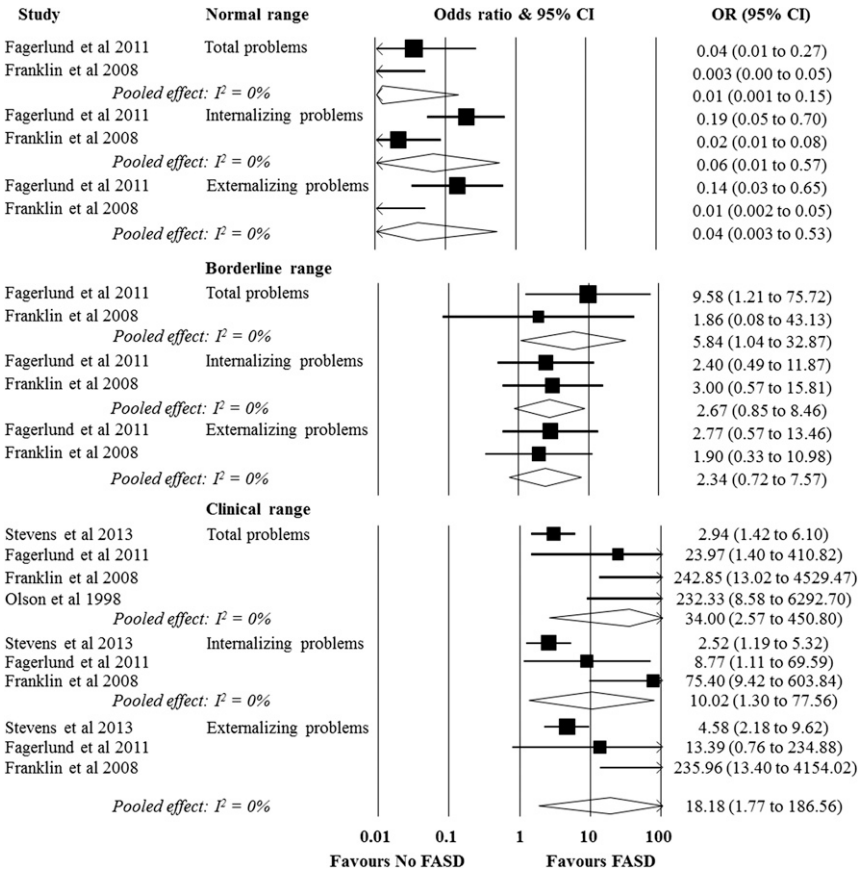

FIGURE 3
CBCL problem ratings (ORs) in the "Clinical" range in FASD versus No FASD.

children in residential care had more Total and Internalizing problems than those in foster/adoptive homes or biological homes, and more Externalizing problems than children in foster/adoptive

homes ($P < .05$). This may reflect the informants' viewpoint or the impact of institutional care. Paley et al[18] used the Parenting Stress Index and reported higher stress levels in foster/adoptive parents (77%) than biological parents (23%), but did not compare CBCL scores between these informants. Only 1 PAE study incorporated home placement as a covariate in their CBCL analyses and found no significant effect of place of residence on behavior.[34]

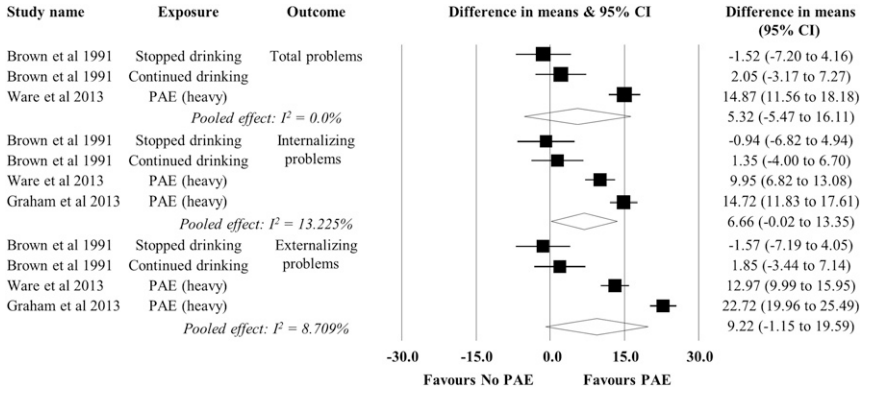

FIGURE 4
CBCL problem ratings (T scores) in PAE versus No PAE.

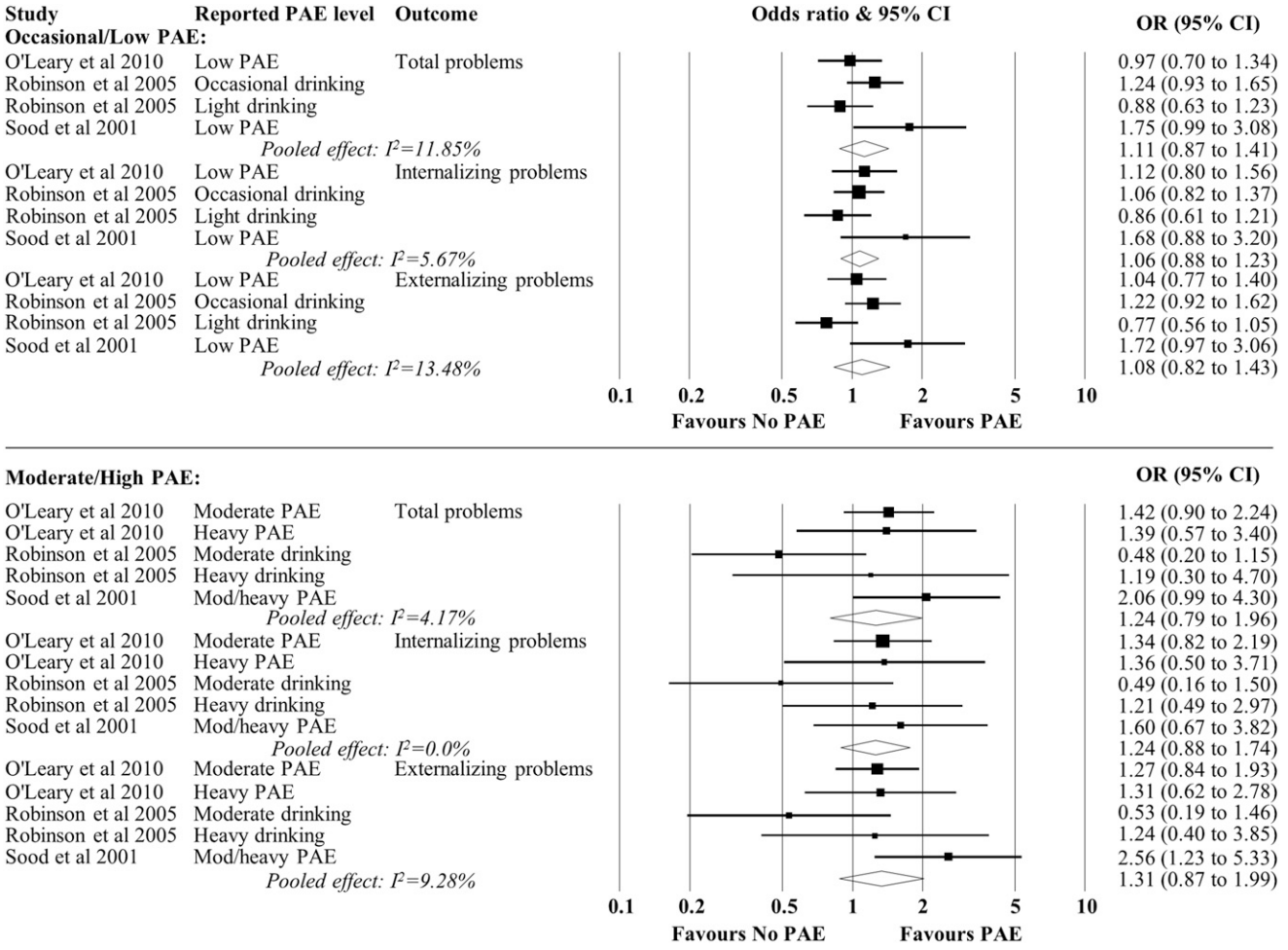

FIGURE 5
CBCL problem ratings (ORs) in the "Clinical" range according to PAE levels reported by authors.

In our qualitative assessment, children with FASD had poorer scores on the following scales: Total competence, Total problems, Internalizing problems, Externalizing problems, School competence, Rule breaking behavior/delinquency, Aggressive behavior, Attention problems, Social problems, and Social competence. These findings are in agreement with previous research reporting that items from the Rule

TABLE 6 Critical Appraisal of Articles

Reference	Defined Sample?	Representative Sample?	Outcome Rater Blinding?	ASEBA Scores of Interest Reported?	Sample Size Calculation?	Statistical Adjustment For Other Prenatal Exposures?	Comparison With Other Psychiatric / Behavior Diagnoses?	Score (/7)
Brown et al 1991[24]	Yes	No	Yes	Yes	No	No	No	3
Ernhart et al 1995[21]	Yes	No	No	Yes	No	Yes	No	3
Olson et al 1998[17]	No	No	No	Yes	No	No	No	1
Mattson and Riley 2000[35]	No	No	No	Yes	No	No	No	1
Sood et al 2001[32]	Yes	No	Yes	Yes	No	No	No	3
Lee et al 2004[27]	No	No	No	Yes	No	Yes	Yes	3
Paley et al 2005[30]	No	Yes	No	Yes	No	No	No	2
Sood et al 2005[33]	Yes	No	No	Yes	No	Yes	No	3
Paley et al 2006[18]	No	No	No	Yes	No	No	No	1
Bada et al 2007[23]	Yes	Yes	No	Yes	No	No	No	3
O'Callaghan et al 2007[28]	Yes	Yes	No	Yes	No	Yes	No	4
Franklin et al 2008[14]	No	No	No	Yes	No	No	No	1
Alati et al 2009[22]	Yes	Yes	No	Yes	No	No	No	3
Chiu et al 2009[25]	No	Yes	No	Yes	No	No	No	2
Greenbaum et al 2009[15]	No	No	No	Yes	No	Yes	Yes	3
Astley 2010[19]	Yes	No	No	Yes	Yes	No	No	3
O'Leary et al 2010[29]	Yes	Yes	No	Yes	No	Yes	No	4
Robinson et al 2010[31]	Yes	Yes	No	Yes	No	Yes	No	4
Fagerlund et al 2011[13]	Yes	No	No	Yes	No	No	No	2
Jirikowic et al 2012[16]	No	No	Yes	Yes	No	No	No	2
Graham et al 2013[26]	Yes	Yes	No	Yes	No	No	Yes	4
Stevens et al 2013[20]	Yes	No	No	Yes	No	No	No	2
Ware et al 2013[34]	Yes	Yes	Yes	Yes	No	No	Yes	5

breaking, Attention problems, and Aggressive CBCL scales discriminated children with FASD from children with ADHD but no PAE, oppositional defiant/conduct disorder, and typically developing healthy control children, although competence items were not examined in those studies.[7] These findings concur with studies that use different behavioral assessments.[36,37] Congruent with our review findings, a major focus in FASD intervention research is to develop and validate interventions for challenging behaviors, social skills deficits, and problems with self-regulation.[38]

No significant problems were observed in children with FASD in Activities competence, Somatic complaints, Withdrawn/depressed, Anxiety/depression, Academic performance (1 TRF study), Adaptive functioning (1 TRF study), or Hyperactive scales. These findings do not concur with studies that use other assessment tools, which report impairments in most of these areas,[37,39] although few studies in this review reported each of these outcomes. It is possible that ASEBA items assessing these constructs are insufficiently sensitive to these deficits in the FASD population.

We found no group differences in behavior among children with and without PAE. The number of studies included was limited and PAE levels were inconsistently defined, so it was difficult to establish a dose-response pattern by standardized definitions. However, when using classifications for PAE levels provided by authors, there was a tendency toward increasing behavior problems with higher levels of PAE. Most articles reported only a few of the possible 78 scales from the CBCL, TRF, and YSR inclusive, and 28 scales were reported in only 1 study, precluding generalization of results.

A potential limitation of this review was the focus on assessments using the ASEBA. However, in many included studies, other measures used concurrently identified similar behavior problems. The review unveiled limitations within the available literature. Study quality was generally low, perhaps due to inclusion of any observational study that reported outcomes of interest, regardless of whether ASEBA scores reported were of primary interest in that study (hence many studies did not calculate sample sizes or statistical adjustments for ASEBA results). Our focus on the ASEBA scales should be set in context for clinicians. We have documented significant behavior problems in FASD; however, challenging behavior is only 1 aspect of central nervous system dysfunction in FASD. Neurocognitive and communication deficits, for example, are also common.

Several studies drew data from the same databases, so some children may have been included more than once. Thus, subjects in different studies may not be independent. Different FASD diagnostic systems

were used in included studies, as several different criteria are used internationally. However, the consistency of behavioral outcomes within our FASD meta-analyses suggests little effect of diagnostic criteria on behavior ratings assessed by using the ASEBA School-Age Forms.

The paucity of data made it difficult to assess several ASEBA scales or to stratify by age. Inconsistencies in reporting of results precluded assessment of publication bias because <10 studies could be included in any meta-analysis. Although risk of bias could not be addressed quantitatively in this review, it is worth mentioning that data from 2 different kinds of studies were included: Longitudinal cohort studies (based on PAE), and case-control studies (based on FASD diagnosis). Findings from separate PAE and FASD meta-analyses are intuitive, and possibly reflect a detection bias favoring greater effects in children with FASD than PAE. This is because women drinking alcohol during pregnancy do not necessarily give birth to a child with behavioral impairments or FASD, whereas children with an FASD diagnosis are more likely to show greater severity of impairments.

There are strengths in the current study, including the meta-analysis of results derived from the most commonly reported behavior assessment tool used to study FASD and PAE. Additionally, we summarized all scales and subscales reported in the literature using the ASEBA School-Age Forms in this systematic review.

Research and Clinical Implications

Overall, this review unveiled behaviors apparently less problematic in FASD, behaviors in need of additional study due to inconsistent findings, or an insufficient number of studies on particular scales. It also revealed consistency in behavior ratings between different studies and by different informants in studies of FASD, but not PAE.

Findings provide direction for future research on FASD and PAE. Descriptive data comprehensively detailing behavioral deficits are crucial to guide research on interventions for the surprisingly prevalent FASD. To improve quality of the knowledge base, there is clear need for consistency in (1) reporting of and stratification according to PAE levels and timing, (2) selection of ASEBA scales, and (3) reporting of ASEBA scores. When feasible, methodology changes are needed. For instance, in ASEBA research, it would be useful to use multiple informants, to blind assessors to PAE/FASD status, and to control for other psychiatric/behavior diagnoses with similar symptoms by using contrast groups (eg, ADHD or IQ comparison groups with no teratogenic exposure). More complete investigation and reporting of the often overlooked ASEBA subscales and Critical items are warranted, especially given their significance for clinical practice.

Behavior problems are almost universal in children with FASD and significantly affect family life and schooling. From a clinical perspective, a clear understanding of behaviors in children with FASD and after PAE is important to families and providers seeking to improve outcomes for affected individuals. This review makes clear that providers should ask about PAE, and refer for assessment/diagnosis of FASD when appropriate. Indeed, FASD should be included in the differential diagnosis of any child with behavioral difficulties. Children with FASD and clinically concerning behavior problems may benefit considerably from early intervention, and knowledge of behaviors associated with FASD will inform diagnosis and management.

ACKNOWLEDGMENTS

The authors thank the librarian at The Children's Hospital at Westmead, Ms Trish Bennett, for her help with the electronic search strategy.

ABBREVIATIONS

ADHD: attention deficit hyperactivity disorder
ASEBA: Achenbach system of empirically based assessment
CBCL: child behavior checklist
CI: confidence interval
FAS: fetal alcohol syndrome
FASD: fetal alcohol spectrum disorders
OR: odds ratio
PAE: prenatal alcohol exposure
TRF: teacher report form
YSR: youth self-report

some articles, and reviewed and revised the final manuscript as submitted; Prof Elliott conceptualized the study, contributed to the design of the study and data collection forms, and reviewed and critically revised the manuscript; and all authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

This trial has been registered with PROSPERO (http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014007040#.VOJ7A8II_MA) (identifier CRD42014007040).

DOI: 10.1542/peds.2015-2542

Accepted for publication Nov 20, 2015

Address correspondence to Tracey W. Tsang, PhD, Kids' Research Institute (APSU), The Children's Hospital at Westmead, Locked Bag 4001, Westmead NSW 2145, Australia. E-mail: tracey.tsang@sydney.edu.au

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

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

FUNDING: The Discipline of Paediatrics and Child Health, Sydney Medical School, The University of Sydney, New South Wales, Australia, is responsible for initiating and managing the review. Drs Tsang and Pinto are funded by a National Health and Medical Research Council project grant (1024474). Prof Elliott is supported by the National Health and Medical Research Council Practitioner Fellowships (457084 and 1021480). Ms Lucas was supported by a fellowship from the Poche Centre for Indigenous Health, Sydney Medical School, The University of Sydney, Sydney, Australia. Prof Carmichael Olson has ongoing salary support from the University of Washington School of Medicine and Seattle Children's Research Institute.

POTENTIAL CONFLICT OF INTEREST: Prof Carmichael Olson was an author of one of the articles included in the review; the other authors have indicated they have no potential conflicts of interest to disclose.

REFERENCES

1. Goh YI, Chudley AE, Clarren SK, et al; Taskforce for Development of FASD Screening Tools. Development of Canadian screening tools for fetal alcohol spectrum disorder. *Can J Clin Pharmacol*. 2008;15(2):e344–e366
2. Mattson SN, Crocker N, Nguyen TT. Fetal alcohol spectrum disorders: neuropsychological and behavioral features. *Neuropsychol Rev*. 2011;21(2):81–101
3. O'Connor MJ, Paley B. Psychiatric conditions associated with prenatal alcohol exposure. *Dev Disabil Res Rev*. 2009;15(3):225–234
4. Hellemans KG, Sliwowska JH, Verma P, Weinberg J. Prenatal alcohol exposure: fetal programming and later life vulnerability to stress, depression and anxiety disorders. *Neurosci Biobehav Rev*. 2010;34(6):791–807
5. 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(4):228–238
6. Nash K, Rovet J, Greenbaum R, Fantus E, Nulman I, Koren G. Identifying the behavioural phenotype in Fetal Alcohol Spectrum Disorder: sensitivity, specificity and screening potential. *Arch Women Ment Health*. 2006;9(4):181–186
7. Koren G, Zelner I, Nash K, Koren G. Foetal alcohol spectrum disorder: identifying the neurobehavioural phenotype and effective interventions. *Curr Opin Psychiatry*. 2014;27(2):98–104
8. Achenbach TM, Rescorla LA. *Manual for the ASEBA School-Age Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families; 2001
9. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol*. 2009;62(10):e1–e34
10. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol*. 2007;36(3):666–676
11. Higgins JPT, Green S, eds. Cochrane handbook for systematic reviews of interventions: version 5.1.0 [updated March 2011]. *The Cochrane Collaboration* 2011; Recommendations on testing for funnel plot asymmetry. Available at: http://handbook.cochrane.org/chapter_10/10_4_3_1_recommendations_on_testing_for_funnel_plot_asymmetry.htm. Accessed July 2, 2014
12. O'Leary CM, Nassar N, Kurinczuk JJ, et al. Prenatal alcohol exposure and risk of birth defects. *Pediatrics*. 2010;126(4). Available at: www.pediatrics.org/cgi/content/full/126/4/e843
13. Fagerlund A, Autti-Rämö I, Hoyme HE, Mattson SN, Korkman M. Risk factors for behavioural problems in foetal alcohol spectrum disorders. *Acta Paediatr*. 2011;100(11):1481–1488
14. Franklin L, Deitz J, Jirikowic T, Astley S. Children with fetal alcohol spectrum disorders: problem behaviors and sensory processing. *Am J Occup Ther*. 2008;62(3):265–273
15. Greenbaum RL, Stevens SA, Nash K, Koren G, Rovet J. Social cognitive and emotion processing abilities of children with fetal alcohol spectrum disorders: a comparison with attention deficit hyperactivity disorder. *Alcohol Clin Exp Res*. 2009;33(10):1656–1670
16. Jirikowic T, Carmichael Olson H, Astley S. Parenting stress and sensory processing: children with fetal alcohol spectrum disorders. *OTJR Occup Participation Health*. 2012;32(4):160–168
17. Olson HC, Feldman JJ, Streissguth AP, Sampson PD, Bookstein FL. Neuropsychological deficits in adolescents with fetal alcohol syndrome: clinical findings. *Alcohol Clin Exp Res*. 1998;22(9):1998–2012
18. Paley B, O'Connor MJ, Frankel F, Marquardt R. Predictors of stress in parents of children with fetal alcohol spectrum disorders. *J Dev Behav Pediatr*. 2006;27(5):396–404
19. Astley SJ. Profile of the first 1,400 patients receiving diagnostic evaluations for fetal alcohol spectrum disorder at the Washington State Fetal Alcohol Syndrome Diagnostic & Prevention Network. *Can J Clin Pharmacol*. 2010;17(1):e132–e164
20. Stevens SA, Nash K, Fantus E, Nulman I, Rovet J, Koren G. Towards identifying a characteristic neuropsychological profile for fetal alcohol spectrum

disorders. 2. Specific caregiver-and teacher-rating. *J Popul Ther Clin Pharmacol*. 2013;20(1):e53–e62

21. Ernhart CB, Greene T, Sokol RJ, Martier S, Boyd TA, Ager J. Neonatal diagnosis of fetal alcohol syndrome: not necessarily a hopeless prognosis. *Alcohol Clin Exp Res*. 1995;19(6):1550–1557
22. Alati R, Najman JM, O'Callaghan M, Bor W, Williams GM, Clavarino A. Fetal growth and behaviour problems in early adolescence: findings from the Mater University Study of Pregnancy. *Int J Epidemiol*. 2009;38(5):1390–1400
23. Bada HS, Das A, Bauer CR, et al. Impact of prenatal cocaine exposure on child behavior problems through school age. *Pediatrics*. 2007;119(2). Available at: www.pediatrics.org/cgi/content/full/119/2/e348
24. Brown RT, Coles CD, Smith IE, et al. Effects of prenatal alcohol exposure at school age. II. Attention and behavior. *Neurotoxicol Teratol*. 1991;13(4):369–376
25. Chiu YN, Gau SSF, Tsai WC, Soong WT, Shang CY. Demographic and perinatal factors for behavioral problems among children aged 4–9 in Taiwan. *Psychiatry Clin Neurosci*. 2009;63(4):569–576
26. Graham DM, Crocker N, Deweese BN, et al; CIFASD. Prenatal alcohol exposure, attention-deficit/hyperactivity disorder, and sluggish cognitive tempo. *Alcohol Clin Exp Res*. 2013;37(suppl 1): E338–E346
27. Lee KT, Mattson SN, Riley EP. Classifying children with heavy prenatal alcohol exposure using measures of attention. *J Int Neuropsychol Soc*. 2004;10(2):271–277
28. O'Callaghan FV, O'Callaghan M, Najman JM, Williams GM, Bor W. Prenatal alcohol exposure and attention, learning and intellectual ability at 14 years: a prospective longitudinal study. *Early Hum Dev*. 2007;83(2):115–123
29. O'Leary CM, Nassar N, Zubrick SR, Kurinczuk JJ, Stanley F, Bower C. Evidence of a complex association between dose, pattern and timing of prenatal alcohol exposure and child behaviour problems. *Addiction*. 2010;105(1):74–86
30. Paley B, O'Connor MJ, Kogan N, Findlay R. Prenatal alcohol exposure, child externalizing behavior, and maternal stress. *Parent Sci Pract*. 2005;5(1):29–56
31. Robinson M, Oddy WH, McLean NJ, et al. Low-moderate prenatal alcohol exposure and risk to child behavioural development: a prospective cohort study. *BJOG*. 2010;117(9):1139–1150
32. 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(2). Available at: www.pediatrics.org/cgi/content/full/108/2/E34
33. Sood BG, Nordstrom Bailey B, Covington C, et al. Gender and alcohol moderate caregiver reported child behavior after prenatal cocaine. *Neurotoxicol Teratol*. 2005;27(2):191–201
34. Ware AL, O'Brien JW, Crocker N, et al; CIFASD. The effects of prenatal alcohol exposure and attention-deficit/hyperactivity disorder on psychopathology and behavior. *Alcohol Clin Exp Res*. 2013;37(3):507–516
35. Mattson SN, Riley EP. Parent ratings of behavior in children with heavy prenatal alcohol exposure and IQ-matched controls. *Alcohol Clin Exp Res*. 2000;24(2):226–231
36. Niccols A. Fetal alcohol syndrome and the developing socio-emotional brain. *Brain Cogn*. 2007;65(1):135–142
37. Kodituwakku PW. Defining the behavioral phenotype in children with fetal alcohol spectrum disorders: a review. *Neurosci Biobehav Rev*. 2007;31(2):192–201
38. Bertrand J; Interventions for Children with Fetal Alcohol Spectrum Disorders Research Consortium. Interventions for children with fetal alcohol spectrum disorders (FASDs): overview of findings for five innovative research projects. *Res Dev Disabil*. 2009;30(5):986–1006
39. Streissguth AP, O'Malley K. Neuropsychiatric implications and long-term consequences of fetal alcohol spectrum disorders. *Semin Clin Neuropsychiatry*. 2000;5(3):177–190

Prenatal Alcohol Exposure, FASD, and Child Behavior: A Meta-analysis

Tracey W. Tsang, Barbara R. Lucas, Heather Carmichael Olson, Rafael Z. Pinto and Elizabeth J. Elliott

Pediatrics originally published online February 22, 2016;

PEDIATRICS®
OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

Prenatal Alcohol Exposure, FASD, and Child Behavior: A Meta-analysis

Tracey W. Tsang, Barbara R. Lucas, Heather Carmichael Olson, Rafael Z. Pinto and Elizabeth J. Elliott

Pediatrics originally published online February 22, 2016;