CADTH

CADTH RAPID RESPONSE REPORT:
SUMMARY WITH CRITICAL APPRAISAL

Fetal Alcohol Spectrum Disorders: A Review of Diagnostic Test Accuracy, Clinical and Cost-Effectiveness of Diagnosis and Treatment, and Guidelines

Service Line:	Rapid Response Service
Version:	1.0
Publication Date:	May 3, 2017
Report Length:	45 Pages

CADTH

Authors: Veronica Poitras, Charlene Argáez

Cite As: Diagnosis, assessment, and treatment of fetal alcohol spectrum disorders: a review of clinical and cost-effectiveness and guidelines. Ottawa: CADTH; 2017 May. (CADTH rapid response report: summary with critical appraisal).

Acknowledgments:

ISSN: 1922-8147 (online)

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Context and Policy Issues

Fetal alcohol spectrum disorder (FASD) is the overarching term for the range of mild to severe physical, mental, behavioural, and cognitive effects that can occur in an individual whose mother consumed alcohol during pregnancy.[1] This spectrum includes: alcohol related neurodevelopmental disorder (ARND), alcohol related birth defects (ARBD), partial fetal alcohol syndrome (pFAS), and fetal alcohol syndrome (FAS).[2] FAS is considered the most severe outcome of prenatal alcohol exposure, and it is diagnosed given the presence of all three of the following criteria: facial dysmorphia (smooth philtrum, thin upper lip, and short distance between the inner and outer corners of the eyes), growth problems (height, weight, or both <10th percentile at any point in time, including before birth), and central nervous system problems (structural, neurologic, or functional).[1] Confirmed absence of alcohol during pregnancy rules out a diagnosis of FASD, but confirmed alcohol use is not required.[1] Diagnostic practices vary considerably; clinical categorization is not standardized.[3,4]

FASD is considered the most common preventable cause of birth defects and developmental disability worldwide.[5] There are no confirmed national statistics on the rates and prevalence of FASD in Canada, but it is estimated that FASD occurs at a rate of 1 out of every 100 live births[5] and that this may vary substantially between populations.[6] For example, in northeastern Manitoba, FASD was found to occur at a rate of 0.72 per 100 live births, compared to 5.5 to 10.1 per 100 in a Manitoba First Nations community.[6] Contrary to common belief however, FASD is not associated with ethnic or cultural background.[6] Rather, risk factors for prenatal alcohol exposure include: higher maternal age, unemployment, lower education level or socioeconomic status; maternal exposure to stress, abuse, undernutrition or under-treated mental health concerns; partner or family member alcohol or drug use at time of pregnancy; and reduced access to prenatal and/or postnatal care.[1,6]

FASD can be diagnosed at birth, but often goes undiagnosed until later in life when behavioural and cognitive effects are more evident.[7] Early diagnosis (e.g., before age 6 years) and intervention are considered critical to improve development and to reduce the likelihood of secondary disabilities,[1,8,9] however access to assessment and treatment is not widely available across Canada and may be particularly limited in rural and remote areas.[10] Because the effects of FASD are so diverse, treatment options are usually multidisciplinary and include medical care, medication, behaviour and education therapy, parent or caregiver training, or alternative approaches.[1] FASD has lifelong direct and indirect costs associated with health, education, and social supports for affected individuals, and is estimated to cost at least $4 to $5.3 billion per year in Canada.[5,10].

The purpose of this report is to examine the diagnostic test accuracy, clinical utility, and cost-effectiveness of diagnosis and/or assessment of FASD, the clinical and cost-effectiveness of treatment of FASD, and the evidence-based guidelines associated with diagnosis, assessment, or treatment of FASD in individuals of any age.

Research Questions

1. What is the diagnostic test accuracy of tools or tests for the diagnosis and/or assessment of fetal alcohol spectrum disorder in individuals of any age?
2. What is the clinical utility of diagnosis and/or assessment of fetal alcohol spectrum disorder in individuals of any age?
3. What is the clinical effectiveness of the treatment of fetal alcohol spectrum disorder in individuals of any age?
4. What is the cost-effectiveness of diagnosis and/or assessment of fetal alcohol spectrum disorder in individuals of any age?
5. What is the cost-effectiveness of the treatment of fetal alcohol spectrum disorder in individuals of any age?
6. What are the evidence-based guidelines associated with the diagnosis/assessment of and treatment for fetal alcohol spectrum disorder in individuals of any age?

Key Findings

Preliminary evidence from ten diagnostic test accuracy studies indicated that a variety of tests or tools, including a decision tree model, checklist, test battery, narrative analysis tool, and computer-assisted landmark-based morphometric face analysis, showed promise as diagnostic tools in comparison to reference standards of unknown accuracy. Overall, there was insufficient evidence to suggest an optimal diagnostic test for FASD, and there remains no “gold standard” for FASD diagnosis. No evidence for the clinical utility or cost-effectiveness of diagnosis and/or assessment of FASD in individuals of any age was identified. Limited evidence from three systematic reviews and one primary study suggested that multi-dimensional treatment strategies (that include physical, mental health, behavioural, cognitive, and/or pharmacologic components) that are individually tailored for patients may be clinically effective. A single economic analysis indicated that the use of an FASD Service Network may result in cost-savings by preventing secondary disabilities (such as crime, homelessness, and mental health problems), however this analysis was based on model components that were estimated from the literature and effectiveness studies are needed. One evidence-based Canadian guideline was identified that provides recommendations regarding the diagnosis of FASD. Based on predominantly high-quality evidence, the guidelines provide strong recommendations for multi-disciplinary diagnosis based on criteria related to facial features, prenatal alcohol exposure, and neurodevelopmental effects.

Methods

Literature Search Methods

A limited literature search was conducted on key resources including PubMed, The Cochrane Library, University of York Centre for Reviews and Dissemination (CRD) databases, ECRI, Canadian and major international health technology agencies, as well as a focused Internet search. No filters were applied to limit the retrieval by study type. Where possible, retrieval was limited to the human population. The search was also limited to English language documents published between January 1, 2007 and April 5, 2017.

Rapid Response reports are organized so that the evidence for each research question is presented separately.

Selection Criteria and Methods

One reviewer screened citations and selected studies. In the first level of screening, titles and abstracts were reviewed and potentially relevant articles were retrieved and assessed for inclusion. The final selection of full-text articles was based on the inclusion criteria presented in Table 1.

Table 1: Selection Criteria

Population	Individuals of any age suspected of having, or diagnosed with, FASD
Intervention	Q1,2,4: Diagnostics and/or assessment of FASD Q3, 5: Treatment of FASD (e.g., programs, services, medications, etc.) Q6: Diagnostics, assessment, and/or treatment of FASD
Comparator	Q1,2,4: Any other diagnostic assessment or tool, or no diagnostic assessment or tool Q3,5: Any other treatment, or no treatment Q6: No comparator
Outcomes	Q1: Diagnostic test accuracy (e.g., sensitivity, specificity, positive and negative predictive values) Q2: Clinical utility outcomes (e.g., rate of detection of patients with FASD, number of patients referred to treatment, proportion of patients treated, improvement in care leading to better clinical outcomes) Q3: Any clinical outcome (e.g., behaviour, social skills, communication, quality of life) Q4,5: Cost-effectiveness Q6: Guidelines
Study Designs	Health technology assessments, systematic reviews, meta-analyses, randomized controlled trials, non-randomized studies, economic evaluations, and guidelines

FASD = fetal alcohol spectrum disorder.

Exclusion Criteria

Articles were excluded if they did not meet the selection criteria outlined in Table 1, they were duplicate publications, or were published prior to 2007. Guidelines from countries other than Canada and the United States were excluded.

Critical Appraisal of Individual Studies

The included systematic reviews were critically appraised using the AMSTAR tool,[11] the primary clinical study was critically appraised using the Downs and Black checklist,[12] diagnostic test accuracy studies were assessed using the QUADAS-2,[13] the economic evaluation was assessed using the Drummond checklist,[14] and guidelines were assessed with the AGREE II instrument.[15] Summary scores were not calculated for the included studies; rather, a review of the strengths and limitations of each included study were described.

Summary of Evidence

Quantity of Research Available

A total of 498 citations were identified in the literature search. Following screening of titles and abstracts, 459 citations were excluded and 39 potentially relevant reports from the

electronic search were retrieved for full-text review. Three potentially relevant publications were retrieved from the grey literature search. Of these potentially relevant articles, 26 publications were excluded for various reasons, and 16 publications met the inclusion criteria and were included in this report. Appendix 1 describes the PRISMA flowchart of the study selection. Additional references of potential interest are provided in Appendix 6.

Summary of Study Characteristics

Additional details regarding the characteristics of included publications are provided in Appendix 2.

Study Design

Three relevant systematic reviews (SRs) were identified.[2,16,17] Two SRs concerned the clinical effectiveness of the treatment of Fetal Alcohol Spectrum Disorder (FASD),[2,16] and the third SR aimed to consider the diagnosis, economic impact, and clinical effectiveness of treatment of FASD.[17] The reviews were published in 2008,[17], 2009,[2] and 2015,[16], and included literature searches from inception to between March and November 2008, or up to January 2009 or December 2014 respectively. Two SRs included both randomized and non-randomized primary studies;[2,16] the third SR included other SRs and narrative reviews.[17] There was overlap in the studies included in the SRs (Appendix 5).

Ten studies regarding the diagnostic test accuracy of tools or tests for the diagnosis and/or assessment of FASD were identified.[3,4,7,9,18-23] One study used a prospective cohort design,[7] and all others were case-control studies. No studies were identified that assessed the clinical utility or cost-effectiveness of diagnosis and/or assessment of FASD.

One before-and-after study regarding the clinical effectiveness of treatment of FASD was identified.[24] One economic study regarding the cost-effectiveness of treatment, that included a cost-benefit analysis, was identified.[25]

One evidence-based guideline, targeted to multidisciplinary diagnostic teams in Canada, was identified regarding the diagnosis of FASD across the lifespan.[26] The guideline was informed by a systematic review of the literature that was published as a supplementary Appendix,[27] and included ratings of the quality of evidence and strength of recommendations according to the Grading of Recommendations Assessment, Development and Evaluation framework.[26] Recommendations were consensus-based.

Country of Origin

The SRs were led by authors based in Australia[2,16,17] and New Zealand.[17] The diagnostic accuracy studies were conducted in Finland,[9,22] South Africa,[3,9,19,20,22] Ukraine,[7] and the United States,[4,18,21-23] The before-and-after treatment study was conducted in the United States,[24] and the economic evaluation was conducted in Canada.[25] The evidence-based guideline for diagnosis was developed in Canada.[26]

Patient Population

The two SRs that concerned the clinical effectiveness of FASD treatment included individuals of any age with prenatal alcohol exposure (PAE) (including fetal alcohol syndrome [FAS], partial FAS [pFAS], alcohol-related neurodevelopmental disorder [ARND], or PAE),[16] or individuals under 18 years of age with a diagnosis of FASD.[2] The third SR included individuals who may have FASD or mothers of individuals who may have FASD.[17]

In the diagnostic accuracy studies, one case-control study included cases with prenatal alcohol exposure and non-exposed controls,[18] five case-control studies included cases with FASD and typically developing controls,[3,4,19,21,23] and three case-control studies included cases with FAS specifically and healthy controls.[9,20,22] The prospective cohort study included mother/child dyads with or without alcohol exposure.[7]

The before-and-after clinical effectiveness study included children with FASD and children with typical development.[24] The economic evaluation included high-risk individuals who were referred to the Alberta FASD Service Networks (SNs) for diagnostic services, and those who were diagnosed with FASD outside of the SNs but were referred to the SN for support, from 2008 to 2011.[25]

The evidence-based guidelines for diagnosis apply to pregnant or postpartum women and individuals at risk of FASD.[26] The intended users are multi-disciplinary diagnostic teams in Canada who have expertise through specialized training and experience.[26]

Interventions and Comparators

In the SRs, a variety of treatment interventions were considered: behavioural,[2,16] educational,[2] social skills and communication,[2] pharmacological,[2] advocacy or support,[16] or any strategy that aimed to improve clinical outcomes in individuals with FASD.[17] One SR additionally considered any strategy that aimed to diagnose an individual who may have FASD, or any strategy to reduce the financial burden of FASD.[17] Eligible comparators included a control group (no treatment, waiting list, usual therapy or placebo) or pre- and post-intervention measurements (for cohort studies),[2] any comparator,[17] or comparators were not specified.[16]

In the diagnostic test accuracy studies, the interventions (index tests) were: a decision tree model including cognitive, intellectual and physical factors;[18] second trimester fetal ultrasound measures,[7] a regression model developed from a test battery including intelligence, cognitive, motor, and behaviour factors;[19] or the Fetal Alcohol Syndrome Diagnostic Checklist.[4] Four studies used various computer-assisted landmark-based morphometric facial analyses,[3,9,20,22] and two studies used a narrative analysis tool[21,23] as the index test. In the study that compared prenatal alcohol-exposed cases with non-exposed controls, the comparator (reference standard) was knowledge of exposure history as obtained through retrospective maternal report, social service, or legal or medical records.[18] Other comparators (reference standards) were various clinical diagnoses based on application of the Four-Digit Diagnostic Code;[21,23] structural features and growth deficiency consistent with the revised Institute of Medicine criteria;[9,22] dysmophological examination at birth and neurobehavioural evaluation at age 6 and/or 12 months;[7] a standard protocol and the Astley Lip-Philtrum Guide 14;[3] sufficient dysmorphology and behaviour problems with confirmation of prenatal alcohol exposure;[20] or clinical diagnosis based on unspecified criteria.[4,19]

In the primary clinical effectiveness study, the before-and-after intervention was 30 minutes of “Sensorimotor Training to Affect Balance, Engagement and Learning” (STABEL), a virtual reality system to train sensory control for balance. The comparator was measurement of outcomes pre-STABEL training.

The economic study evaluated the break-even effectiveness (i.e., the effectiveness level at which the intervention became cost-saving) of 12 Alberta FASD SNs compared to no SNs.

The eligible comparators were not specified in the guidelines for diagnosis.[26]

Outcomes

The outcomes considered in the SRs related to clinical effectiveness of FASD treatment were the following: developmental status;[2,16] self-regulation and attentional control;[16] cognitive status;[2] specific skills (e.g., math knowledge, skills, or reasoning; nonverbal reasoning; reading comprehension);[16], social skills;[16] quality of life;[2] parenting skills;[16] employment;[2] contact with the law;[2], substance abuse;[2] support, education and advocacy;[16] supporting parents who have FASD;[16] and reduction in severity of primary and/or secondary disabilities or deficits associated with FASD.[17] One SR also considered the sensitivity and specificity of FASD diagnosis, and cost-effectiveness of strategies to reduce FASD, as relevant outcomes.[17]

In most of the diagnostic test accuracy studies, the outcomes were classification accuracy, sensitivity, and/or specificity.[4,9,18-23] In one study, the outcome was agreement between clinical categorization and classifications from face shape alone,[3] and in the study that used ultrasound measures as the index test the outcome was the proportion of variance in FASD accounted for by combined and individual ultrasound measures.[7]

In the primary clinical effectiveness study, the outcomes of interest were sensory attention and postural control.[24]

The relevant outcomes in the economic evaluation were: cost of secondary disabilities (crime, homelessness, mental health problems, school disruption [< 18 y old] or unemployment [≥ 18 y]) with the absence of the intervention; cost of the intervention; and break-even effectiveness.[25]

The outcomes of interest were not specified in the guidelines for diagnosis.[26]

Summary of Critical Appraisal

Additional details regarding the strengths and limitations of included publications are provided in Appendix 3.

Systematic Reviews

Strengths common to all three SRs included the use of comprehensive literature searches, and provision of a list and some key characteristics of included studies.[2,16,17] In contrast, none of the included SRs provided details on other key characteristics, such as age at diagnosis or comorbidities.[2,16,17] Study screening and selection, and data extraction, were performed in duplicate in two SRs,[2,16] but in the third SR study selection was performed by a single reviewer and only data extraction was checked for accuracy by a second reviewer.[17] Consideration of the scientific quality of evidence varied across the SRs; quality was assessed in one SR using the Effective Public Health Project assessment tool,[16] the second SR considered some characteristics affecting quality but was not comprehensive,[2] and the third SR did not include a formal assessment or documentation of scientific quality of but mentioned strengths and limitations of included studies in the narrative summary.[17] No SRs included meta-analyses; two SRs determined that it was not clinically appropriate to combine studies,[2,16] and the third SR included other reviews that did not conduct meta-analyses.[17] In addition, none of the SRs assessed the likelihood of publication bias or reported conflict of interest for included studies.[2,16,17] The review authors declared no conflict of interest in two of the SRs,[2,16] but conflict of interest was not reported for the third.[17] In all three SRs, conclusions were based on statistical significance; clinical significance was not considered.[2,16,17]

Diagnostic Test Accuracy Studies

Most of the diagnostic test accuracy studies evaluated diagnostic tests that are still in preliminary stages of development. As a result, most of the studies used index tests that are unavailable for clinical use (e.g., regression equations, 3-dimensional face matching algorithms).[3,7,9,19-23] All but one[7] of the included studies used a case-control design, and the method of participant sampling was unclear or inappropriate in eight studies.[3,9,18-23] Importantly, there is no consensus on the most appropriate reference standard for FASD, and eight different reference standards, with unknown accuracy, were used across the included studies. In four studies, "difficult-to-diagnose" patients were purposefully excluded (e.g., those without a strict FAS diagnosis, or those without an FASD diagnosis but who may have had some prenatal alcohol exposure),[9,19,21,23] and in five studies the index test was both developed and evaluated in the same study sample.[3,9,19,20,22] The study that used ultrasound measurements as the index test[7] has unknown applicability because FASD is usually diagnosed at or after birth.

In terms of strengths, "difficult-to-diagnose" patients were purposefully included in the study by Goh et al. 2016, and the decision tree model (index test) that was being tested had been developed in a different sample.[18] In seven studies, all patients received the same reference standard;[3,7,9,18-20,22] in seven studies all patients were included in the analysis,[4,7,18,20-23] and in four studies no inappropriate exclusion criteria were reported.[4,7,18,23]

Clinical Effectiveness Study

The before-and-after study was well-reported: it had a clearly described objective; provided participant inclusion and exclusion criteria; and clearly described characteristics of the population, intervention, comparator, outcomes, and main findings with statistical estimates of variability and exact probability values (p-values) reported.[12] The study was limited, however, by use of a small convenience sample with no performance of a power calculation, unknown validity and reliability of the main outcome measure, absence of a time-control group, and no blinding of the examiner to group or pre- or post-training status.[12]

Economic Evaluation

One SR, published in 2008, included a literature search for publications related to the cost-effectiveness of treatment of FASD, however no eligible studies were identified.[17]

The single economic evaluation was a cost-benefit analysis.[25] Strengths included a clear research question and appropriate intervention and comparator, outcome measures, analysis perspective, and form of economic evaluation used. The methods of estimation of numbers and costs of secondary disabilities were provided, and the costs of the intervention were measured directly. The costs were converted to 2012 Canadian dollars and details of adjustment for inflation were provided. The main results were clearly presented. A limitation of the economic study was that model components were all estimated from the literature, and not based on effectiveness studies. In addition, some of the model assumptions may not be valid; for instance, probabilities that reflect a secondary disability occurring once in an individual's lifetime were treated as yearly probability in the analysis, which may have overestimated the costs of secondary disabilities if the SN was not present.

Guideline

The evidence-based guideline for FASD diagnosis was of high quality overall.[26] The guideline had a clearly defined scope and purpose, stakeholder involvement, rigour of development, and evidence of editorial independence.[26] However, although systematic methods were used to search for evidence, the criteria for selecting the evidence were not clearly described.[26] The guideline was also limited by lack of provision of a procedure for updating the guideline, minimal consideration of the potential resource implications of applying the recommendations, and absence of monitoring and/or auditing criteria.[26] Future plans for implementation were mentioned in the guideline, but these implementation efforts were not currently underway at the time of guideline publication in February, 2016.[26]

Summary of Findings

1. *What is the diagnostic test accuracy of tools or tests for the diagnosis and/or assessment of fetal alcohol spectrum disorder in individuals of any age?*

One SR included a literature search for publications related to the strategies to identify and/or diagnose individuals who may have FASD, however no eligible studies were identified.[17] Ten primary studies regarding the accuracy of tools or tests for the diagnosis and/or assessment of FASD were identified.[3,4,7,9,18-23] Most of the studies evaluated diagnostic tests that were still in preliminary stages of development and that are unavailable for clinical use (e.g., regression equations, 3-dimensional face matching algorithms).[3,7,9,19-23] A decision tree model,[18] checklist,[4] test battery (including intelligence, cognitive, motor, and behaviour factors),[19] narrative analyses tool,[21,23] and computer-assisted landmark-based morphometric facial analyses[3,9,20,22] all showed promise as diagnostic tools, in comparison to reference standards of unknown accuracy. In contrast, second trimester fetal ultrasound measures were not able to correctly classify children with FASD.[7] A detailed summary of findings is provided in Appendix 4.

2. *What is the clinical utility of diagnosis and/or assessment of fetal alcohol spectrum disorder in individuals of any age?*

No relevant evidence regarding the clinical utility of diagnosis and/or assessment of FASD in individuals of any age was identified; therefore, no summary can be provided.

3. *What is the clinical effectiveness of the treatment of fetal alcohol spectrum disorder in individuals of any age?*

Three SRs[2,16,17] and one primary study[24] were identified that provided evidence on the clinical effectiveness of the treatment of FASD. In general, the evidence regarding the clinical effectiveness of treatment was limited.

In the first SR, by Reid et al. 2015, 32 primary studies were included, and results could not be pooled due to clinical heterogeneity.[16] Interventions included behavioural treatment, advocacy, or support, and were primarily targeted at individuals with FASD in early to middle childhood. In general, evidence suggested that interventions may provide benefit for the outcomes that are targeted in those with FASD, including self-regulation and attentional control, specific skills such as math knowledge or literacy skills, social skills, and parenting skills.[16] There was also some evidence for the effectiveness of education, support, and advocacy services.[16]

In the second SR, 12 primary studies were included that examined educational, pharmacological, social skills and communication, or behavioural interventions.[2] A meta-analysis could not be performed due to clinical heterogeneity, and there was substantial overlap with the review by Reid et al. 2015[16] (see Appendix 5). In terms of distinct findings, results from two primary studies provided limited evidence for the effectiveness of stimulant medication in decreasing hyperactivity and impulsivity in children with FAS or pFAS and attention-deficit hyperactivity disorder, but no impact on attention.[2]

In the third SR, two SRs and two narrative reviews regarding the clinical effectiveness of treatment of FASD were included.[17] A wide range of treatments were examined including, but not limited to, nursing interventions, case management, education, counselling, advocacy, school-based interventions, cognitive-behavioural strategies, and targeted skill training. The authors concluded that the evidence did not support any particular treatment, but rather that multi-dimensional treatment strategies that are tailored for each individual patient are needed.[17]

The before-and-after clinical effectiveness study found that postural control and sensory attention were worse in children with FASD compared to children with typical development.[24] Following 30 minutes of virtual reality training (Sensorimotor Training to Affect Balance, Engagement and Learning), children in both groups had reduced postural control, which the authors speculated may have been related to fatigue.[24]

4. *What is the cost-effectiveness of diagnosis and/or assessment of fetal alcohol spectrum disorder in individuals of any age?*

One SR included a literature search for publications related to the cost-effectiveness of diagnosis of FASD, however no eligible studies were identified.[17] No other relevant evidence regarding the cost-effectiveness of diagnosis and/or assessment of FASD in individuals of any age was identified; therefore, no summary can be provided.

5. *What is the cost-effectiveness of the treatment of fetal alcohol spectrum disorder in individuals of any age?*

One SR included a literature search for publications related to the cost-effectiveness of treatment of FASD, however no eligible studies were identified.[17]

The single relevant economic evaluation reported that there are substantial costs associated with disabilities that are secondary to FASD, such as crime, homelessness, and mental health problems; with no FASD Service Network (SN) available in Alberta, it was estimated that the secondary disabilities would cost C$22.85 million per year.[25] The cost of running the FASD SN was C$6.12 million per year, and the break-even effectiveness (i.e., effectiveness level at which the SNs became cost-saving) was estimated at 28% (25% to 32%). Therefore, if the program is 28% effective in preventing secondary disabilities, the cost to run the SNs would be offset by the savings in prevented secondary disabilities.[25]

6. *What are the evidence-based guidelines associated with the diagnosis/assessment of and treatment for fetal alcohol spectrum disorder in individuals of any age?*

One evidence-based Canadian guideline was identified that provides recommendations for diagnosis of FASD across the lifespan.[26] In general, the guideline recommends a multi-disciplinary diagnostic approach, involving a complex physical and neurodevelopmental assessment.[26]

According to the guideline, a diagnosis of “FASD with sentinel facial features” may be made if the following criteria are met: prenatal alcohol exposure confirmed or unknown, presentation with three sentinel facial features, and evidence of impairment in at least three of the identified neurodevelopmental domains (or evidence of microcephaly in infants or young children) (see Appendix 4 for details). A diagnosis of “FASD without sentinel facial features” may be made if the following criteria are met: evidence of impairment in at least three of the identified neurodevelopmental domains, and confirmation of prenatal alcohol exposure at a level known to be associated with neurodevelopmental impairment.[26]

It is also recommended that a designation (not diagnosis) of “at risk for neurodevelopmental disorder and FASD” be given to individuals when the criteria for diagnosis are not met, but there is confirmation of prenatal alcohol exposure at a level known to be associated with neurodevelopmental impairment, and there is some evidence of neurodevelopmental disorder with a plausible explanation as to why the full diagnostic criteria for substantial impairment were not met. The “at risk” designation can be considered for individuals with all three sentinel facial features who do not have evidence of the other diagnostic criteria, as long as the absence of prenatal alcohol exposure has not been confirmed.[26]

Note that an important distinction between these guidelines and the previous iteration of the Canadian guidelines is the use of FASD as a diagnostic term.[26] Please see Appendix 4 for the complete recommendations. A diagnostic algorithm (flow-chart) was included in the guideline to facilitate diagnosis.[26]

Limitations

No evidence was identified regarding the clinical utility or cost-effectiveness of diagnosis and/or assessment of FASD. Regarding diagnostic test accuracy, 10 different index tests were compared to eight different reference standards in 10 diagnostic test accuracy studies,[3,4,7,9,18-23] and it is not possible to compare results across studies. Most of the diagnostic tests that were evaluated are still in the preliminary stages of development and are not available for clinical use. In general, there is no consensus on the most appropriate reference standard, and the accuracy of the reference standards used in the diagnostic test accuracy studies is unknown. All diagnostic test accuracy studies were conducted outside of Canada; given that ethnicity may influence the clinical presentation of FASD,[9,22] findings from these studies may not be applicable to the Canadian setting. In addition, most diagnostic test accuracy studies were conducted in those in early to middle childhood[3,7,19-23], however presentation of characteristic facial features that are used in diagnosis diminishes with age[20] and results may not be generalizable to adolescents and adults.

With respect to treatment of FASD, there was substantial variability in the populations (e.g., alcohol exposed, FASD, pFAS, FAS), interventions (e.g., behavioural, pharmacological, educational), comparators (e.g., no treatment, delayed treatment, or placebo), and outcomes (e.g., measures of physical, mental, or cognitive health) in the included systematic reviews[2,16,17] and single primary study.[24] In addition, the diagnostic criteria used for FASD varied across studies and were not always reported, which limits the applicability of findings to other clinical settings.[2] As with the diagnostic test accuracy studies, most studies regarding the clinical effectiveness of treatment of FASD were conducted in those in early to middle childhood. However, the physical, mental, behavioural, and cognitive effects of FASD endure across the lifecycle,[16] and the treatment needs and effects of treatment may be different in older age groups.

No evidence was found regarding specific barriers or access to, or human resources associated with, diagnosis, assessment or treatment of FASD. Similarly, no evidence was identified regarding Indigenous populations in Canada; this is an important gap because Indigenous and non-Indigenous populations may differ with respect to risk factors associated with FASD,[28] prevalence of FASD,[28] and validity of diagnostic tools.[29]

Evidence regarding the cost-effectiveness of treatment of FASD was limited. Only one cost-benefit analysis was identified, and it reported on the break-even effectiveness of a specific FASD Service Network in Alberta, Canada.[25]

Finally, the guidelines provide predominantly strong recommendations for the diagnosis of FASD based on high quality evidence.[26] However, the views and preferences of the target population were minimally considered. In addition, the guidelines recommend a multi-disciplinary diagnostic approach that is likely to be time- and resource-intensive, however the resource implications of applying the recommendations were inadequately considered and advice and/or tools on how to implement the recommendations were not provided.[26]

Conclusions and Implications for Decision or Policy Making

This report identified evidence on the diagnostic accuracy of tools or tests for the diagnosis and/or assessment of FASD, the clinical and cost-effectiveness of treatment, and guidelines regarding the diagnosis of FASD in individuals of any age. No evidence was identified for the clinical utility or cost-effectiveness of diagnosis and/or assessment of FASD in individuals of any age.

Ten studies evaluated the diagnostic accuracy of tests or tools that were still in preliminary stages of development.[3,4,7,9,18-23] Results indicated that a decision tree model,[18] checklist,[4], test battery (including intelligence, cognitive, motor and behaviour factors),[19] narrative analyses tool,[21,23] and computer-assisted landmark-based morphometric facial analysis[3,9,20,22] all showed promise as diagnostic tools, in comparison to reference standards of unknown accuracy.[3,4,7,9,18-23] There is no “gold standard” for the diagnosis of FASD.[3,4]

With respect to clinical effectiveness of treatment, evidence from three SRs[2,16,17] and one primary study[24] provided limited but growing evidence for the effectiveness of a range of interventions. Due to clinical heterogeneity across studies, there was insufficient evidence in favour of any specific intervention for individuals with FASD, however because the effects of FASD are so diverse it was suggested that multi-dimensional treatment strategies that are tailored for each individual patient are needed.[17] These treatment strategies may include physical, behavioural, specific skills, cognitive, pharmacologic, and/or mental health components.[2,16,17] Most interventions targeted individuals with FASD in early to middle childhood, and there is a particular need for intervention studies in adolescents and adults.

Evidence from a single economic study in Alberta, Canada, estimated the break-even effectiveness of an FASD Service Network as 28% (25% to 32%). In other words, if the FASD Service Network is 28% effective in preventing secondary disabilities (such as crime, homelessness, and mental health problems), the cost to run the Service Network would be offset by the savings from prevented secondary disabilities.[25]

One evidence-based Canadian guideline was identified that provides recommendations regarding the diagnosis of FASD.[26] Based on predominantly high-quality evidence, the guidelines provide strong recommendations for multi-disciplinary diagnosis based on

criteria related to facial features, prenatal alcohol exposure, and neurodevelopmental effects.[26]

References

1. Fetal Alcohol Spectrum Disorders (FASDs) [Internet]. Atlanta (GA): Centers for Disease Control and Prevention, National Center on Birth Defects and Developmental Disabilities, Division of Birth Defects and Developmental Disabilities; 2016 Oct [cited 2017 Apr 11]. Available from: https://www.cdc.gov/ncbddd/fasd/index.html

2. Peadon E, Rhys-Jones B, Bower C, Elliott EJ. Systematic review of interventions for children with Fetal Alcohol Spectrum Disorders. BMC Pediatr [Internet]. 2009 [cited 2017 Apr 11] May 25;9:35. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2698825

3. Suttie M, Foroud T, Wetherill L, Jacobson JL, Molteno CD, Meintjes EM, et al. Facial dysmorphism across the fetal alcohol spectrum. Pediatrics [Internet]. 2013 Mar [cited 2017 Apr 11];131(3):e779-e788. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3581841

4. Burd L, Klug MG, Li Q, Kerbeshian J, Martsolf JT. Diagnosis of fetal alcohol spectrum disorders: a validity study of the fetal alcohol syndrome checklist. Alcohol. 2010 Nov;44(7-8):605-14.

5. Evaluation of the fetal alcohol spectrum disorder (FASD) initiative 2008-2009 to 2012-2013 [Internet]. Ottawa (ON): Public Health Agency of Canada; 2014 [cited 2017 Apr 11] Mar. Available from: http://www.phac-aspc.gc.ca/about_apropos/evaluation/reports-rapports/2013-2014/efasdi-eietaf/index-eng.php

6. Chudley AE, Conry J, Cook JL, Loock C, Rosales T, LeBlanc N, et al. Fetal alcohol spectrum disorder: Canadian guidelines for diagnosis. CMAJ [Internet]. 2005 [cited 2017 Apr 11] Mar 1;172(5 Suppl):S1-S21. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC557121

7. Montag AC, Hull AD, Yevtushok L, Zymak-Zakutnya N, Sosyniuk Z, Dolhov V, et al. Second-trimester ultrasound as a tool for early detection of fetal alcohol spectrum disorders. Alcohol Clin Exp Res. 2016 Nov;40(11):2418-25.

8. 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 Aug;25(4):228-38.

9. Fang S, McLaughlin J, Fang J, Huang J, utti-Ramo I, Fagerlund A, et al. Automated diagnosis of fetal alcohol syndrome using 3D facial image analysis. Orthod Craniofac Res [Internet]. 2008 Aug [cited 2017 Apr 11];11(3):162-71. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3428207

10. FASD fact sheet [Internet]. Vancouver (BC): CanFASD; 2010 [cited 2017 Apr 11]. Available from: https://canfasd.ca/media/fasd-fact-sheet/

11. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol [Internet]. 2007 [cited 2017 Apr 11];7:10. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1810543/pdf/1471-2288-7-10.pdf

12. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health [Internet]. 1998 [cited 2017 Apr 11] Jun;52(6):377-84. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1756728/pdf/v052p00377.pdf

13. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011 Oct 18;155(8):529-36.

14. Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions [Internet]. Version 5.1.0. London (England): The Cochrane Collaboration; 2011 Mar [cited 2017 Apr 11] Figure 15.5.a: Drummond checklist (Drummond 1996). Available from: http://handbook.cochrane.org/chapter_15/figure_15_5_a_drummond_checklist_drummond_1996.htm

15. Brouwers M, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in healthcare. CMAJ [Internet]. 2010 [cited 2017 Apr 11]Dec;182(18):E839-E842. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3001530/pdf/182e839.pdf

16. Reid N, Dawe S, Shelton D, Harnett P, Warner J, Armstrong E, et al. Systematic review of fetal alcohol spectrum disorder interventions across the life span. Alcohol Clin Exp Res. 2015 Dec;39(12):2283-95.

17. Elliott L, Coleman K, Suebwongpat A, Norris S. Fetal alcohol spectrum disorders (FASD): systematic reviews of prevention, diagnosis and management. HSAC Report [Internet]. 2008 Nov [cited 2017 Apr 11];1(9). Available from: http://www.healthsac.net/downloads/publications/HSAC07_FASD_FINALv3.pdf

18. Goh PK, Doyle LR, Glass L, Jones KL, Riley EP, Coles CD, et al. A decision tree to identify children affected by prenatal alcohol exposure. J Pediatr [Internet]. 2016 [cited 2017 Apr 11]Oct;177:121-7. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5291174

19. Kalberg WO, May PA, Blankenship J, Buckley D, Gossage JP, Adnams CM. A practical testing battery to measure neurobehavioral ability among children with FASD. Int J Alcohol Drug Res [Internet]. 2013 Nov 26 [cited 2017 Apr 11];2(3):51-60. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4170949

20. Mutsvangwa TE, Meintjes EM, Viljoen DL, Douglas TS. Morphometric analysis and classification of the facial phenotype associated with fetal alcohol syndrome in 5- and 12-year-old children. Am J Med Genet A. 2010 Jan;152A(1):32-41.

21. Thorne JC, Coggins T. A diagnostically promising technique for tallying nominal reference errors in the narratives of school-aged children with foetal alcohol spectrum disorders (FASD). Int J Lang Commun Disord. 2008 Sep;43(5):570-94.

22. Moore ES, Ward RE, Wetherill LF, Rogers JL, Autti-Rämö I, Fagerlund A, et al. Unique facial features distinguish fetal alcohol syndrome patients and controls in diverse ethnic populations. Alcohol Clin Exp Res. 2007 Oct;31(10):1707-13.

23. Thorne JC, Coggins TE, Carmichael OH, Astley SJ. Exploring the utility of narrative analysis in diagnostic decision making: picture-bound reference, elaboration, and fetal alcohol spectrum disorders. J Speech Lang Hear Res. 2007 Apr;50(2):459-74.

24. McCoy SW, Jirikowic T, Price R, Ciol MA, Hsu LY, Dellon B, et al. Virtual sensorimotor balance training for children with fetal alcohol spectrum disorders: feasibility study. Phys Ther [Internet]. 2015 Nov [cited 2017 Apr 11];95(11):1569-81. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4625061

25. Thanh NX, Moffatt J, Jacobs P, Chuck AW, Jonsson E. Potential impacts of the Alberta fetal alcohol spectrum disorder service networks on secondary disabilities: a cost-benefit analysis. J Popul Ther Clin Pharmacol [Internet]. 2013 [cited 2017 Apr 11];20(2):e193-e200. Available from: http://www.jptcp.com/articles/potential-impacts-of-the-alberta-fetal-alcohol-spectrum-disorder-service-networks-on-secondary-disabilities-a-costbenefit-analysis.pdf

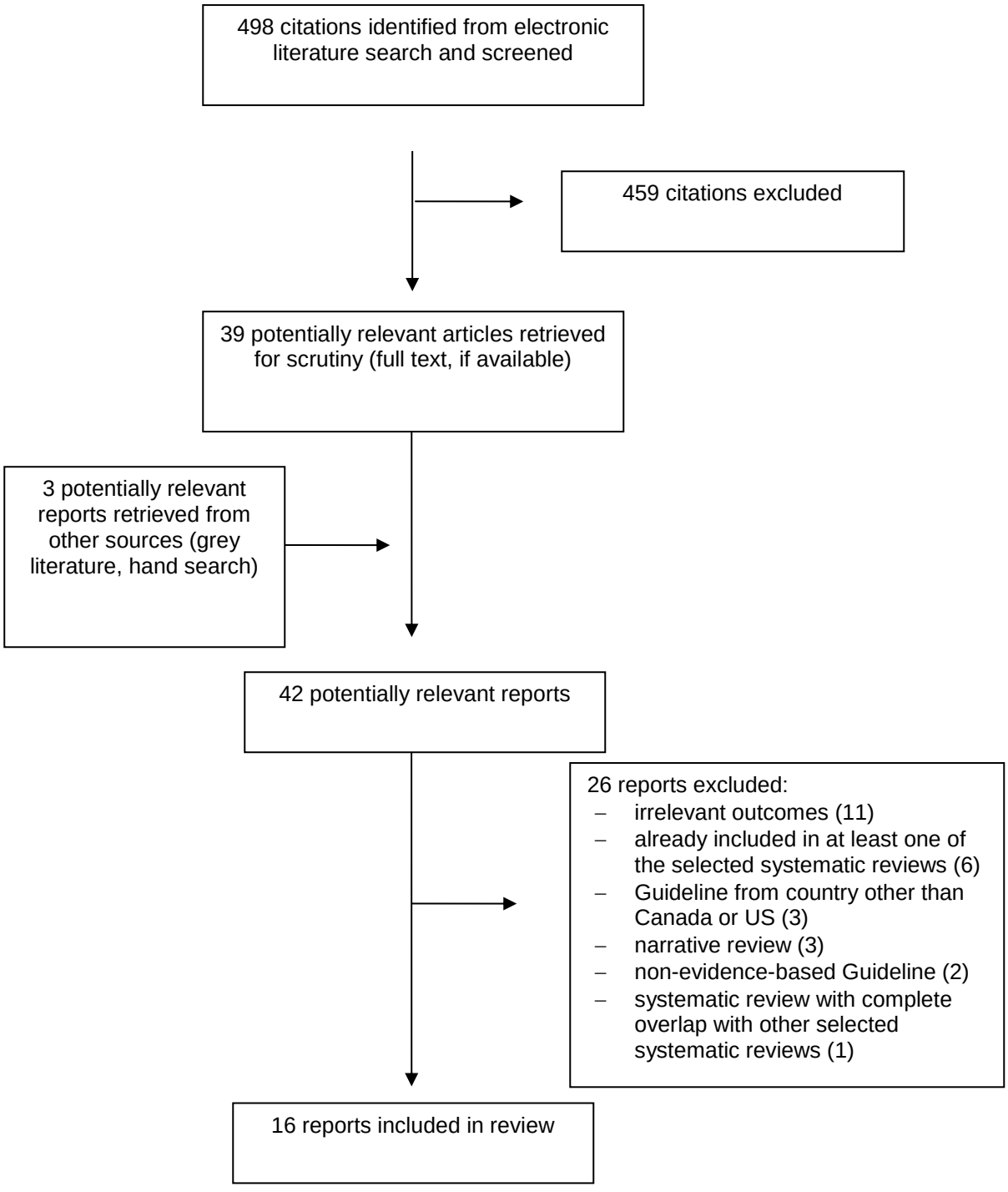
26. Cook JL, Green CR, Lilley CM, Anderson SM, Baldwin ME, Chudley AE, et al. Fetal alcohol spectrum disorder: a guideline for diagnosis across the lifespan. CMAJ [Internet]. 2016 Feb 16 [cited 2017 Apr 11];188(3):191-7. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4754181

27. Cook JL, Green CR, Lilley CM, Anderson SM, Baldwin ME, Chudley AE, et al. Fetal alcohol spectrum disorder: a guideline for diagnosis across the lifespan [supplemental appendix]. CMAJ [Internet]. 2016 Feb 16 [cited 2017 Apr 11];188(3):191-7. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4754181

28. Pacey M. Fetal alcohol syndrome and fetal alcohol spectrum disorder among Aboriginal Peoples: a review of prevalence [Internet]. Prince George (BC): National Collaborating Centre for Aboriginal Health; 2009 Sep [cited 2017 Apr 11]. Available from: http://www.nccah-ccnsa.ca/Publications/Lists/Publications/Attachments/34/2010_01_27_FASDAboriginalReviewPrevalence_EN_Updated_Web.pdf

29. Tait CL. Fetal alcohol syndrome among Aboriginal People in Canada: review and analysis of the intergenerational links to residential schools [Internet]. Ottawa (ON): Aboriginal Healing Foundation; 2014 [cited 2017 Apr 11]. Available from: http://www.phac-aspc.gc.ca/about_apropos/evaluation/reports-rapports/2013-2014/efasdi-eietaf/index-eng.php

Appendix 1: Selection of Included Studies



Appendix 2: Characteristics of Included Publications

Table A1: Characteristics of Included Systematic Reviews

Author, Publication Year, Country	Types and Numbers of Primary Studies Included	Population Characteristics	Intervention(s)	Comparator(s)	Clinical Outcomes, Length of Follow-up

Reid, et al. 2015[16]
 	
Australia
	32 primary studies included:
	
	CCT, n = 16;
	case study, n = 4;
	cohort, n = 4;
	RCT, n = 3;
	retrospective case-file analysis, n = 3;
	cohort analytic, n = 1;
	interrupted time series, n = 1
		Individuals of any age with PAE, including FAS, pFAS, ARND, or PAE	Behavioural treatment, advocacy or support	Not specified	Quantitative measures of functioning: developmental; self-regulation and attentional control; specific skills; social skills; parenting skills; support, education and advocacy; supporting parents who have FASD.
					
					Follow-up unclear, or ranged from 1 week to 3 years

Peadon, et al. 2009[2] **Australia**	12 primary studies included: RCT, n = 6; pre-post intervention, n = 4; CCT, n = 1; Quasi-RCT, n = 1	Individuals under age 18 y with a diagnosis of a FASD	Behavioural, speech, occupational, physiotherapies, early intervention programmes, psychosocial and educational, pharmacological	Control group (no treatment, waiting list, usual therapy, placebo), or pre- and post-intervention measurements	Measures of physical and mental health, developmental or cognitive status, quality of life, educational attainment, employment, contact with the law, and substance abuse Follow-up unclear, or ranged from day 3 of intervention to 10 months
Elliott, et al. 2008[17] **Australia and New Zealand**	Diagnosis: None eligible Economic: None eligible Management: Systematic reviews, n = 2 Narrative reviews, n = 2	Individuals who may have FASD or mothers of individuals who may have FASD	Diagnosis: Any strategy to identify and/or diagnose an individual who may have FASD Economic: Any strategy to reduce the financial burden of FASD Management: Any strategy to improve clinical outcomes in individuals with FASD	Any comparator	Diagnosis: Sensitivity and specificity of FASD diagnosis Economic: Cost-effectiveness of strategies to reduce FASD Management: Reduction in severity of primary and/or secondary disabilities or deficits associated with FASD

ARND = alcohol-related neurodevelopmental disorder; CCT = controlled clinical trial; FAS = fetal alcohol syndrome; FASD = fetal alcohol spectrum disorder; PAE = prenatal alcohol exposure; pFAS = partial fetal alcohol syndrome; RCT = randomized controlled trial.

Table A2: Characteristics of Included Diagnostic Accuracy Studies

First Author, Publication Year, Country	Study Design	Patient Characteristics	Reference Standard	Index Test	Outcomes
Goh, et al. 2016[18] **United States**	Case-control	N = 434 Children (5-7 y), n = 165 Mean age: 6.6 y AE, n = 55 Non-AE, n = 110 Adolescents (10-16 y), n = 289 Mean age: 13.4 y AE, n = 98 Non-AE, n = 191	Prenatal AE determined by knowledge of exposure history (>13 drinks/wk or >4 drinks/occasion during pregnancy; or suspected such exposure in a child with an FAS diagnosis) obtained through retrospective maternal report, social service, legal or medical records	Decision tree model including the following: CBCL score, IQ, physical exam for features of FAS, VABS-II score, physical exam for extended features Two routes through the decision tree were tested: 1) *psychologist route*: patient presents with neurobehavioral problems 2) *pediatrician route*: patient presents with suspicion of prenatal AE or concerning physical or behaviour features	Classification accuracy, sensitivity, specificity, PPV, NPV
Montag, et al. 2016[7] **Ukraine**	Longitudinal prospective cohort	233 mother/child dyads; pregnant women were assigned to "alcohol-exposed" and "no-or-low-alcohol exposure" groups 1:1 recruitment ratio with sequential enrolment (after a woman enrolled in the alcohol-exposed group, the next no- or low-exposure woman was invited) Infant diagnosis based on reference standard: FAS, n = 4 FASD, n = 11	Dysmorphological examination (at birth) and neurobehavioral evaluation using the BSID-II (at approximately 6 and/or 12 mo of age) *FAS classification*: prenatal AE, at least 2 key facial features plus growth deficiency or microcephaly or both *FASD classification:* prenatal AE, 1 key facial feature plus growth deficiency and microcephaly, or a BSID-II score <85 and 1 key facial feature, 1	Second trimester fetal ultrasound with measures of TCD, OFD, CCD, FTD, IOD, OOD, and OD	Proportion of variance in FASD accounted for by combined and individual ultrasound measures

First Author, Publication Year, Country	Study Design	Patient Characteristics	Reference Standard	Index Test	Outcomes
			growth abnormality plus at least 1 additional alcohol-related feature		
Kalberg, et al. 2013[19] **South Africa**	Case-control	N = 113 FASD, n = 61 Mean age ± SD: 78.5 ± 7.6 mo (including: FAS, n = 37; pFAS, n = 16; ARND, n = 8) Control, n = 52 Mean age ± SD: 88.3 ± 11.0 mo	Diagnosis made at a "case conference based on dysmorphology exams, developmental testing, and maternal interviews documenting alcohol exposure and other risk factors" p. 5	Regression model developed from a test battery including: intelligence tests, perceptual motor tests, planning test, memory tests (spatial, short term, long term, working), and a behaviour checklist	Classification accuracy
Suttie, et al. 2013[3] **South Africa**	Case-control	N = 192 FAS, n = 22 pFAS, n = 26 Nonsyndromal HE, n = 75 Non-AE, n = 69 Range of mean ages ± SD across groups: 10.0 ± 1.5 y to 10.6 ± 2.4 y	Diagnosis by two expert dysmorphologists, blinded to prenatal AE history, using a standard protocol and the Astley Lip-Philtrum Guide14	Dense surface modeling and signature analyses of 3-dimensional facial photographs	Agreement between clinical categorization and classifications from face shape alone
Burd, et al. 2010[4] **United States**	Case-control	N = 658 FAS, n = 152; Other-FASD (pFAS or ARND), n = 167; No-FASD, n = 339 Mean age not reported Note: the "no-FASD" group included those referred for evaluation with ADHD, chromosomal abnormalities, various syndromes, familial	Clinical diagnosis in genetic/ dysmorphic clinics	FASDC	Accuracy, sensitivity, specificity Optimal set of checklist variables for diagnostic categorization

First Author, Publication Year, Country	Study Design	Patient Characteristics	Reference Standard	Index Test	Outcomes
		neuropsychoatric disorders, or those who were evaluated for but not diagnosed with FASD			
Mutsvangwa, et al. 2010[20] **South Africa**	Case-control	N = 34 FAS, n = 17 Age 5 y, n = 4 Age 12 y, n = 13 Controls, n = 17 Age 5 y, n = 11 Age 12 y, n = 6	Clinical diagnosis of FAS based on "sufficient dysmorphology", falling approximately 2 SDs below the mean on either verbal or non-verbal IQ tests, and having "substantial" behavioural problems and confirmation of prenatal AE	Landmark-based morphometric analysis of facial phenotype using "leave-one-out cross-validation"	Classification accuracy, sensitivity, specificity
Fang, et al. 2008[9] **Finland and South Africa**	Case-control	N = 149 FAS, n = 86 Control, n = 63 Two ethnic samples: FC, CC Age range: 2.8 to 21 y	Clinical diagnosis based on structural features and growth deficiency, consistent with the revised IOM criteria	Computer algorithm classification based on 3-dimensional facial image analysis	Classification accuracy, sensitivity, specificity
Thorne and Coggins 2008[21] **United States**	Retrospective case-control	N = 32 FASD, n = 16 TD, n = 16 Mean age: 9.11 y	Clinical diagnosis using the Four-Digit Diagnostic Code	rNRE from oral narrative analysis *or* Rate of ANR/TW calculated with the SECS system of analysis	AUC, classification accuracy, number of false positives, number of false negatives
Moore, et al. 2007[22] **Finland, South Africa, United States**	Case-control	N = 276 FAS, 43% Control, 57% Divided into 4 groups based on ancestry: CC, FC, AA, NAC	Clinical diagnosis based on structural features and growth deficiency, consistent with the revised IOM criteria	Computerized anthropometry based on laser-obtained 3-dimensional facial images	Classification accuracy, sensitivity, specificity

First Author, Publication Year, Country	Study Design	Patient Characteristics	Reference Standard	Index Test	Outcomes
Thorne, et al. 2007[23] **United States**	Case-control	N = 32 FASD, n = 16 TD, n = 16 Mean age: 9.11 y	Clinical diagnosis using the Four-Digit Diagnostic Code	Rate of ANR/TW calculated with the SECS system of analysis	ROC cut-point, AUC, sensitivity, specificity, efficiency, number correctly classified

AA = African American; ADHD = attention-deficit hyperactivity disorder; AE = alcohol exposure; ANR/TW = rate of ambiguous nominal reference errors; ARND = alcohol-related neurodevelopmental deficits; AUC = area under the (receiver-operating characteristic) curve; BSID-II = Bayley Scales of Infant Development, 2nd ed.; CBCL = Child Behavior Checklist; CC = Cape Coloured; CCD = caval-calvarial distance; FAS = fetal alcohol syndrome; FASD = fetal alcohol spectrum disorder; FASDC = Fetal Alcohol Syndrome Diagnostic Checklist; FC = Finnish Caucasian; FTD = frontothalamic distance; HE = heavily exposed; IOD = interorbital distance; IOM = Institute of Medicine; IQ = intelligence quotient; NAC = North American Caucasian; NPV = negative predictive value; OD = orbital diameter; OFD = occipital frontal diameter; OOD = outer orbital diameter; pFAS = partial fetal alcohol syndrome; PPV = positive predictive value; rNRE = rate of nominal reference errors; SD = standard deviation; SECS = Semantic Elaboration Coding System; TCT = transverse cerebellar diameter; TD = typically developing; VABS-II = Vineland Adaptive Behavior Scales-II; y = years.

Table A3: Characteristics of Included Clinical Effectiveness Studies

First Author, Publication Year, Country	Study Design	Patient Characteristics	Intervention(s)	Comparator(s)	Clinical Outcomes
Westcott McCoy, et al. 2015[24] **United States**	Before-and-after study	Children with FASD, diagnosed using the 4-digit diagnostic code, n = 11 Mean age ± SD: 137.0 ± 23.0 mo Children with typical development, n = 11 Mean age ± SD: 133.5 ± 26.9 mo Age range: 8 to 16 y	30 min of STABEL training (a virtual reality system to train sensory control for balance); 3 x 6 min blocks of training interspersed with rest	Measurement of outcomes pre-STABEL training	Sensory attention (as assessed by sensory attention fraction and entrainment gain) and postural control (anterior-posterior and medial-lateral postural sway velocity) post-STABEL training (assessed with the MuMBER system)

FASD = fetal alcohol spectrum disorder; mo = months; MuMBER = Multimodal Balance Entrainment Response; SD = standard deviation; STABEL = Sensorimotor Training to Affect Balance, Engagement and Learning.

Table A4: Characteristics of Included Economic Studies

First Author, Publication Year, Country	Type of Analysis, Perspective, Time Horizon	Study Population	Intervention, Comparator, Outcomes	Main Assumptions
Thanh, et al. 2013[25] Canada	Cost-benefit analysis; break-even effectiveness (i.e., the effectiveness level at which the intervention became cost-saving) Societal perspective 1-year time horizon	High-risk individuals who were referred to the SN for diagnostic services, and those who were diagnosed with FASD outside of the SN but were referred to the SN for support, from 2008/9 to 2010/11 N = 1,275 Adults, n = 471 Children, n = 804	12 Alberta FASD SNs (purpose: diagnose FASD, provide support for those with FASD, raise awareness, and provide FASD prevention interventions) vs. no SNs Outcomes: cost of secondary disabilities (crime, homelessness, mental health problems, school disruption [< 18 y old] or unemployment [≥ 18 y]) with the absence of the intervention; cost of the intervention; break-even effectiveness	• Individuals that leave the SN are at the same risk for secondary disabilities as those in the scenario where the SN is not present • Probability of staying in the SN for women who participate in the prevention component of the SN was used as a proxy for probability of staying in the SN • Probabilities of secondary disabilities were retrieved from a systematic literature search and converted to 1-year probabilities if required • Probabilities that reflect a secondary disability occurring once in an individual's lifetime were treated as yearly probability in the analysis • Probability of homelessness for people with FASD was not available, so the occurrence of being homeless among people with mental illness was used as a proxy • Probability of having FASD among high-risk individuals referred to the SN was assumed to be 67.5% • Productivity cost of unemployment was the average wage per person per 2 months in Alberta in 2011 • Secondary disabilities were treated as mutually exclusive (to avoid double-counting) • Criminal justice cost of crime per homeless person was used as a proxy for criminal justice cost of crime per person with FASD

FASD = fetal alcohol spectrum disorder; SN = Service Network.

Table A5: Characteristics of Included Guidelines

Objectives			Methodology			
Target Population, Intended Users	Intervention and Practice Considered	Major Outcomes Considered	Evidence Collection, Selection and Synthesis	Evidence Quality Assessment	Recommendations Development and Evaluation	Guideline Validation
Cook, et al. 2016[26b]						
Target population: Pregnant or postpartum women and individuals at risk of FASD **Intended users:** Multi-disciplinary diagnostic teams in Canada, who have expertise through specialized training and experience	Multidisciplinary diagnostic process	Not specified	Electronic database searches, online survey administered to 35 diagnostic centres across Canada, 6 focus groups with national and international experts	Evidence rated using the GRADE approach	Expert consensus based on review of literature and strength of recommendation assigned according to GRADE	Internal and external peer review

FASD = fetal alcohol spectrum disorder; GRADE = Grading of Recommendations Assessment, Development and Evaluation.

Appendix 3: Critical Appraisal of Included Publications

Table A6: Strengths and Limitations of Systematic Reviews using AMSTAR[11]

Strengths	Limitations
Reid, et al. 2015[16]	
• Research question and inclusion criteria published prior to conduct of study • Comprehensive literature search performed, including database searching and a systematic search for non-peer-reviewed intervention trials • List of included studies provided • Study selection and data extraction performed by two independent reviewers • Some key characteristics of included studies provided (e.g., sample size, age, approach, follow-up) • Scientific quality of included studies assessed, using the Effective Public Health Project assessment tool, and used appropriately in formulating conclusions • Clinical appropriateness of combining studies was considered and results were not pooled • Review authors declared no conflict of interest	• List of excluded studies not provided • Consensus procedure for disagreements in study selection and data extraction not described • Some key characteristics of included studies not provided (e.g., age at diagnosis, socioeconomic data, comorbidities) • Inclusion of low-quality study designs, such as case studies and pilot studies that were underpowered to detect intervention effects • No numerical data were presented in the narrative synthesis • Statistical significance not consistently clear or explicit • Conclusions based on statistical significance; clinical significance not considered • Narrative summary based on direction of association, with no consideration of clinical significance • No assessment of publication bias • Conflict of interest not reported for the included studies
Peadon, et al. 2009[2]	
• Comprehensive literature search performed, including database searching, hand-searching (reference lists, review articles, conference proceedings), and contacting individuals conducting research on FASD • List of included studies provided • Study selection and data extraction performed by two independent reviewers, with a third reviewer resolving conflicts • Some key characteristics of included studies provided (e.g., sample size, age, type of FASD diagnosis) • Some characteristics affecting scientific quality of included studies assessed (e.g., blinding of outcome assessment, use of standardized measures); however, assessment was not comprehensive • Clinical appropriateness of combining studies was considered and results were not pooled • Review authors declared no conflict of interest	• No reference to a protocol, ethics approval, or pre-determined research objectives to indicate that the research question and inclusion criteria were established *a priori* • No formal grey literature search conducted (limited to hand-searching) • List of excluded studies not provided • Some key characteristics of included studies not provided (e.g., age at diagnosis, socioeconomic data, comorbidities) • Some characteristics of study quality assessed (method of randomization, allocation concealment, intention-to-treat analysis, blinding of outcome assessment, use of standardized measures, follow-up), but no overall summary of scientific quality of included studies provided • Conclusions based on statistical significance; clinical significance not considered • Scientific quality of the included studies was not adequately considered in the analysis and conclusions of the review • No assessment of publication bias • Conflict of interest not reported for the included studies
Elliott, et al. 2008[17]	
• Research questions and inclusion criteria were developed prior to the conduct of the review • Comprehensive literature search performed, including database searches and hand-searching (health technology assessment websites, reference lists, selection of relevant journals, conference abstracts) • Lists of included and excluded studies provided	• No formal grey literature search conducted • Study selection was performed by a single reviewer and was not double-checked by a second reviewer • Some key characteristics of included studies not provided (e.g., age at diagnosis, socioeconomic data, comorbidities) • Conclusions based on statistical significance; clinical significance not considered

Strengths	Limitations
• Data extraction and critical appraisal were performed by one reviewer and checked by another • Some key characteristics of included studies provided (e.g., sample size, age, type of FASD diagnosis) • Scientific quality of the included studies was used appropriately in formulating conclusions	• Strengths and limitations of included studies mentioned in narrative summary, but no formal assessment or documentation of the scientific quality of the included studies • Publication bias was not formally considered • Conflict of interest was not reported for the systematic review or the included studies

FASD = fetal alcohol spectrum disorder.

Table A7: Strengths and Limitations of Diagnostic Test Accuracy Studies using QUADAS-2[13]

Strengths	Limitations
Goh, et al. 2016[18]	
• "Difficult-to-diagnose" patients purposefully included • No inappropriate exclusion criteria reported • The decision tree (index test) was developed in one sample and then tested in an independent sample of patients • All patients received the same reference standard • All patients were included in the analysis	• Case-control design may exaggerate diagnostic accuracy • Method of participant sampling unclear • Unclear accuracy of reference standard to correctly classify AE • Unclear whether the decision tree (index test) was completed independently of knowledge of the results of the reference test
Montag, et al. 2016[7]	
• Appropriate method of participant sampling; participants were sampled in 1:1 fashion such that after each AE woman was enrolled the next non-AE woman was invited to participate • No inappropriate exclusion criteria reported • The index test results were interpreted without knowledge of the results of the reference standard; ultrasonographers were blinded to participants' prenatal AE status • The reference standard results were interpreted without knowledge of the results of the index test; the clinicians (dysmorphologists/geneticists and neurobehavioral examiners) who performed the reference standard tests were blinded to the index test results • There was an appropriate interval between the index test and reference standard; although in general the index test and reference standard are performed at the same time, in this case the purpose was to evaluate whether a prenatal test could be used diagnostically • All patients received the same reference standard • All patients were included in the analysis	• Applicability concerns: FASD is usually diagnosed at or after birth, however the index test was conducted prenatally • Unclear test accuracy of reference standard to correctly classify FASD • Classification of mothers into the "alcohol-exposed" and "no-or-low-alcohol exposure" groups was done via self-reported exposure data, which could have contributed to misclassification of exposure groups • Index test not available for external/clinical use
Kalberg, et al. 2013[19]	
• All patients received the same reference standard • Unclear whether index test results were interpreted without knowledge of the results of the reference standard, however because the index test is objective this is unlikely to introduce bias	• Case-control design may exaggerate diagnostic accuracy • Unclear whether method of participant sampling was appropriate; control children were initially randomly selected, but an additional 25 children "who met age and sex criteria" (p. 4) were then added to the sample • Possible inappropriate exclusions; control children whose

Strengths	Limitations
	mothers were found to have consumed alcohol during pregnancy were excluded • Unclear test accuracy of reference standard to correctly classify FASD • The index test (regression model) was tested in the same sample in which it was developed • Only the percent correctly classified for the total sample was reported; the percent correctly classified for each group, and the percent misclassified, were not reported • The test battery did not include a test of mathematical ability or a comprehensive measure of behaviour, which would be expected to be important for discriminating between children with or without FASD • Index test not available for external/clinical use
Suttie, et al. 2013[3]	
• Unclear whether index test results were interpreted without knowledge of the results of the reference standard, however because the index test is objective this is unlikely to introduce bias • All patients received the same reference standard	• Case-control design may exaggerate diagnostic accuracy • Unclear whether method of participant sampling was appropriate; participants were recruited from 2 longitudinal cohorts and identified by screening children in a school • Unclear whether inappropriate exclusions were avoided • Not all patients were included in the analysis; in the face signature analysis, 16 individuals with FAS or pFAS were omitted “due to insufficient controls for normalization” p. e784 • Unclear test accuracy of reference standard to correctly classify pFAS and FAS • The index test (facial analysis) was tested in the same sample in which it was developed • Index test not available for external/clinical use
Burd, et al. 2010[4]	
• Appropriate method of participant selection; participants were consecutive patients seen for diagnostic evaluation at genetic and dysmorphology clinics; however, enrollment took place over a 19 y span • Inappropriate exclusions were avoided • Unclear whether index test results were interpreted without knowledge of the results of the reference standard, however because the index test is objective this is unlikely to introduce bias	• Case-control design may exaggerate diagnostic accuracy • Data were collected over a long period of time (1984 to 2003) and it is possible that the reference standard (clinical diagnosis) changed over time; unclear if all patients received the same reference standard • Unclear test accuracy of reference standard to correctly classify FAS, other-FASD, or non-FASD
Mutsvangwa, et al. 2010[20]	
• Unclear whether index test results were interpreted without knowledge of the results of the reference standard, however because the index test is objective this is unlikely to introduce bias • All patients received the same reference standard • All patients were included in the analysis	• Case-control design may exaggerate diagnostic accuracy • Method of participant sampling unclear • Unclear whether inappropriate exclusions were avoided • Small, unbalanced sample • Unclear test accuracy of reference standard to correctly classify FAS • The index test (facial analysis) was tested in the same sample in which it was developed • Index test not available for external/clinical use

Strengths	Limitations
Fang, et al. 2008[9]	
• Unclear whether the index test results were interpreted without knowledge of the results of the reference standard, however because the index test is automated this is unlikely to introduce bias • All patients received the same reference standard • All patients were included in the analysis	• Case-control design may exaggerate diagnostic accuracy • Method of participant sampling unclear • Inappropriate exclusions; only participants designated as clearly "FAS" or "no FAS" were included (i.e., "difficult-to-diagnose" patients on the FASD spectrum were excluded, which would be expected to inflate the test accuracy) • Unclear test accuracy of reference standard to correctly classify FAS • The index test (facial analysis) was tested in the same sample in which it was developed • Index test not available for external/clinical use
Thorne and Coggins 2008[21]	
• Unclear whether the index test results were interpreted without knowledge of the results of the reference standard, however because the index test is objective this is unlikely to introduce bias • All patients were included in the analysis	• Case-control design may exaggerate diagnostic accuracy • Method of participant sampling unclear • Not all patients received the same reference standard; TD participants did not undergo the same clinical assessment (reference standard) as the children with FASD • Inappropriate exclusions; only children with clearly impaired language performance deficits were included (i.e., "difficult-to-diagnose" patients were excluded, which would be expected to inflate the test accuracy) • Unclear test accuracy of reference standard to correctly classify FASD • Index test not available for external/clinical use
Moore, et al. 2007[22]	
• Unclear whether the index test results were interpreted without knowledge of the results of the reference standard, however because the index test is objective this is unlikely to introduce bias • All patients received the same reference standard • All patients were included in the analysis	• Case-control design may exaggerate diagnostic accuracy • Method of participant sampling unclear • Inappropriate exclusions; only participants designated as clearly "FAS" or "no FAS" were included (i.e., "difficult-to-diagnose" patients on the FASD spectrum were excluded, which would be expected to inflate the test accuracy) • The index test (facial analysis) was tested in the same sample in which it was developed • Index test not available for external/clinical use
Thorne, et al. 2007[23]	
• No inappropriate exclusion criteria reported • Unclear whether the index test results were interpreted without knowledge of the results of the reference standard, however because the index test is objective this is unlikely to introduce bias • All patients were included in the analysis	• Case-control design may exaggerate diagnostic accuracy • Method of participant sampling unclear • Not all patients received the same reference standard; TD participants did not undergo the same clinical assessment (reference standard) as the children with FASD • Unclear test accuracy of reference standard to correctly classify FASD • Index test not available for external/clinical use

AE = alcohol exposure; FAS = fetal alcohol syndrome; FASD = fetal alcohol spectrum disorder; pFAS = partial fetal alcohol syndrome; TD = typically developing.

Table A8: Strengths and Limitations of Clinical Effectiveness Studies using the Downs and Black Checklist[12]

Strengths	Limitations
Westcott McCoy, et al. 2015[24]	
• Study objective clearly described • Participant inclusion and exclusion criteria provided • Characteristics of the included participants clearly described • Intervention, comparator, outcomes, and main findings clearly described • Statistical estimates of variability in the data provided for the main findings • Actual probability values (p-values) reported • No retrospective, unplanned analyses reported • Appropriate statistical tests used to assess main outcomes	• A convenience sample was used; children with FASD were recruited from a clinical registry and children with TD were recruited from the community • There was no time-control group (i.e., repeat testing without STABEL training to account for possible learning effects on the outcome measures) • No attempt was made to blind the examiner to group (FASD or TD), or to pre- or post-STABEL training status • Validity and reliability of the main outcome measures are mentioned, but not reported, as “unpublished data” • Power calculation not performed • Small sample size and limited external validity

FASD = fetal alcohol spectrum disorder; STABEL = Sensorimotor Training to affect Balance, Engagement and Learning; TD = typical development.

Table A9: Strengths and Limitations of the Economic Study using Drummond[14]

Strengths	Limitations
Thanh, et al. 2013[25]	
• Clearly stated research question, economic importance of research question, intervention and comparator, outcome measures, analysis perspective, and form of economic evaluation used • Choice of form of economic evaluation justified in relation to the question addressed • Methods of estimation of numbers and costs of secondary disabilities provided • Costs of the intervention were measured directly; costs of the secondary disabilities were retrieved from Alberta or Canadian sources • The currency and price date and details of adjustment for inflation or currency conversion were provided • Sufficient details provided regarding the analytic model • Main results clearly presented; conclusions followed from the data reported and were accompanied by important caveats	• Model components were all estimated from the literature, and not based on effectiveness studies • Model assumptions may not be valid (e.g., the probability of individuals with FASD staying in the SN may be different than the probability of staying in the SN for women participating in the prevention component) • Productivity changes (occurrences and costs of unemployment) were not reported separately • Probabilities that reflect a secondary disability occurring once in an individual’s lifetime were treated as yearly probability in the analysis; this may have overestimated costs of secondary disabilities if the SN was not present • A 1-year time horizon was used, and actual 1-year costs of the SN were used, but the study population included individuals who were referred to the SN over a 3-year time span • A range for the break-even effectiveness, identified via sensitivity analysis, was reported; however, confidence intervals around the main study variables were not consistently reported • For sensitivity analysis, the ranges over which variables varied were reported, however insufficient details of the approach used were provided

FASD = fetal alcohol spectrum disorder; SN = Service Network.

Table A10: Strengths and Limitations of Guidelines using AGREE II[15]

Item	Guideline
	Cook et al. 2016[26]
Domain 1: Scope and Purpose	
1. The overall objective(s) of the guideline is (are) specifically described.	✓
2. The health question(s) covered by the guideline is (are) specifically described.	✓
3. The population (patients, public, etc.) to whom the guideline is meant to apply is specifically described.	✓
Domain 2: Stakeholder Involvement	
4. The guideline development group includes individuals from all relevant professional groups.	✓
5. The views and preferences of the target population (patients, public, etc.) have been sought.	✓
6. The target users of the guideline are clearly defined.	✓
Domain 3: Rigour of Development	
7. Systematic methods were used to search for evidence.	✓
8. The criteria for selecting the evidence are clearly described.	X
9. The strengths and limitations of the body of evidence are clearly described.	✓
10. The methods for formulating the recommendations are clearly described.	✓
11. The health benefits, side effects, and risks have been considered in formulating the recommendations.	✓
12. There is an explicit link between the recommendations and the supporting evidence.	✓
13. The guideline has been externally reviewed by experts prior to its publication.	✓
14. A procedure for updating the guideline is provided.	X
Domain 5: Applicability	
18. The guideline describes facilitators and barriers to its application.	✓
19. The guideline provides advice and/or tools on how the recommendations can be put into practice.	X
20. The potential resource implications of applying the recommendations have been considered.	X
21. The guideline presents monitoring and/or auditing criteria.	X
Domain 6: Editorial Independence	
22. The views of the funding body have not influenced the content of the guideline.	✓
23. Competing interests of guideline development group members have been recorded and addressed.	✓

✓ = yes; X = no or unclear.

Appendix 4: Main Study Findings and Author's Conclusions

Table A11: Summary of Findings of Included Systematic Reviews

Main Study Findings	Author's Conclusion
Reid, et al. 2015[16]	
Interventions Targeting Individuals with FASD **Developmental Outcomes in Infants** • 1 study found that children with PAE scored in the average (typically developing) range for development following an intensive 3-y in-home intervention service; 1 study found no effect on developmental outcomes (children scored below normative benchmarks) **Self-Regulation and Attentional Control** • 3 studies reported gains in executive functioning, and 1 study found "some evidence of" changes in gray matter volume in brain regions related to self-regulation, following 12 wk of the ALERT program (1-1.5 h/wk) • 1 study found significant improvements on measures of sustained and selective attention, math, and reading fluency, and a significant decrease in distractibility and reaction time, after 16 h (over 9 wk) of the Computerised Progressive Attention Program • 1 study found significant improvements in auditory and visual sustained attention and on nonverbal reasoning tasks following 6 h (12 x 30 min) of the "pay attention training protocol" and "additional visual search tasks"[16]; there was a trend for improved alternating attention (non-significant) • 1 study found significant improvement in behaviour ratings, but no change in cognitive functioning, after ~40 h (1 h/wk x 10 mo) cognitive control therapy **Specific Skills** • 3 studies implemented the MILE program (for 4 h, 15 wk, or duration not specified); 3 found significant gains in math knowledge, skills, or reasoning; 1 found improvements in problem behaviour; and 1 case study (5 participants) showed improved nonverbal reasoning and reading comprehension in 3 out of 5 children • 2 studies found that children were able to learn fire or street safety skills from computer games (duration not specified) • 1 study found significant improvement in language and literacy skills, but not general scholastic skills, after 38 h (1 h/wk x 9 mo) of a classroom-based literacy training intervention • 1 study found significant improvements in digit span (a memory-related measure) after 3 sessions (over 10 days) of group rehearsal training • 1 case study (1 child) found improvements in the number of words spelled correctly after 10 h (over 6 wk) of practicing a spelling procedure • 1 study found no impact of 24 h of a motor skills program (2 x 1.5 h/wk for 8 wk) on cortisol levels, which were higher in	*"[T]he body of literature reviewed showed that it is possible to make improvements across many domains of functioning." (*p 2292)[16] "*There is growing evidence for interventions that improve outcomes for early to middle childhood. However, a lack of research exists outside of this developmental period. This lack of research is concerning given the potential positive impact of early intervention, for individuals and, financially, for governments. In addition, the lack of interventions for adolescents and adults further highlights the widening developmental gap and the potential influence of secondary disabilities for this at-risk population.*"(p2283)[16] "*In summary, the parent-based intervention studies provide promising evidence that parents and caregivers benefit from support in managing their children's behavior and that this improvement is accompanied by improvements in parent/caregiver well-being." (*p2291)[16]

Main Study Findings	Author's Conclusion
the FASD vs. control group in the afternoon and evening **Social Skills** • 1 case study found improvements in social communication at 10 h over 6 wk of a social communication intervention • 3 studies implemented an adaptation of the CFT program (12 x 90 min); 3 found improvements in social skills; 1 found decreased problem behaviours; 1 found lower hostile attribution; 1 found improvement in self-esteem • 1 study reported that 7 of 11 parents showed improvements in social skills following a community-based group (90 min/wk x 7 wk) • 1 case study (n=4) found children were not able to imitate a block-building task after observing a 4-min video of the task <u>Interventions to Support Parents, Caregivers, and Others</u> **Parenting Skills** • 1 study found improved parental self-efficacy, parent needs and parent self-care, and reduced child behaviour problems, in families receiving the FMF intervention (90 min every 2 wk for 9-11 mo) compared to standard care • 1 study found reductions in child behavioural problems and parenting stress with a parent-only parenting support program, and greater reductions with a group adaptation of PCIT (both 1 h/wk x 14 wk) • 1 study delivered a parent-education program (2 x 2 h) in 3 formats: standard care (information packet), group workshops, internet training; knowledge of behavioural learning principles increased in all groups; knowledge of FASD and parent advocacy significantly improved in workshop and internet groups; child behaviour improved in standard care and workshop but not internet groups **Support, Education, and Advocacy** • 1 study found significant decreases in needs (e.g., housing and transport), caregiver stress, and increases in goals (e.g., self-care and health) in families who received the CF program; improvements were proportional to duration in the CF program • 1 study found improved placement stability (significantly fewer placement changes) following implementation of the "promising practices" intervention for children with out-of-home-care • 1 study found increased adaptive skills, significantly decreased school problems, and no change in academic achievement for children whose teachers participated in a professional development intervention (2 full-day and 4 half-day workshops, and weekly mentor-teacher meetings over 1 y) • 1 study provided qualitative support (no statistical significance) for a favorable impact of the Key Worker and Parent Support Program (length not stated) on parenting confidence, stress, parent/caregiver challenges, childcare, and child behaviour problems	

Main Study Findings	Author's Conclusion
Supporting Parents with FASD • 1 study found reductions in clients' needs (e.g., housing, mental health issues) and increases in goals (e.g., self-care and health) following a 3-y "step-by-step" program • 1 study found reduced alcohol and drug use, and increases in obtaining stable housing and use of contraception, medical, and mental health care series following a 12-mo modified PCAP-home visitation case management program	
Peadon, et al. 2009[2]	
Pharmacological Interventions • 1 study found that methylphenidate (0.6 mg/kg per dose to nearest 2.5 mg) significantly improved hyperactivity and impulsivity, but not attention, compared to placebo (lactose or Vitamin C); adverse effects were common with methylphenidate (3/4 children had reduced appetite; 2/4 children mild stomach aches; 2/4 children headaches) • 1 study found that stimulant medication (usual dose: methylphenidate, n = 8; pemoline, n = 2; dexedrine, n = 1) significantly improved hyperactivity scores, but not performance or attention, compared with placebo (colour-matched capsule); adverse effects were not reported Other Interventions All other included interventions were also included in the review by Reid et al. 2015[16]; findings are not repeated here (see Appendix 5 for overlap between systematic reviews).	"*[W]e found limited evidence for specific interventions for children with FASD. […] [S]timulant medication may decrease hyperactivity and impulsivity but does not improve attention[.]*" (p6)[2]
Elliott, et al. 2008[17]	
Diagnosis No eligible studies were identified. Economic No eligible studies were identified. Management **Systematic review findings:** • 1 systematic review was included; all 3 interventions were also included in the review by Reid, et al. 2014[16]; findings are not repeated here (see Appendix 5 for overlap between systematic reviews). **Narrative review findings:** • 1 narrative review identified publications that recommended the following interventions: nursing interventions; case management (e.g., financial assistance, medical, speech, behavioural care, social needs, educational services); health teaching, counselling and consultation; advocacy. No conclusions were made. • 1 narrative review included school-based interventions; the author suggested that executive functioning limitations must first be addressed, after which interventions focusing on teaching new skills using behavioural principles such as	Diagnosis *"This top level review of the postnatal screening and diagnostic literature did not identify any publications that evaluated the accuracy of the diagnostic criteria. Therefore, there is no evidence that any one criterion is the most appropriate." (*p155)[17] Economic "*There are a small number of published studies which examine the economics of FASD. The majority of these provide estimates of the economic burden of FASD (or specifically FAS) in Canada or the US." (*p185) [17]. No studies evaluated cost-effectiveness of diagnosis, assessment, or treatment of FASD. Management "*It is critical that management strategies are specifically tailored for each individual patient. Therefore, it is clinically inappropriate to recommend a single management strategy for all individuals with FASD. Generally, individuals with FASD benefit from a broad management plan, which requires the support of clinical staff, caregivers and teachers. Individuals need access to multiple services (e.g., physical, occupational, speech, behavioural, mental health) and caregivers may require assistance to ensure that children are able to access all of these*

Main Study Findings	Author's Conclusion
positive reinforcement and natural consequences can be undertaken. "The use of cognitive-behavioural strategies, such as social skill straining, emotion identification, coping skills, anger management, and self-talk" (p171)[17] were suggested to be helpful. • 1 narrative review recommended: neuropsychological testing, academic achievement, behavioural profile, adaptive skills, learning environment, and functional behaviour should be assessed so that interventions can be tailored to meet the needs of the individual with FASD. Six specific interventions were discussed: structure and systematic teaching, visual structure, environmental structure, task structure, cognitive control therapy, family role.	*services. Older children need practical interventions, such [as] improving skills of daily living, specific job skills and money management." (*p174)[17] "*There is no evidence to recommend any specific intervention for individuals with FASD." (*p175)[17]

CF = Coaching Families; CFT = Child Friendship Training; FASD = fetal alcohol spectrum disorder; FMF = Families Moving Forward; MILE = The Math Interactive Learning Experience; PAE = prenatal alcohol exposure; PCAP = Parent-Child Assistance Program; PCIT = Parent-Child Interaction Therapy; RCT = randomized controlled trial.

Table A12: Summary of Findings of Included Diagnostic Accuracy Studies

Main Study Findings	Author's Conclusion
Goh, et al. 2016[18]	
Psychologist Route *Child* Overall accuracy: 84.5% Sensitivity: 70.7% Specificity: 93.5% PPV: 87.9% NPV: 82.9% *Adolescent:* Overall accuracy: 84.7% Sensitivity: 79.3% Specificity: 87.6% PPV: 77.4% NPV: 88.7% **Pediatrician Route** *Child* Overall accuracy: 82.1% Sensitivity: 63.8% Specificity: 93.4% PPV: 85.7% NPV: 80.7% *Adolescent:* Overall accuracy: 79.5% Sensitivity: 81.3% Specificity: 78.3% PPV: 71.4% NPV: 86.2%	"*The decision tree developed in this study differentiates children affected by prenatal alcohol exposure from nonexposed children by the use of physical and neurobehavioral measures, even in the absence of a diagnosis of FAS." (*p125)[18] "*[T]he combination of both neurobehavioral and dysmorphology measures resulted in the greatest accuracy rates; reliance solely on dysmorphology measures resulted in reduced overall accuracy, because many children affected by prenatal alcohol exposure present without dysmorphology." (*p125)[18] "*The ability to accurately distinguish individuals affected by prenatal alcohol exposure is of great clinical importance, given the high rates of missed diagnosis and misdiagnosis." (*p127)[18] "*The decision tree model distinguished children affected by prenatal alcohol exposure from nonexposed control subjects, including those with other behavioral concerns or conditions." (*p 121)[18]

Main Study Findings	Author's Conclusion
Accuracy rates in post-hoc analyses with different clinical scenarios: • Limiting non-AE group to include only those with no known neurobehavioral problems: psychologist route, 89.2%; pediatrician route, 83.3% • Limiting non-AE group to include only those with known neurobehavioral problems: psychologist route, 67.2%; pediatrician route, 75.1% • Limiting AE group by excluding those with FAS: psychologist route, 84.7%; pediatrician route, 77.4% • Removing dysmorphology measures from the model: psychologist route, 80.2%; pediatrician route, 79.0% • Removing neurobehavioral measures from the model: pediatrician route: 68.6%	
Montag, et al. 2016[7]	

• Children classified as having FASD (using dysmorphological examination and BSID-II scores) had significantly different ultrasound measurements: longer IOD and lower FTD/IOD, OFD/IOD, and FTD/OD ratios (all $P < 0.05$)
• In regression analysis, the proportion of the variance in FASD accounted for by the ultrasound variables was:
  ○ Total sample: 5.8%
  ○ AE group only: 9.6%
	"*While some [ultrasound] measurements are associated with FASD outcome and specific dysmorphology measures, it is difficult to predict which children will be classified with an FASD."* (p2424)[7]

Kalberg, et al. 2013[19]	

• The final regression model contained these variables:
  ○ Digit Span A+B (A: assesses short-term sequential auditory memory, attention; B: assesses conceptual set shifting and working memory, logical memory, later recall)
  ○ Absurd situation (assesses recognition of right and wrong, impulse control)
  ○ Word associations (assesses ability to organize, abstract, and find relationships that are not immediately obvious)
• The overall ability of the model to correctly classify participants as FASD or control was 77.6%
 	
Note: Numerical values reported in the abstract do not match those reported in the results section. The more conservative estimates, from the results section, are reported here.
	"*A brief, practical set of tests can discriminate children with and without FASD and provide useful information for interventions for affected children."* (p2)[19]

Suttie, et al. 2013[3]	

• Face classification achieved significant agreement with clinical categories for non-AE vs FAS alone:
  ○ Agreement for "complete face" ranged from 0.97 to 1.00 depending on the type of analysis;
  ○ Agreement for "face profile" ranged from 0.92 to 0.93
• Agreement between face classification and clinical categories was lower when pFAS was added (i.e., non-AE vs FAS+PFAS):
	"*Heat maps and morphing visualizations of face signatures may help clinicians detect facial dysmorphism across the fetal alcohol spectrum. Face signature graphs show potential for identifying nonsyndromal heavily exposed children who lack the classic facial phenotype but have cognitive impairment."* (pe779)[3]
	
	"*Computer-based facial analysis shows potential for recognizing FAS facial characteristics, but without an accurate test for FASD, studies inducing classification schemes only assess*

Main Study Findings	Author's Conclusion
○ Agreement for "complete face" ranged from 0.89 to 0.91; ○ Agreement for "face profile" ranged from 0.92 to 0.93	*agreement with clinical categorization, which is not standardized." (*pe785)[3]
Burd, et al. 2010[4]	
• The FASDC differentiated between the FAS and non-FASD groups with 98.8% accuracy, 98.7% sensitivity, and 98.8% specificity • Logistic models for participants with scores in the FASD range (i.e., Other-FASD) were differentiated from FAS using data on phenotype and alcohol exposure with 81.8% accuracy, 84.8% sensitivity, and 79.7% specificity • Logistic models for participants with scores in the No-FASD range were differentiated *without* using data on phenotype or alcohol exposure, but using other descriptive data from the FASDC (maternal characteristics, birth records, demographic data) with 78.4% accuracy, 64.5% sensitivity, and 80.7% specificity ○ When the FASDC alcohol exposure and phenotype scores were added into the model, the accuracy and specificity decreased, and sensitivity increased (to 74.6%, 73.5% and 80.7% respectively)	"*The FASDC scores produce diagnostic groupings that approximate expert clinical judgment. The tool may be useful in other clinical settings for the diagnosis of FASD or as an FASD registry or research database." (*p605)[4] "*For the other-FASD group, the additional variables may be a useful strategy to improve the categorical diagnosis of subjects with decreased severity of their physical phenotype." (*p611)[4]
Mutsvangwa, et al. 2010[20]	
• Results of the discriminant function analysis cross-validation classification: ○ Age 5: accuracy, 95.46%; sensitivity, 100%; specificity, 90.91% ○ Age 12: accuracy, 80.13; sensitivity, 76.92%; specificity, 83.3%	*"[T]he differences in facial shape between FAS individuals in different age groups were more pronounced than for control individuals, supporting the notion that FAS facial anomalies diminish with age." (*p32)[20] "*Geometric morphometric analysis of stereo-photogrammetrically derived 3D facial landmarks allows visualization of the facial anomalies associated with FAS, as well as classification of facial shapes." (*p32)[20] "*This type of 3D shape analysis is a new tool in assessing the FAS facial phenotype and we have presented results that hold promise for objective diagnosis of FAS." (*p40)[20]
Fang, et al. 2008[9]	
FC sample – Automated classification • Overall accuracy, 92.6% • Sensitivity, 88.2% • Specificity, 100% CC sample – Automated classification • Overall accuracy, 90.9% • Sensitivity, 91.7% • Specificity, 90.0% Combined sample – Automated classification • Overall accuracy, 80.0% • Sensitivity, 82.8%	"*A new 3D facial image analysis technique was developed to identify children with FAS. Laser scans of facial images were collected, processed, and analyzed using computer graphics, machine learning, and pattern recognition techniques to determine the facial features that best discriminate FAS and control subjects. It provides an automated and potentially more accurate and efficient means to identify individuals with FAS."* (p13)[9] "*[F]acial features automatically selected by the algorithm to distinguish FAS and control faces vary among different ethnic populations. Prenatal alcohol exposure not only leads to the specific dysmorphic features outlined in the criteria for FAS –*

Main Study Findings	Author's Conclusion
• Specificity, 76.2%	*short palpebral fissure, thin upper lip and vermillion border – but also to other more subtle yet, by 3D imaging, identifiable features that make the overall gestalt of a FASD face." (*p9)[9] "*[T]he classification rate was lower in a sample of mixed races, indicating that ethnicity plays a significant role in the features that help to identify individuals who are prenatally exposed to alcohol." (*p9)[9] *"These results suggest 3D images are a useful addition to the current dysmorphology and clinical evaluation of individuals suspected of prenatal alcohol exposure. Importantly, this approach shows great promise as part of growing efforts to develop novel telemedicine applications which would allow better clinical care in remote locations."* (p9-10)[9]
Thorne and Coggins 2008[21]	
FASD vs. TD classification: • rNRE: AUC, 0.90; 95% CI, 0.73 to 0.97; % accuracy, 88% (false positives, 3; false negatives, 1) • ANR/TW: AUC, 0.86; 95% CI, 0.70 to 0.96; % accuracy, 81% (false positives, 4; false negatives, 2) **FAS/pFAS vs. TD and other FASD classification:** • rNRE: AUC, 0.98; 95% CI, 0.85 to 0.99; % accuracy, 97% (false positives, 1; false negatives, 0) • ANR/TW: AUC, 1.00; 95% CI, 0.89 to 1.00; % accuracy, 100% (false positives, 0; false negatives, 0)	"*The proposed system for calculating the rNRE was highly accurate at predicting which narratives were produced by children with FASD (versus TD, 88% overall accuracy), and which were produced by children with FAS/pFAS (versus all others, 97% overall accuracy), and outperformed all other methods tested." (*p570-571)[21] "*The strong predictive accuracy demonstrated in this study provides empirical evidence that the system proposed in this feasibility study has sufficient sensitivity and diagnostic utility to warrant further development for use with children suspected of prenatal alcohol exposure."* (p571)[21] "*The evidence presented in this study reveals that the proposed method for calculating nominal reference errors is reliable and has sufficient diagnostic utility to merit further study with children suspected of prenatal alcohol exposure."* (p589)[21]
Moore, et al. 2007[22]	
Discriminant analysis was performed separately in each of the 4 ethnic groups, and different facial measurements were included in each. **AA:** • Included features: palpebral fissure and philtrum length • Overall accuracy, 79% • Sensitivity, 73% • Specificity, 85% **CC:** • Included features: minimal frontal width, bizygomatic width, inner canthal width, philtrum length, ear length • Overall accuracy, 92% • Sensitivity, 94% • Specificity, 91% **FC:**	"*We found measurements that reflected reduced size of the eye orbit to be a consistent feature discriminating FAS and controls across each study population. However, each population had a unique, although often overlapping, set of variables which discriminated the 2 groups [those with FAS and controls], suggesting important ethnic differences in the presentation of FAS."* (p1707)[22] "*We have demonstrated that a unique set of facial measurements can be obtained from remotely gathered e-dimensional images and used to separate FAS from control individuals in each population we studied."* (p1712)[22]

Main Study Findings	Author's Conclusion
• Included features: age, bitragal width, inner canthal width, outer canthal width, palpebral fissure length, midfacial depth, nasal length, nasal bridge length, ear length • Overall accuracy, 93% • Sensitivity, 96% • Specificity, 91% **NAC:** • Included features: inner canthal width, outer canthal width • Overall accuracy, 77% • Sensitivity, 74% • Specificity, 81%	
Thorne, et al. 2007[23]	
FASD vs. TD Classification • ANR/TW: AUC, 0.86; 95% CI, 0.74 to 0.99; sensitivity, 87.5%; specificity, 75.0%; efficiency, 81.3% • At the cut-point identified in the ROC analysis, 14/16 children from the FASD group and 12/16 from the TD group were correctly identified • The cut-point was >0.165, which is interpreted to mean more than 1.65% of total words being ANRs indicates a positive test for FASD status	"*[T]he rate of ambiguous nominal reference was highly accurate in classifying children with an FASD regardless of their performance on the standardized language task (area under the ROC curve = .863, confidence interval = .991).*" (p459)[23] "*Results support further study of the diagnostic utility of narrative analysis using discourse level measures of elaboration and children's strategic use of reference.*" (p459)[23]

3D = three dimensional; AA = African American; AE = alcohol exposed; ANR/TW = rate of ambiguous nominal reference errors; AUC = area under the (receiver-operating characteristic) curve; BSID-II = Bayley Scales of Infant Development, 2nd ed.; CC = Cape Coloured; FASD = fetal alcohol spectrum disorder; FASDC = Fetal Alcohol Syndrome Diagnostic Checklist; FC = Finnish Caucasian; FTD = frontothalamic distance; IOD = interorbital distance; NAC = North American Caucasian; NPV = negative predictive value; OD = orbital diameter; OFD = occipital frontal diameter; PPV = positive predictive value; rNRC = rate of nominal reference errors; ROC = receiver-operating characteristic; TD = typically developing.

Table A13: Summary of Findings of Included Clinical Effectiveness Studies

Main Study Findings	Author's Conclusion
McCoy, et al. 2015[24]	
Postural control: Both children with FASD and those with TD had significantly higher anterior-posterior and medial-lateral postural sway velocity after STABEL practice, indicating decreased postural stability. This may have been related to fatigue from the long testing protocol. **Sensory attention:** Sensory attention (as assessed by "sensory attention fraction" and "entrainment gain") was not different from pre- to post-STABEL training in either group of children. Children with FASD had higher entrainment gain than children with TD both before and after STABEL training, suggesting poorer balance.	"*We demonstrated that the STABEL system was feasible for school-aged children with and without postural and balance control deficits. Our preliminary results indicated the STABEL system provoked sensory (vestibular) responses during balance practice, but group immediate effects on sensory attention were limited. Analysis of individual responses and patterns of change, however, suggest grounds for further study of the STABEL system using a larger sample and dose.*" (p1580)[24]

FASD = fetal alcohol spectrum disorder; MuMBER = Multimodal Balance Entrainment Response; STABEL = Sensorimotor Training to Affect Balance, Engagement and Learning; TD = typical development.

Table A14: Summary of Findings of Included Economic Evaluations

Main Study Findings	Author's Conclusion
Thanh, et al. 2013[25]	
• 471 adults and 804 children were served within the Alberta FASD SN • If no SN was in place, the secondary disabilities would cost C$22.85 million per year (C$8.62 million per year for adults; C$14.24 million per year for children) • The cost of the SN was C$6.12 million per year • The break-even effectiveness (i.e., effectiveness level at which the SNs became cost-saving) was estimated at 28% (range identified via sensitivity analysis: 25% to 32%)	*"[T]he economic and social burden associated with secondary disabilities is significant and there are economic opportunities to reduce the resource burden on social resources and programs, while improving services for persons with FASD." (*pe199)[25] *"[T]he analysis can be regarded as a conservative estimate, and shows that a moderate reduction in utilization associated with secondary disabilities will allow a break-even result." (*pe199)[25] "*[T]he economic and social burden associated with secondary disabilities will "pay-off" if the effectiveness of the program in reducing secondary disabilities is 28%." (*pe193)[25]

FASD = Fetal Alcohol Spectrum Disorder; SN = Service Network.

Table A15: Summary of Recommendations in Included Guideline

Findings and Recommendations	Quality of Evidence, Strength of Recommendation
Cook, et al. 2016[26]	
Recommendations for screening, referral and support	
"All pregnant and postpartum women should be screened for alcohol use with validated measurement tools by service providers who have received appropriate training in their use. Women at risk of heavy alcohol use should receive early, brief interventions (i.e., counselling and/or other services)." (p193)[26]	Quality of Evidence: High Strength of Recommendation: Strong
"Referral of individuals for a possible FASD diagnosis should be made whenever there is evidence of, or suspected prenatal alcohol exposure at levels associated with, physical or developmental effects." (p193)[26]	Quality of Evidence: Moderate Strength of Recommendation: Strong
"Abstinence from alcohol should be recommended to all women during pregnancy to ensure the safest outcome for the fetus, and appropriate support should be provided, as indicated." (p193)[26]	Quality of Evidence: High Strength of Recommendation: Strong
Recommendations for medical assessment	
"The diagnostic process should include compiling a social and medical history and complete physical examination." (p193)[26]	Quality of Evidence: High Strength of Recommendation: Strong
"Confirmation of prenatal alcohol exposure requires documentation that the biological mother consumed alcohol during the index pregnancy based on: reliable clinical observation; self-report; reports by a reliable source; medical records documenting positive blood alcohol concentrations; alcohol treatment; or other social, legal or medical problems related to drinking during the pregnancy. The presence of all three facial features has such high specificity to alcohol exposure and FASD that confirmation of alcohol exposure is not	Quality of Evidence: Moderate Strength of Recommendation: Strong

Findings and Recommendations	Quality of Evidence, Strength of Recommendation
required. The presence of fewer than three facial features does not have the same degree of specificity and therefore requires other confirmation." (p193)[26]	
Recommendations for assessment of facial features	
"The following three sentinel facial features must be present because of their specificity to prenatal alcohol exposure: • Palpebral fissure length ≥ 2 SDs below the mean (< third percentile). • Philtrum rated 4 or 5 on 5-point scale of the University of Washington Lip–Philtrum Guide. • Upper lip rated 4 or 5 on 5-point scale of the University of Washington Lip–Philtrum Guide." (p193)[26]	Quality of Evidence: High Strength of Recommendation: Strong
Recommendations for neurodevelopmental assessment	
"A diagnosis of FASD is made only when there is evidence of pervasive brain dysfunction, which is defined by severe impairment in three of more of the following neurodevelopmental domains: motor skills; neuroanatomy/neurophysiology; cognition; language; academic achievement; memory; attention; executive function, including impulse control and hyperactivity; affect regulation; and adaptive behaviour, social skills or social communication."(p193)[26]	Quality of Evidence: High Strength of Recommendation: Strong
"Severe impairment is defined as a global score or a major subdomain score on a standardized neurodevelopmental measure that is ≥ 2 SDs below the mean, with appropriate allowance for test error. In some domains, large discrepancies among subdomain scores may be considered when a difference of this size occurs with a very low base rate in the population (≤ 3% of the population). Clinical assessment with converging evidence from multiple sources and DSM-V diagnostic criteria for certain disorders may also be considered in specific domains that are not easily assessed by standardized tests. For example, in the affect regulation domain, the following diagnoses may be taken as an indication of severe impairment: major depressive disorder (with recurrent episodes), persistent depressive disorder, disruptive mood dysregulation disorder, separation anxiety disorder, selective mutism, social anxiety disorder, panic disorder, agoraphobia or generalized anxiety disorder." (p193)[26] A domain-by-domain discussion of how these criteria are operationalized is outlined in Appendix 1 of their report.[26]	Quality of Evidence: Moderate Strength of Recommendation: Strong
Recommendations for nomenclature and diagnostic criteria	
"A diagnosis of FASD may be made if an individual meets either of the two sets of criteria below: **FASD with sentinel facial features** • Simultaneous presentation of the three sentinel facial features (see [above]); AND • Prenatal alcohol exposure confirmed or unknown; AND • Evidence of impairment in three or more of the identified neurodevelopmental domains (see [above]) or, in infants and young children, evidence of microcephaly. OR **FASD without sentinel facial features**	Quality of Evidence: High Strength of Recommendation: Strong

Findings and Recommendations	Quality of Evidence, Strength of Recommendation
• Evidence of impairment in three or more of the identified neurodevelopmental domains (see [above]); AND • Confirmation of prenatal alcohol exposure, with the estimated dose at a level known to be associated with neurodevelopmental effects." (p193)[26] "**At risk for neurodevelopmental disorder and FASD, associated with prenatal alcohol exposure** This is not a diagnosis; this is a designation that should be given to individuals when: • There is confirmation of prenatal alcohol exposure, with the estimated dose at a level known to be associated with neurodevelopmental effects; • Central nervous system criteria from [above] are not met; • There is some indication of neurodevelopmental disorder in combination with a plausible explanation as to why the neurodevelopmental assessment results failed to meet the criteria for substantial impairment (e.g., patient was too young; incomplete assessment). This designation may also be considered for individuals with all three sentinel facial features as described [above] who do not yet have documentation or evidence of the requisite three or more neurodevelopmental domain criteria or true microcephaly. This designation should never be considered when prenatal alcohol exposure is confirmed absent." (p194)[26]	
Recommendations for the diagnostic team	
"Core team members across the lifespan: *For infants (< 18 mo)* • Pediatrician/physician • Child development specialist who has the skill set to conduct physical and functional assessments (e.g., • speech-language pathologist, physiotherapist, occupational therapist, clinical psychologist) *For preschoolers (18 mo–5 yr)* • Pediatrician/physician • Occupational therapist • Speech-language pathologist • Psychologist *For school-aged children (6 yr–age of majority)* • Pediatrician/physician with expertise in FASD and differential diagnosis • Occupational therapist • Speech-language pathologist • Psychologist *For adults* • Physician • Psychologist • Speech-language pathologist/psychologist with expertise in language assessment" (p194)[26]	Quality of Evidence: High Strength of Recommendation: Strong
Recommendations for special considerations in the neurodevelopmental assessment of infants and young children	
"Infants and young children with all three sentinel facial features and microcephaly should be given a diagnosis of "FASD with	Quality of Evidence: High

Findings and Recommendations	Quality of Evidence, Strength of Recommendation
sentinel facial features"; these children have a high risk of neurodevelopmental disorder. They should also be referred to a clinical geneticist." (p194)[26]	Strength of Recommendation: Strong
"Infants and young children with all three facial features may be given a diagnosis of "FASD with sentinel facial features" if they undergo a comprehensive neurodevelopmental assessment and show deficits in three or more brain domains. Infants and young children with confirmed prenatal alcohol exposure may be given a diagnosis of "FASD without sentinel facial features" if they undergo a comprehensive neurodevelopmental assessment and show deficits in three or more brain domains." (p194)[26]	Quality of Evidence: Moderate Strength of Recommendation: Strong
"Infants and young children with confirmed prenatal alcohol exposure but who do not meet the criteria for FASD should be designated as "At risk for neurodevelopmental disorder and FASD, associated with prenatal alcohol exposure." Those with all three facial features but no microcephaly should be referred to clinical genetics." (p194)[26]	Quality of Evidence: High Strength of Recommendation: Strong
"A complete neurodevelopmental assessment should be recommended at an age-appropriate time for all infants and young children with confirmed prenatal alcohol exposure and/or all three facial features." (p194)[26]	Quality of Evidence: High Strength of Recommendation: Strong
Recommendations for special considerations in the neurodevelopmental assessment of adolescents and adults	
"Recommendations following the assessment must address basic and immediate needs of the client, and assist them in accessing required resources." (p194)[26]	Quality of Evidence: Moderate Strength of Recommendation: Strong
Recommendations for management and follow-up	
"Individuals with FASD and their caregivers should be linked to resources that can improve outcomes. However, just because availability of services is limited, an individual should not be denied an assessment and management plan. Often the diagnosis is the impetus that leads to the development of resources." (p194)[26]	Quality of Evidence: Low Strength of Recommendation: Strong
"When young adults are transitioning to independent living situations, it may require that they undergo a reassessment to identify changes in their adaptive function and to make subsequent adjustments to their management plan." (p194)[26]	Quality of Evidence: Low Strength of Recommendation: Strong

DSM-V = Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition; FASD = fetal alcohol spectrum disorder; SD = standard deviation.

Note: A diagnostic algorithm (flow-chart) was included in the guideline to facilitate diagnosis.[26]

Appendix 5: Overlap between Included Systematic Reviews

Primary Study Author, Publication Year	Systematic Review Author, Publication Year		
	Reid et al. 2015[16]	Peadon et al. 2009[2]	Elliott et al. 2008[17]
Adnams et al. 2007	•	•	
Adnams et al. reported in Riley et al. 2003	•	•	•
Clark et al. 2014	•		
Coles et al. 2007	•	•	
Denys et al. 2011	•		
Grant et al. 2004	•		
Green 2007			•
Gryiec et al. 2004	•		
Gurwitch et al. reported in Bertrand 2009	•		
Hume et al. 2009	•		
Kable et al. 2007	•	•	
Kable et al. 2012	•		
Kable et al. 2015	•		
Kalberg and Buckley 2007			•
Kartin et al. 2002	•		
Keil et al. 2010	•		
Keiver et al. 2015	•		
Kerns et al. 2010	•		
Leenaars et al. 2012	•		
Loomes et al. 2008	•	•	
Meyer 1998	•	•	
Millians and Coles 2014	•		
Nash et al. 2015	•		
O'Connor et al. 2006	•	•	
O'Connor et al. 2012	•		
Oesterheld et al. 1998		•	•
Olson et al. reported in Bertrand 2009	•		
Padgett et al. 2006	•	•	
Pelech et al. 2013	•		
Premji et al. 2007			•
Snyder et al. 1997		•	•
Soh et al. 2015	•		
Sparks-Keeny et al. 2011	•		
Timler et al. 2005	•	•	
Vernescu 2008	•	•	
Wells et al. 2015	•		
Yazdani et al. 2009	•		

Appendix 6: Additional References of Potential Interest

Framework for Developing Interventions

Kodituwakku PW, Kodituwakku EL. From research to practice: an integrative framework for the development of interventions for children with fetal alcohol spectrum disorders. Neuropsychol Rev. 2011 Jun;21(2):204-23.

Guidelines

Guideline summary: guidelines for the identification and management of substance use and substance use disorders in pregnancy. In: National Guideline Clearinhouse [Internet]. Rockville (MD): Agency for Healthcare Research and Quality; 2014 [cited 2017 Apr 11]. Guideline summary available from: https://www.guideline.gov/summaries/summary/48894/guidelines-for-the-identification-and-management-of-substance-use-and-substance-use-disorders-in-pregnancy?q=fetal+alcohol+syndrome *Population of interest is pregnant women.*

Hoyme HE, Kalberg WO, Elliott AJ, Blankenship J, Buckley D, Marais AS, et al. Updated clinical guidelines for diagnosing fetal alcohol spectrum disorders. Pediatrics. 2016 Aug;138(2). *Not based on a systematic review.*

Landgraf MN, Nothacker M, Kopp IB, et al. The diagnosis of fetal alcohol syndrome. Dtsch Arztebl Int 2013;110:703-710. *German guidelines.*

Tao L, Temple V, Casson I, Kirkpatrick SM. Health watch table - fetal alcohol spectrum disorder (FASD) [Internet]. Toronto (ON): Surrey Place Centre; 2013 [cited 2017 Apr 11]. Available from: http://www.surreyplace.on.ca/documents/Primary%20Care/HWT_FASD_Feb_2014.pdf *Unclear methodology.*

Watkins RE, Elliott EJ, Wilkins A, Mutch RC, Fitzpatrick JP, Payne JM, et al. Recommendations from a consensus development workshop on the diagnosis of fetal alcohol spectrum disorders in Australia. BMC Pediatr [Internet]. 2013 [cited 2017 Apr 11];13:156. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3849849. *Australian guidelines.*

Narrative Reviews and Reports – 2012 to 2017 and/or specific to Aboriginal Peoples

Bax AC, Geurts CD, Balachova TN. Improving recognition of children affected by prenatal alcohol exposure: detection of exposure in pediatric care. Curr Dev Disord Rep. 2015 Sep 1;2(3):165-74.

Coriale G, Fiorentino D, Di LF, Marchitelli R, Scalese B, Fiore M, et al. Fetal alcohol spectrum disorder (FASD): neurobehavioral profile, indications for diagnosis and treatment. Riv Psichiatr. 2013 Sep;48(5):359-69

Kable JA, Mukherjee RA. Neurodevelopmental disorder associated with prenatal exposure to alcohol (ND-PAE): a proposed diagnostic method of capturing the neurocognitive phenotype of FASD. Eur J Med Genet. 2017 Jan;60(1):49-54.

Leibson T, Neuman G, Chudley AE, Koren G. The differential diagnosis of fetal alcohol spectrum disorder. J Popul Ther Clin Pharmacol. 2014;21(1):e1-e30.

Memo L, Gnoato E, Caminiti S, Pichini S, Tarani L. Fetal alcohol spectrum disorders and fetal alcohol syndrome: the state of the art and new diagnostic tools. Early Hum Dev. 2013 Jun;89 Suppl 1:S40-S43.

Murawski NJ, Moore EM, Thomas JD, Riley EP. Advances in diagnosis and treatment of fetal alcohol spectrum disorders: from animal models to human studies. Alcohol Res. 2015;37(1):97-108

Pacey M. Fetal alcohol syndrome and fetal alcohol spectrum disorder among Aboriginal peoples: a review of prevalence. Prince George (BC): National Collaborating Centre for Aboriginal Health; Sep 2009 [cited 2017 Apr 11]. Available from: http://www.nccah-ccnsa.ca/Publications/Lists/Publications/Attachments/34/2010_01_27_FASDAboriginalReviewPrevalence_EN_Updated_Web.pdf

Pruett D, Waterman EH, Caughey AB. Fetal alcohol exposure: consequences, diagnosis, and treatment. Obstet Gynecol Surv. 2013 Jan;68(1):62-9.

Salmon J. Diagnosing fetal alcohol disorder: historical to present day challenges. JSM Schizophrenia.1(1): 1004; 2016.

Tait CL. Fetal alcohol syndrome among Aboriginal people in Canada: review and analysis of the intergenerational links to residential schools. Ottawa (ON): Aboriginal Healing Foundation; 2003 [cited 2017 Apr 11]. Available from: http://www.phac-aspc.gc.ca/about_apropos/evaluation/reports-rapports/2013-2014/efasdi-eietaf/index-eng.php

Rural Diagnostic Services Model

McFarlane A, Rajani H. Rural FASD diagnostic services model: Lakeland Centre for fetal alcohol spectrum disorder. Can J Clin Pharmacol. 2007;14(3):e301-e306.

Screening - Tool Kit and Cost-Effectiveness

Hopkins RB, Paradis J, Roshankar T, Bowen J, Tarride JE, Blackhouse G, et al. Universal or targeted screening for fetal alcohol exposure: a cost-effectiveness analysis. J Stud Alcohol Drugs. 2008 Jul;69(4):510-9.

National screening tool kit for children and youth identified and potentially affected by FASD [Internet]. Ottawa (ON): Canadian Association of Paediatric Health Centres; 2010 [cited 2017 Apr 11]. http://ken.caphc.org/xwiki/bin/download/FASDScreeningToolkit/National+Screening+Tool+Kit+for+Children+and+Youth+Identified+and+Potentially+Affected+by+FASD/FASDToolKit-BoundEN.pdf

Treatment Improvement Protocol

Substance Abuse and Mental Health Services Adminstration (SAMHSA) [Internet]. Addressing fetal alcohol spectrum disorders (FASD) [Treatment improvement protocol (TIP); no. 58]. Rockville (MD): U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration; 2014 [cited 2017 Apr 11]. Available from: http://www.healthsac.net/downloads/publications/HSAC07_FASD_FINALv3.pdf