

# Comparison of the 4-Digit Code, Canadian 2015, Australian 2016 and Hoyme 2016 fetal alcohol spectrum disorder diagnostic guidelines

Susan J Astley Hemingway<sup>\*1,2</sup>, Julia M Bledsoe<sup>2</sup>, Allison Brooks<sup>1</sup>, Julian K Davies<sup>2</sup>, Tracy Jirikowic<sup>3</sup>, Erin Olson<sup>4</sup>, John C Thorne<sup>5</sup>

<sup>1</sup>Department of Epidemiology, University of Washington, Seattle, USA; <sup>2</sup>Department of Pediatrics, University of Washington, Seattle, USA;

<sup>3</sup>Department of Rehabilitation Medicine, University of Washington, Seattle, USA, <sup>4</sup>Department of Education, University of Washington,

<sup>5</sup>Department of Speech and Hearing, University of Washington, Seattle, USA.

## ABSTRACT

**Background:** As clinicians strive to achieve consensus worldwide on how best to diagnose fetal alcohol spectrum disorders (FASD), the most recent FASD diagnostic systems show convergence and divergence. Applying these systems to a single clinical population illustrates the contrasts between them, but validation studies are ultimately required to identify the best system.

**Methods:** The 4-Digit-Code, Hoyme 2016, Canadian 2015 and Australian 2016 FASD diagnostic systems were applied to 1,392 patient records evaluated for FASD at the University of Washington. The diagnostic criteria and tools, the prevalence and concordance of diagnostic outcomes, and validity measures were compared between the systems.

**Results:** The proportion diagnosed with fetal alcohol syndrome (FAS) and FASD varied significantly (4-Digit-Code 2.1%, ≤79%; Hoyme 6.4%, 44%, Australian 1.8%, 29%; Canadian 1.8%, 16%). Eighty-two percent were diagnosed FASD by at least one system; only 11% by all four systems. Key factors contributing to discordance include: requiring high alcohol exposure; excluding growth deficiency; relaxing the facial criteria; requiring brain criteria that prevent diagnosis of infants/toddlers; and excluding moderate dysfunction from the spectrum. Primate research confirms moderate dysfunction (1-2 domains ≤2 standard deviations) is the most prevalent outcome caused by PAE (FAS 5%, severe dysfunction 31%, moderate dysfunction 59%). Only the 4-Digit-Code replicated this diagnostic pattern.

**Conclusion:** The needs of individuals with FASD are best met when diagnostic systems provide accurate, validated diagnoses across the lifespan, the full spectrum of outcome, the full continuum of alcohol exposure; and utilize diagnostic nomenclature that accurately reflects the association between outcome and alcohol exposure.

## INTRODUCTION

Alcohol is a well-recognized teratogen and both human and animal research indicates that the impact of prenatal alcohol exposure (PAE) manifests as a spectrum of developmental variations in severity and type of dysfunction across individuals [1-3]. These outcomes vary significantly based on timing and dosage of exposure as well as the presence of other risk factors and are typically characterized as including physical impacts (i.e., growth deficiency, facial dysmorphism and structural brain abnormalities) as well as functional impairment of the central nervous system (CNS). This

spectrum of outcome was found in early primate studies on the impact of prenatal alcohol exposure. For example Clarren et al. [4] document the distribution of developmental outcomes when the only risk factor present was PAE. In that study, the primates had been exposed weekly to binge exposures equivalent to a six-pack of beer for the first 3, 6 or entire 24 weeks of gestation (mean maternal peak plasma ethanol concentrations ranged from 176 to 271 mg/dl). The primate model confirmed that PAE causes a spectrum of outcomes; the most common outcome of PAE was moderate CNS dysfunction (the 4-Digit Code equivalent of Neurobehavioral Disorder/Alcohol Exposed ND/AE) (59% of primates) followed

**Correspondence to:** Susan J. Astley Hemingway, Departments of Epidemiology and Pediatrics, University of Washington, Seattle, WA 98195 USA; E-mail: astley@uw.edu

**Received:** September 25, 2019, **Accepted:** October 10, 2019, **Published:** October 17, 2019

**Citation:** Hemingway SJA, Bledsoe JM, Brooks A, Davies JK, Jirikowic T, Olson E, et al. (2019) Comparison of the 4-Digit Code, Canadian 2015, Australian 2016 and Hoyme 2016 Fetal Alcohol Spectrum Disorder Diagnostic Guidelines. *Adv Pediatr Res* 6:31. doi: 10.35248/2385-4529.19.6.31

**Copyright:** © 2019 Hemingway SJA, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

by more severe CNS impairment (the 4-Digit Code equivalent of Static Encephalopathy/Alcohol Exposed SE/AE) (31%); notably sentinel physical impacts (the 4-Digit Code equivalent of FAS/PFAS) were found in only 5% of primates under these controlled conditions, and a similar number of primates exhibited little to no impacts of PAE (5%).

These facts present a challenge to public health systems seeking ways to best capture this spectrum of outcomes in order to appropriately diagnose and serve individuals that may have been impacted by PAE. As the field of fetal alcohol spectrum disorders (FASD) strives to achieve consensus worldwide on how best to meet this diagnostic challenge, the most recent versions of published guidelines (4-Digit Code, 2004 [5]) Canadian, 2015 [6], Hoyme, 2016 [7], and Australian, 2016 [8] show both convergence and divergence in their approach. For example, the new Canadian and Australian diagnostic systems have many features in common with one another and have adopted the facial criteria of the 4-Digit Code, but diverge substantially from the 4-Digit Code and Hoyme systems by removing growth deficiency from their diagnostic criteria [9] and adopting a nomenclature (FASD with the face, and FASD without the face) that reflects a dichotomy rather than a spectrum of outcomes. The 4-Digit Code [5] and Hoyme [7] criteria continue to generate a spectrum of diagnoses under the umbrella of FASD (fetal alcohol syndrome (FAS), partial FAS (PFAS), Alcohol Related Neurodevelopmental Disorder (ARND), Static Encephalopathy/Alcohol Exposed (SE/AE), Neurobehavioral Disorder/Alcohol Exposed (ND/AE), and Alcohol Related Birth Defects (ARBD)) and maintain the 3 original core diagnostic criteria (growth deficiency, facial anomalies, and CNS abnormalities). The 4-Digit Code and Hoyme systems differ significantly in their diagnostic nomenclature, diagnostic tools, and the specific criteria used to generate each diagnosis. The Canadian [6] and Hoyme systems require high PAE; the 4-Digit Code and Australian systems require confirmed PAE at any reported level. The Canadian and Australian [8] systems do not include moderate dysfunction under the umbrella of FASD; the 4-Digit Code and Hoyme systems do. The contrasts in these systems create confusion for clinicians faced with diagnosing FASD. Applying these systems to a single clinical population illustrates the contrasts between them, but validation studies are ultimately required to identify the best system.

### The objectives of this study were to:

1. Compare the tools, nomenclature and criteria used by the four diagnostic systems.
2. Administer each system to the records of 1,392 patients to:
  - a. Compare the prevalence of FASD diagnoses produced by each system.
  - b. Assess diagnostic discordance/concordance between the four systems.
  - c. Assess and compare the diagnostic performance (validity) of each system.

A comprehensive comparison of the 4-Digit Code and Hoyme 2016 systems was conducted in 2017 [10]. This study expands

the comparison to include all four diagnostic systems using the same clinical population of 1,392 patients. Key findings from the published comparison of the 4-Digit Code and Hoyme systems are included in this report, but the Reader is referred to the previous publication [10] for more detail. Since contrasts in the diagnostic tools and criteria used by each system impact our application of each system to our study population, the methods and results for Objective 1 are presented first, followed by the methods and results for Objective 2.

## OBJECTIVE 1. COMPARISON OF THE TOOLS, NOMENCLATURE AND CRITERIA USED BY THE FOUR SYSTEMS

### Methods

The following tools, nomenclature and criteria used by the four diagnostic systems were compared.

#### Lip-philtrum guides

The 4-Digit Code introduced two guides: Lip-Philtrum Guide 1 for Caucasians and all races with thinner upper lips like Caucasians, and Lip-Philtrum Guide 2 for African Americans and all races with thicker upper lips like African Americans (Figures 1A and 1B). These Lip-Philtrum Guides were adopted for use by the Canadian and Australian systems. Hoyme 2016 introduced two different lip/philtrum guides: the North American Lip/Philtrum Guide [7] produced from a U.S. white population and the South African Mixed Race Lip/Philtrum Guide [5] produced from a Cape Coloured (mixed race) population in the Western Cape Province (Figures 1C and 1D).

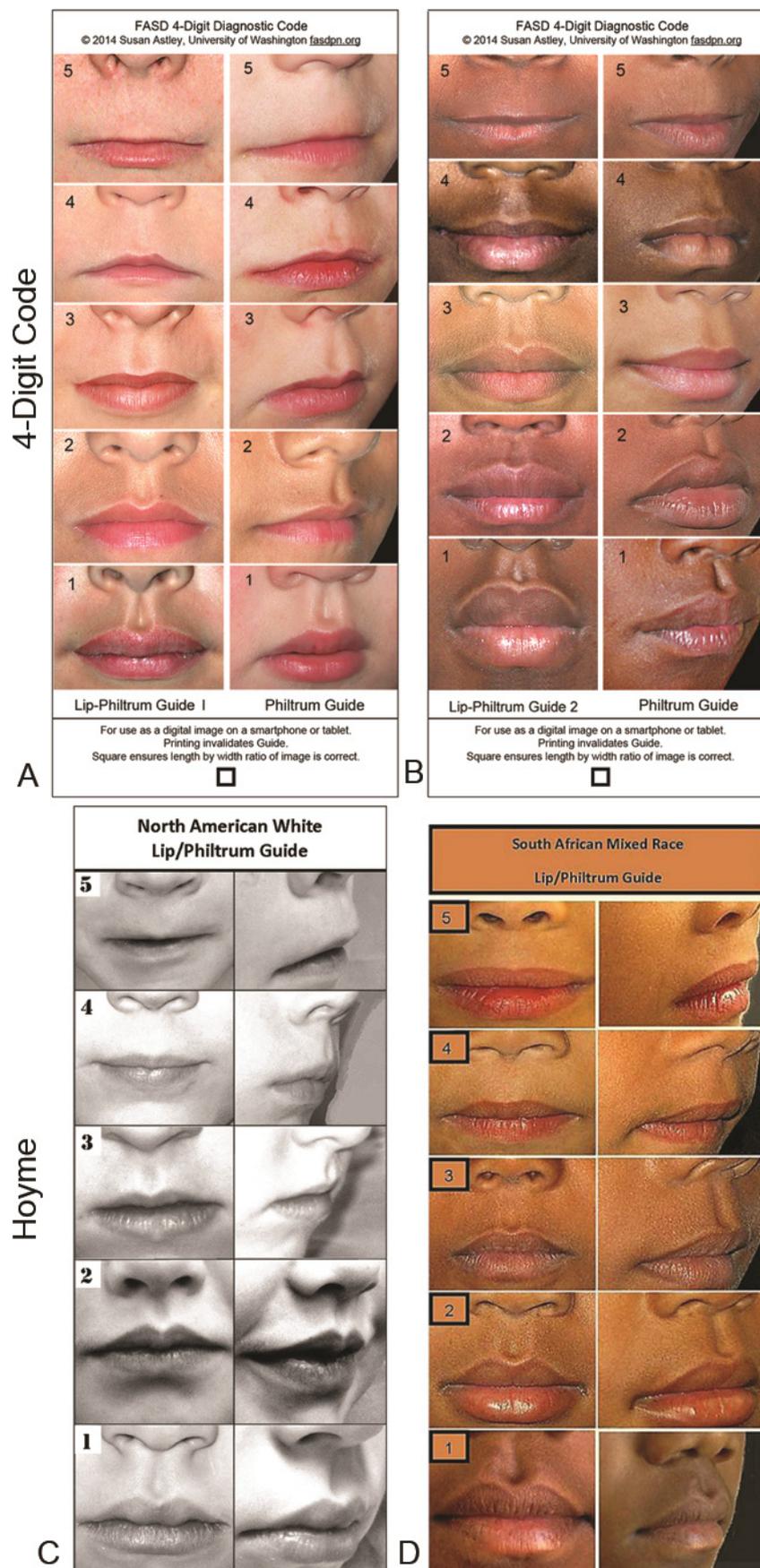
The Rank 1-5 lips depicted on the 4-Digit Code Caucasian and Hoyme et al. [7] North American guides were compared using the objective, quantitative measure of lip thinness called lip circularity ( $\text{perimeter}^2/\text{area}$ ) generated by the FAS Facial Photographic Analysis Software [12]. Circularity is computed by outlining the vermilion border of the upper lip with the computer mouse (Figure 2C); the thinner the lip, the bigger the circularity.

#### PFL normal growth charts

The 4-Digit Code uses the Stromland Scandinavian PFL normal growth charts for all races except African American [13]. The Stromland PFL norms cover the full lifespan (birth to adult). These same charts were used for the Hoyme system. The Canadian and Australian systems use the Stromland charts for patients <6 years of age and the Clarren Canadian PFL [14] charts for patients 6 years of age and older.

#### Facial analysis software

The 4-Digit Code advises measuring the facial features from 2D digital photos using the FAS Facial Photographic Analysis Software [12]. The Canadian and Australian systems also encourage the use of the FAS Facial Photographic Analysis Software. The authors of the Hoyme system “feel direct examinations of facial features are more practical in an office setting”. Since empirical studies have already confirmed the superior accuracy of the photo versus direct method of facial measurement [13,15], a formal assessment of photo versus direct measurement of facial features was not repeated in this study.



**Figure 1.** Lip/Philtrum Guides.

The 4-Digit Code [5] introduced two guides in 1999: A) Lip-Philtrum Guide 1 for Caucasians and all races with thinner upper lips like Caucasians, and B) Lip-Philtrum Guide 2 for African Americans and all races with thicker upper lips like African Americans. Hoyme introduced two different lip/philtrum guides: C) the North American Lip/Philtrum Guide in 2016 [7] produced from a U.S. white population (reproduced with permission from Pediatrics [7] copyright 2019 by the AAP) and D) the South African Mixed Race Lip/Philtrum Guide in 2015 [11] produced from a Cape Coloured (mixed race) population in the Western Cape Province (reproduced with permission from AJMG [11] copyright 2019 by John Wiley & Sons).

## Diagnostic nomenclature and criteria

Tables were created to illustrate the key contrasts between the diagnoses generated by each system, the nomenclature assigned to each diagnosis, and the diagnostic criteria.

## Results

### Contrasts in lip-philtrum guides

Astley et al. [10] confirmed the Hoyme lip philtrum guides differ significantly from the 4-Digit Code Lip-Philtrum Guides resulting in substantially relaxed FAS facial features relative to the 4-Digit Code.

The Hoyme 2016 North American White Lip/Philtrum Guide does not match the “Caucasian” 4-Digit Code Lip-Philtrum Guide 1 (Figure 2A).

**Philtrum:** The Rank 1 through 5 philtrums depicted on both the 4-Digit and Hoyme guides appeared broadly equivalent by visual inspection.

**Upper Lip:** Lip thinness is measured using the objective measure of upper lip thinness ( $\text{circularity} = \text{perimeter}^2 / \text{area}$ ). Circularity confirmed the Hoyme Rank 1, 2, 3, and 4 lips were equivalent to the 4-Digit Ranks 2, 2, 3, and 2 respectively (Figure 2A). The image depicting the vermilion portion of the Hoyme Rank 5 upper lip is not sufficiently clear to judge its level of equivalency with the 4-Digit Code Rank 5 lip. Circularity, as demonstrated in a video link (Figure 2C) confirms the Hoyme Rank 4 lip is substantially thicker than the 4-Digit Code Rank 4 lip (e.g., it is equivalent to the 4-Digit Code Rank 2 lip (Figure 2A). Unlike the 4-Digit Code Lip-Philtrum Guide, the lips pictured on the Hoyme Guide do not become progressively thinner with increasing Rank and no lip on the Hoyme Guide is equivalent to the 4-Digit Ranks 1 or 4. Despite the contrasts between the two lip/philtrum guides, both are intended for use on North American Caucasian populations and thus were used to address Objective 2 below

The Hoyme et al. South African Mixed Race Lip/Philtrum Guide (Figure 1D) does not match the “African American” 4-Digit Code Lip-Philtrum Guide 2 in Figure 1B.

**Philtrum:** The Rank 1 through 5 philtrums depicted on both guides appeared broadly equivalent by visual inspection.

**Upper Lip:** The objective measure of upper lip thinness ( $\text{circularity} = \text{perimeter}^2 / \text{area}$ ) confirmed the Hoyme Rank 1, 2, 3, 4 and 5 lips were equivalent to the 4-Digit Ranks 2, 3, 3, 3 and 3 respectively (Figure 1 in Astley et al. [10]). Unlike the 4-Digit Code Lip-Philtrum Guide, the lips pictured on the Hoyme Guide do not become progressively thinner with increasing Rank. There are no lip images on the Hoyme Guide that correspond to the 4-Digit Ranks 1, 4 or 5. The Hoyme Rank 5 lip is thicker (circularity 40.1) than the Hoyme Rank 4 lip (circularity 46.0). Most importantly, the Hoyme Rank 4 lip (the clinical cut-off for FAS) is thicker than the 4-Digit Rank 4 lip. The Hoyme Rank 4 lip is equivalent to the 4-Digit Rank 3 lip. The Hoyme Rank 5 lip (circularity 40.1) is substantially thicker than the 4-Digit Rank 5 lip (circularity 80).

Based on our findings here and the findings of Hoyme et al. [11], the South African Mixed Race Lip/Philtrum Guide is not appropriate for use on an African American population and thus was not used to address Study Objective 2. The study population for Objective 2 was adjusted accordingly (as described below) to accommodate this finding

### Contrasts in Diagnostic Categories and Nomenclature

The key contrasts in the diagnostic categories and nomenclature used by each system are highlighted in Table 1.

### Contrasts in diagnostic criteria

Key contrasts in diagnostic criteria are highlighted in red font in Table 2.

## Discussion

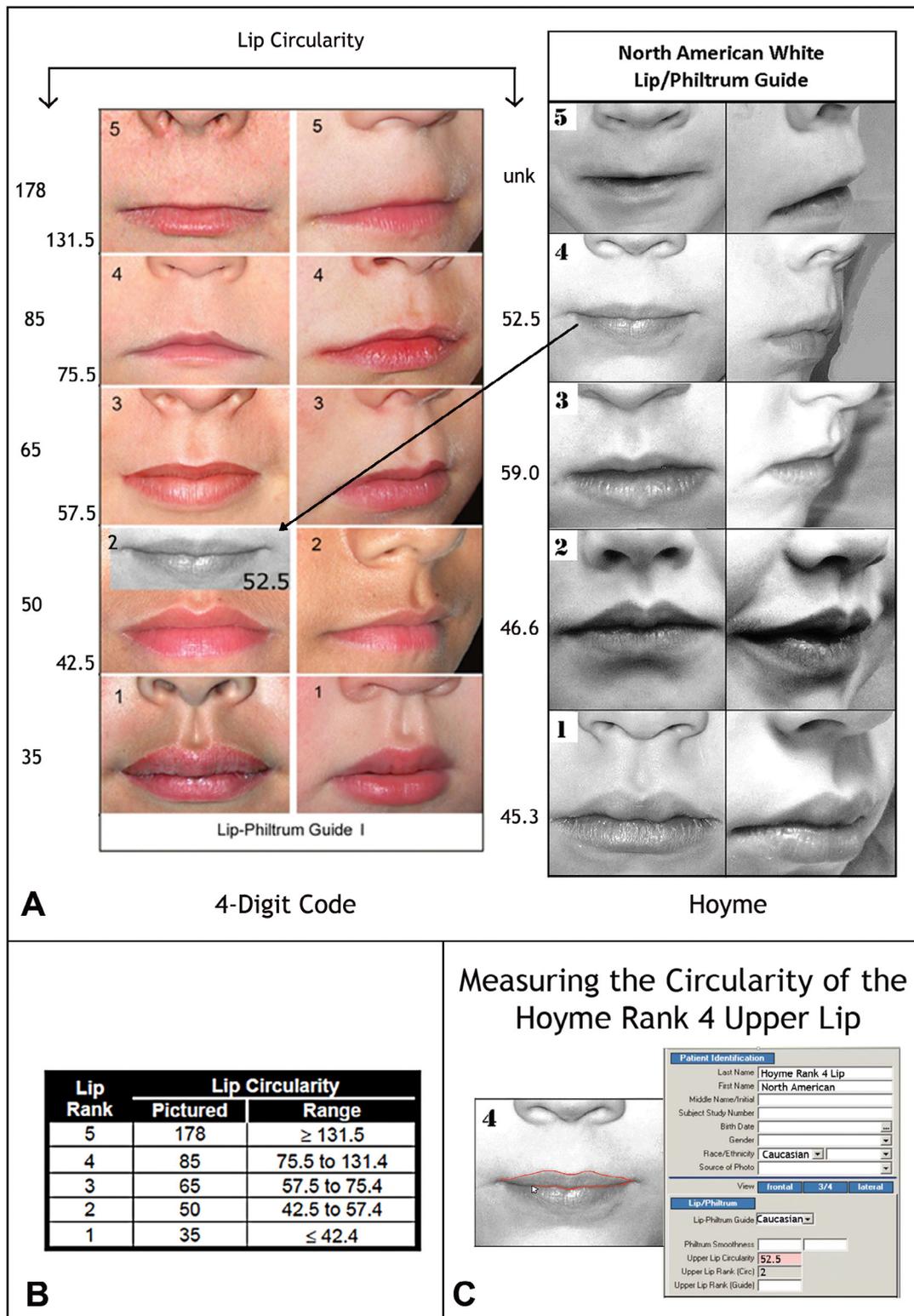
### Growth deficiency

The Hoyme criteria use the same cut-off (prenatal or postnatal height and/or weight  $\leq 10$ th percentile) to define growth deficiency as the 4-Digit Code, but the Hoyme criteria classify growth deficiency on a dichotomous scale (present/absent), whereas the 4-Digit Code ranks growth deficiency on a 4-point ordinal scale with emphasis on short stature. The 4-Digit Code method for ranking growth deficiency is confirmed to be highly predictive of CNS dysfunction among individuals with PAE and appears to differentiate growth deficiency (postnatal short stature) significantly associated with PAE from growth deficiency (low birth weight) significantly associated with prenatal tobacco exposure [9]. Rank 3 and Rank 4 growth deficiency was confirmed to be as highly correlated with, and predictive of, severe brain dysfunction as the 4-Digit Code Rank 4 FAS facial phenotype. Individuals with Rank 3 or 4 growth deficiency had a two to three-fold increased risk for severe brain dysfunction. Sixty percent of patients with Rank 4 growth deficiency had severe brain dysfunction. Growth deficiency is so highly predictive of severe CNS dysfunction among infants/toddlers with PAE, it becomes a vital clinical tool for identifying and qualifying infants/toddlers for early intervention. The Canadian and Australian systems removed growth deficiency from their FASD diagnostic guidelines.

### Facial phenotype

When compared to the 4-Digit Code Rank 4 FAS facial phenotype (used by the 4-Digit Code, Canadian and Australian systems), the Hoyme FAS facial phenotype is substantially relaxed. This is best illustrated using the 4-Digit Code Facial ABC-Score printed on the backside of the 4-Digit Code “Caucasian” Lip-Philtrum Guide 1 in Figure 3A. The 4-Digit Code Rank 4 FAS facial phenotype is defined by a single ABC-Score (Facial ABC-Score CCC, Face Rank 4) (Figure 3A). The three letters “CCC” reflect the magnitude of expression of the short PFL, smooth philtrum, and thin upper lip in that order. C reflects severe expression in the FAS range, B reflects moderate expression, and A reflects normal expression. The Hoyme FAS facial criteria are relaxed relative to the 4-Digit Code in three ways:

1. Only 2 of 3 cardinal features are required.
2. The PFL is relaxed from the 3rd percentile to the 10th percentile.



**Figure 2.** The Hoyme North-American White Lip/Philtrum Guide differs from the 4-Digit Code “Caucasian” Lip-Philtrum Guide 1. The Ranks 1 through 5 philtrums depicted on both Guides appears broadly equivalent, but the upper lips are substantially different. A) Lip circularity (perimeter<sup>2</sup>/area) is printed to the left of each guide. B) The range of circularities that define each 4-Digit Code Lip Rank are presented in the Lip Circularity table printed on the backside of the 4-Digit Code Lip-Philtrum Guide. C) The FAS Facial Photographic Analysis Software [12] computes circularity when the User outlines the vermilion border of the upper lip (click on video link for demonstration <http://depts.washington.edu/fasdpn/movie/Fig2Cvideo.mp4>). Lip circularity confirms the Hoyme Rank 1, 2, 3, and 4 lips are equivalent to the 4-Digit Ranks 2, 2, 3, and 2 respectively. The vermilion portion of the Hoyme Rank 5 upper lip is not sufficiently clear to judge its level of equivalency with the 4-Digit Code Rank 5 lip. There is no lip image on the Hoyme Guide that reflects the 4-Digit Rank 1 or Rank 4 lips. The lips on the 4-Digit Guide become progressively thinner (circularity becomes progressively larger) with increasing Rank. This is not true for the Hoyme Guide. The circularity of the Hoyme Rank 4 lip (the clinical cut-off for FAS) is 52.5, confirming it falls within the circularity range (42.5 to 57.5) of the 4-Digit Code Rank 2 lip. The black and white overlay (A) of the Hoyme Rank 4 lip on the 4-Digit Code Guide 1 demonstrates both visually and numerically that the Hoyme Rank 4 lip is substantially thicker than the 4-Digit Code Rank 4 lip. This analysis confirms the Hoyme North American White Lip/Philtrum Guide is not a valid tool for use with the FASD 4-Digit Diagnostic Code. (North American White Lip/Philtrum Guide used with permission from the American Academy of Pediatrics).

**Table 1.** Diagnostic categories and overlap of nomenclature used by 4 FASD diagnostic systems.

4-Digit Code 2004 [5]	Hoyme et al., 2016 [7]	Canadian 2015 [6]	Australian 2016 [8]
FAS Alcohol Exposed or Unknown	FAS Alcohol Exposed or Unknown	FASD with the Face Alcohol Exposed or Unknown	FASD with the Face Alcohol Exposed or Unknown
pFAS Alcohol Exposed	pFAS Alcohol Exposed or Unknown	FASD without the Face High Alcohol Exposure	FASD without the Face Alcohol Exposed
SE/AE Static Encephalopathy Alcohol Exposed	ARND High Alcohol Exposure Must be ≥ 3 years old		
ND/AE Neurobehavioral Disorder Alcohol Exposed	ARBD High Alcohol Exposure		

**Table 2.** Key contrasts in diagnostic criteria between the four systems.

Criteria	4-Digit Code 2004 [5]	Hoyme et al., 2016 [7]	Canadian 2015 [6]	Australian 2016 [8]
<b>Growth</b>	≤ 10th percentile. Growth: normal, mild, moderate, severe. <b>Emphasis on short stature</b>	≤ 10th percentile Growth: <b>normal/abnormal</b>	<b>Excluded</b>	<b>Excluded</b>
<b>FAS Face</b>	<b>All 3 features PFL ≤ 3rd percentile. Lip &amp; Philtrum Rank 4 or 5 on 4-Digit Code Lip-Philtrum Guides. Face: normal, mild, mod, severe. Specificity: ~95%. Photo Software confirmed more accurate than direct exam.</b>	<b>2 of 3 features. PFL≤10th percentile. Lip &amp; Philtrum Rank 4 or 5 on Hoyme Lip/Philtrum Guides. Face: normal, abnormal. Specificity: ~71%. “We feel that direct exams are more practical in an office setting”</b>	<b>All 3 features. PFL ≤ 3rd percentile. Lip &amp; Philtrum Rank 4 or 5 on 4-Digit Code Lip-Philtrum Guides. Face: normal, abnormal. Specificity: ~95%. Photo Software recommended.</b>	<b>All 3 features. PFL ≤ 3rd percentile. Lip &amp; Philtrum Rank 4 or 5 on 4-Digit Code Lip-Philtrum Guides. Face: normal, abnormal. Specificity: ~95%. Photo Software recommended.</b>
<b>Alcohol Related Birth defects (ARBD)</b>	<b>Excluded</b>	<b>Cardiac: atrial septal defects, aberrant great vessels, ventricular septal defects, conotruncal heart defects; Skeletal: radioulnar synostosis, vertebral segmentation defects, large joint contractures, scoliosis; Renal: aplastic/hypoplastic/dysplastic kidneys, “horseshoe” kidneys/ureteral duplications; Eyes: strabismus, ptosis, retinal vascular anomalies, optic nerve hypoplasia; Ears: conductive hearing loss, neurosensory hearing loss</b>	<b>Excluded</b>	<b>Excluded</b>
<b>Brain structure</b>	Structural/neurological abnormalities. <b>OFC ≤ 3rd percentile. Structure alone meets CNS criteria.</b>	Structural/neurological abnormalities. <b>OFC ≤ 10th percentile. Structure alone does not meet CNS criteria.</b>	Structural/neurological abnormalities. <b>OFC ≤ 3rd percentile. Structure alone does not meet CNS criteria. Serves as 1 of 3 brain domains.</b>	Structural/neurological abnormalities. <b>OFC ≤ 3rd percentile Structure alone does not meet CNS criteria. Serves as 1 of 3 brain domains.</b>
<b>Brain Function</b>	Severe: 3 or more domains ≤ -2 SDs. <b>Moderate: 1-2 domains ≤ -2 SDs and/or 1 or more domains ≤ 1.5 SDs. Function: normal, moderate, severe.</b>	Moderate to Severe <b>1 or more domains ≤ -1.5 SDs. Function: normal, abnormal.</b>	Severe: 3 or more domains ≤ -2 SDs. Function: normal, abnormal.	Severe: 3 or more domains ≤ -2 SDs. Function: normal, abnormal.
<b>Alcohol</b>	Confirmed Exposure (at any reported level) or Unknown Exposure (if 4-Digit Rank 4 FAS face present).	<b>Confirmed High Exposure (≥ 6 drinks/wk for (≥ 2 weeks) or (≥ 3 drinks/occasion, (≥ 2 occasions) or Unknown Exposure (if Hoyme FAS face present).</b>	<b>Confirmed High Exposure (≥ 7 drinks/week) or (≥ 4 drinks/occasion, ≥ 2 occasions) or Unknown Exposure (if 4-Digit Rank 4 FAS face present).</b>	Confirmed Exposure (at any reported level) or Unknown Exposure (if 4-Digit Rank 4 FAS face present).
<b>Children</b>	Diagnostic criteria do not vary with age.	Children ≤ 3 yrs, brain criteria for FAS and PFAS relaxed to developmental delay ≤ -1.5 SDs. Not eligible for a diagnosis of ARND.	Children ≤ 6 yrs: FASD with Face=3 facial features and microcephaly.	Children ≤ 6 yrs: FASD with Face=3 facial features and microcephaly with confirmed or unknown PAE.

## At Risk

5.2.1: Prenatal alcohol exposure, with the estimated dose at a level known to be associated with neurodevelopmental effects; Central nervous system criteria from FASD with or without the Face are not met; and 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).  
 5.22: All 3 facial features present but do not yet have documentation or evidence of the requisite 3 or more neurodevelopmental domain criteria or true microcephaly.  
 7.3: Infants and young children with 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."

Individuals who, despite assessment, fail to meet criteria for FASD at the current time, but may nevertheless potentially have FASD. Example include: Neurodevelopmental assessment is incomplete or inconclusive. Despite confirmed PAE, neurodevelopmental impairment is present in fewer than 3 domains. Neurodevelopmental impairment is present in 3 or more domains, but impairment is not sufficiently severe to meet criteria. Comprehensive, age-appropriate neurodevelopmental assessment is impossible or unavailable e.g., in infants and young children. These individuals may be considered 'at risk of FASD' and require follow-up and reassessment. Confirmed or unknown PAE, <6 yrs, all 3 facial features, do not meet neurodevelopmental criteria and do not have microcephaly.

Readers are referred to the published guidelines for each system for how these criteria are used to generate diagnoses under the umbrella of FASD. Key contrasts are in red font.

3. A Rank 4 or 5 thin upper lip is required, but as illustrated in our analysis above, the Rank 4 lip on the Hoyme North American Lip/Philtrum Guide is equivalent to the Rank 2 lip on the 4-Digit Lip-Philtrum Guide 1.

This results in almost every 4-Digit Code Facial ABC-Score meeting the relaxed Hoyme facial criteria (Figure 3B) including 13 of the 15 ABC-Scores that depict the 4-Digit Code Rank 2 (mild) facial phenotype and 3 of the 8 ABC-Scores that depict the complete absence of all three FAS facial features (Rank 1). Clinically, the 4-Digit Code classifies Rank 1 and 2 facial phenotypes as being within the normal range. The practical clinical impact of this relaxation is illustrated in Figure 3C in which an adolescent with high function (e.g., FSIQ 123) and confirmed absence of PAE met the Hoyme criteria for the full FAS facial phenotype.

In addition to the contrasts in facial criteria, the scales of measurement used to clinically classify the facial phenotype also differ. The 4-Digit Code documents the full continuum of expression of the FAS facial phenotype (Face Ranks 1 through 4); a continuum confirmed to be highly predictive of CNS dysfunction [9,16]. Patients with the Rank 3 facial phenotype have a 2-fold increased risk of severe brain dysfunction, whereas patients with the full Rank 4 FAS facial phenotype have a 5-fold increased risk of severe brain dysfunction. In contrast, the Hoyme system documents the facial phenotype as present (equivalent to 4-Digit Face Ranks 2, 3 and 4 and half of Rank 1) and absent (equivalent to the other half of Rank 1) (Figure 3B). The Canadian and Australian systems adopted the 4-Digit Code Rank 4 FAS facial phenotype

using the 4-Digit Code Lip-Philtrum Guides, but like the Hoyme et al. system, documents the phenotype as present (4-Digit Code Face Rank 4) or absent (4-Digit Code Face Ranks 1-3). The clinical and research impact of dichotomizing the FAS facial phenotype is illustrated below in Objective 2C.

### CNS abnormalities

CNS Functional Abnormalities: The Hoyme criteria that define neurobehavioral impairment appear broadly equivalent to the 4-Digit Code criteria for moderate to severe CNS dysfunction (CNS Ranks 2 and 3). The Canadian and Australian systems adopted the criteria introduced by the 4-Digit Code for severe CNS dysfunction (CNS Rank 3: 3 or more domains of function, 2 or more SDs below the mean). The Canadian and Australian systems exclude moderate dysfunction (the 4-Digit Code equivalent of ND/AE) from under the umbrella of FASD. CNS Structural Abnormalities: The Hoyme criteria for deficient brain growth, abnormal morphogenesis, or abnormal neurophysiology were equivalent to the 4-Digit Code criteria for CNS structural and neurological abnormalities (CNS Rank 4) with the exception of the cut-off used to define microcephaly (Hoyme criteria:  $\leq 10$ th percentile; 4-Digit Code:  $\leq 3$ rd percentile). The Canadian and Australian systems adopted the 4-Digit Code criteria for CNS structural abnormalities, but unlike the 4-Digit Code, do not allow structural abnormalities alone to meet the CNS criteria for FASD. Rather, CNS structural abnormalities must be accompanied by at least two domains of function 2 or more SDs below the mean to meet the Canadian and Australian CNS criteria for FASD.

### FAS Facial Phenotype 4-Digit Code

3 Features Required.

1. PFL  $\leq 3^{rd}$  %
2. Philtrum Rank 4 or 5
3. Lip Rank 4 or 5




FACE TABLES				
5-Point Rank for Philtrum or Lip	Z-scores for Palpebral Fissure Length	ABC-Scores for:		
		Palpebral Fissure	Philtrum	Upper Lip
4 or 5	$\leq -2$ SD	C	C	C
3	$> -2$ SD and $\leq -1$ SD	B	B	B
1 or 2	$> -1$ SD	A	A	A

4-Digit Diagnostic Rank	Level of Expression of FAS Facial Features	Palpebral Fissure – Philtrum – Lip ABC-Score Combinations
3	Moderate	CCB, CBC, BCC
2	Mild	CCA, CAC, CBB, CBA, CAB, CAA BCB, BCA, BBC, BAC ACC, ACB, ACA, ABC, AAC
1	None	BBB, BBA, BAB, BAA ABB, ABA, AAB, AAA

### FAS Facial Phenotype Hoyme 2016

Only 2 features required and 2 of the 3 relaxed relative to 4-Digit Code.

1. PFL  $\leq 10^{th}$  %
2. Philtrum Rank 4 or 5
3. Lip Ranks 2-5

Lip Thickness

See facial image in Figure 2 in Hoyme et al., 2016 [7]



FACE TABLES				
5-Point Rank for Philtrum or Lip	Z-scores for Palpebral Fissure Length	ABC-Scores for:		
		Palpebral Fissure	Philtrum	Upper Lip
4 or 5	$\leq -2$ SD	C	C	C
3	$> -2$ SD and $\leq -1$ SD	B	B	B
1 or 2	$> -1$ SD	A	A	A

4-Digit Diagnostic Rank	Level of Expression of FAS Facial Features	Palpebral Fissure – Philtrum – Lip ABC-Score Combinations
3	Moderate	CCB, CBC, BCC
2	Mild	CCA, CAC, CBB, CBA, CAB, CAA BCB, BCA, BBC, BAC ACC, ACB, ACA, ABC, AAC
1	None	BBB, BBA, BAB, BAA ABB, ABA, AAB, AAA

### North American White Lip/Philtrum Guide

5	Unk
4	52.5
3	59.0
2	46.6
1	45.3

Confirmed absence of Prenatal Alcohol, FSIQ 123

### Hoyme 2016 FAS Facial Phenotype



\* PFL 9th percentile  
\* Philtrum Rank 3  
\* Lip Circularity 85 (Rank 4-5)

Guideline	FAS Facial Features Present			
	FAS Face	PFL	Philtrum	Lip
Hoyme	Yes	Yes	No	Yes
4-Digit	No	No	No	Yes
	Rank 2	B	B	C

**Figure 3.** The Hoyme FAS facial phenotype is substantially relaxed relative to the 4-Digit Code. (A) The 4-Digit Code FAS facial phenotype is defined by the Facial ABC-Score “CCC” as depicted in the Face Table on the backside of Lip-Philtrum Guide 1. (B) The relaxed criteria for the Hoyme FAS facial phenotype results in almost every 4-Digit Code Facial ABC-Score meeting the relaxed Hoyme facial criteria [10]. The prevalence of the FAS facial phenotype was 10-fold higher using the Hoyme criteria (n=552; 40%) compared to the 4-Digit Code (n=54; 4%). (C) The practical clinical impact of this relaxation is illustrated in which an adolescent with high function (e.g., FSIQ 123) and confirmed absence of PAE met the Hoyme criteria for the full FAS facial phenotype. Copyright Susan Astley Hemingway, University of Washington.

## Alcohol exposure

The Hoyme 2016 criteria for documented PAE are more stringent than the 4-Digit Code and include thresholds ( $\geq 6$  drinks/week for  $\geq 2$  weeks during pregnancy or  $\geq 3$  drinks per occasion on  $\geq 2$  occasions during pregnancy). The 4-Digit Code requires a confirmed exposure, but does not set thresholds because: 1) recall and reporting of quantity, frequency, and timing of exposure have been confirmed highly unreliable in a clinical setting (especially in populations like the FASDPN clinic where 85% of patients are not in their birth mother's care at the time of the evaluation); 2) details on quantity, frequency and timing are often unavailable; 3) exposure below a designated threshold has not been confirmed safe for all fetuses [17]; and 4) a recent twin study confirmed risk is not just determined by amount of exposure-fetal genetics modifies risk [18]. The Hoyme system allows FAS and PFAS to be diagnosed when exposure is unknown because the Hoyme FAS facial phenotype is required to be present. The Hoyme FAS facial phenotype, however, is only 71% [19] specific to PAE. The 4-Digit Code allows FAS to be diagnosed when exposure is unknown because FAS requires the presence of the Rank 4 FAS facial phenotype and the Rank 4 face is confirmed to be highly specific (95% specificity) to PAE [20]. The Australian system adopted the alcohol exposure criteria used by the 4-Digit Code. The Canadian system, in contrast, requires high exposure ( $\geq 7$  drinks/week or  $\geq 4$  drinks per occasion on  $\geq 2$  occasions) when the Rank 4 FAS facial phenotype is absent. In the current study population, of the 1,177 with confirmed PAE, only 46% met the Hoyme or Canadian threshold for high exposure.

## OBJECTIVE 2. COMPARISON OF DIAGNOSTIC OUTCOMES ACROSS THE FOUR SYSTEMS

### Methods

#### Study population

The records of 1,392 patients were drawn from 1,522 consecutive patients that received an FASD diagnostic evaluation at the University of Washington Fetal Alcohol Syndrome Diagnostic & Prevention Network (FASDPN). The diagnostic evaluations were performed by interdisciplinary teams between 1993 and 2012 using the FASD 4-Digit Code [5]. The interdisciplinary teams included a medical doctor, psychologist, occupational therapist, speech language pathologist, social worker, family advocate, and public health professional [17,21]. All patients with one or both birth parents African American (130 of the 1,522) were excluded from the study because it was unclear which PFL normal growth chart to use for African Americans when applying the Hoyme system [22] and our findings in Astley et al. [10] and those reported by Hoyme [11] confirm the South African Mixed Race Lip/Philtrum Guide is inappropriate for use on an African American population.

Historically, all records resulting from each patient's FASD diagnostic evaluation have been entered into a research database since 1992 with University of Washington Human Subjects approval and patient consent. Over 95% of patients provide consent for their clinical data to be used for research purposes. Patients' records include the following standardized 4-Digit Code data forms: the New Patient Information Form, the FASD

Diagnostic Form, digital facial photos, and the FAS Facial Photographic Analysis Report [5,12]. These data are entered into a research database shortly after the patient's FASD diagnostic evaluation reflecting the tools and growth norms available at that time. Over the decades the 4-Digit Code has evolved (1st edition 1997, 3rd edition 2004) [5,23-25], new tools have been developed like the FAS Facial Photographic Analysis Software (Version 1.0 in 2004, Version 2.1 in 2016) [12], and new more accurate growth norms have been adopted (CDC [26] and WHO [27] growth charts and Stromland Scandinavian PFL charts [28]).

For the purposes of research, all patients' clinical 4-Digit Codes are updated to "research" 4-Digit Codes to reflect the most current tools and norms available at the time of the research study. For this study, all 4-Digit Codes were updated to reflect the most current 2004, 3rd edition of the 4-Digit Code [5].

### Application of the diagnostic tools and norms

The following tools and norms were used to update the 4-Digit Code FASD diagnoses and generate the Hoyme [7], Canadian [6] and Australian [8] FASD diagnoses

#### Growth

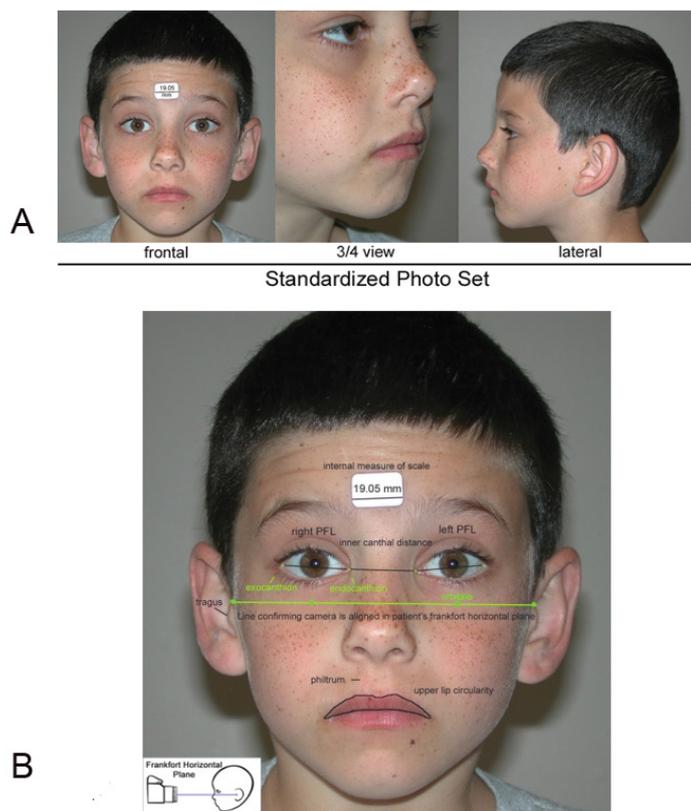
The Hoyme criteria use the same cut-off (prenatal or postnatal height and/or weight  $\leq 10$ th percentile) to define growth deficiency as the 4-Digit Code, thus all patients with 4-Digit Code Growth Ranks 2,3 or 4 were classified as meeting the Hoyme growth deficiency criteria.

**Height and weight normal growth charts:** Height and weight percentiles were generated from the Hall [29] birth weight and length growth charts by gestational age; the World Health Organization (WHO) [27] height and weight growth charts for children 0-2 years of age, and the Centers for Disease Control (CDC) 2000 [26] height and weight growth charts for patients 2 years of age and older. The height percentile was adjusted for mid-parental height [30] when both parents' heights were reported. The Canadian and Australian systems excluded growth deficiency as a criterion for FASD.

#### Facial features

At the time of each patient's FASD diagnostic evaluation, three standardized, digital facial photographs (Figure 4) were taken and measured using the FAS Facial Photographic Analysis Software [12]. As a result, each patient's research record included the following facial measures: PFLs in millimeters, philtrum smoothness (Rank 1 to 5 on the 4-Digit Code Lip-Philtrum Guide1) and upper lip circularity (perimeter<sup>2</sup>/area) and corresponding Lip Rank (Rank 1 to 5 on the 4-Digit Code Lip-Philtrum Guide 1).

**Palpebral fissure length:** For the 4-Digit Code and Hoyme systems, PFL z-scores were updated to reflect the Stromland Scandinavian PFL growth charts [28]. The Stromland charts are confirmed valid for use on a North American population [13] and address the full age span (birth through adult) represented in our study population. In addition, the Stromland PFL growth charts were generated from digital images, thus meeting the recommendation by Hoyme [7] that PFLs measured from photos should be compared to PFL normal growth charts generated from photos. The Hoyme system cites the Canadian PFL charts [14] as one of several published norms obtained from 2-dimensional photography that one may use, but



**Figure 4.** The FAS Facial Photographic Analysis Software [12] was Used to Measure the 3 FAS Facial Features.

A) The palpebral fissure length (PFL), philtrum smoothness, and upper lip thinness are measured from three standardized, digital photographs. B) Standardization includes proper rotation, exposure, focus, and facial expression. An internal measure of scale (a 3/4 inch (19.05 mm) paper sticker) is placed on the forehead to measure the PFLs in millimeters. A video demonstration of the software can be viewed at this link: <http://depts.washington.edu/fasdnpn/movie/software1024-768cd2.mp4>. Copyright Susan Astley Hemingway, University of Washington.

the Canadian norms start at 6 years of age. As demonstrated in Astley et al. [13] transition from the Stromland PFL norms to the Canadian PFL norms at 6 years of age results in an abrupt, artificial decrease in the prevalence of short PFLs due to the discrepancy between the two norms. To avoid this artifact, the Stromland PFL charts that span the entire lifespan were used for the Hoyme system. In accordance with the Canadian and Australian systems, the Canadian PFL growth charts [14] were used for patients 6 years of age and older. The Stromland growth charts [28] were used for patients less than 6 years of age.

**Philtrum smoothness and upper lip thinness:** The 4-Digit Code “Caucasian” Lip-Philtrum Guide 1 (Figure 1A) was used to Rank philtrum smoothness and upper lip thinness for the 4-Digit Code, Canadian and Australian systems. The Hoyme North American Lip/Philtrum Guide (Figure 1C) was used to rank philtrum smoothness and lip thinness for the Hoyme et al. system. Since the images depicting the Rank 1 through 5 philtrums on the 4-Digit Code and Hoyme guides appeared broadly equivalent (per Objective 1), the philtrum rank assigned at the time of diagnosis using the 4-Digit Code guide was the same philtrum rank assigned to the patient using the Hoyme guide (Figure 1C) (e.g., if the patient had a Rank 4 philtrum using the 4-Digit Code guide, they received a Rank 4 philtrum using the Hoyme guide). In contrast, the analyses in Objective 1 [10] confirmed the Rank 1 through 5

images depicting upper lip thinness did not match between the 4-Digit Guide 1 and the Hoyme North American Guide (Figure 2A). The 4-Digit Code uses the full range of Lip Ranks 1-5 to classify the FAS facial phenotype on a 4-point Likert scale from normal (Face Rank 1) to severe FAS (Face Rank 4). In contrast, the Hoyme FAS/PFAS facial criteria measure lip thinness on a dichotomous scale (thin:  $\geq$  Rank 4, not thin:  $<$ Rank 4 on the Hoyme North American Lip/Philtrum Guide (Figure 1C) to classify the FAS/PFAS facial phenotype on a dichotomous scale (present, absent). To accurately and objectively identify which patients met the Hoyme diagnostic criteria for a thin upper lip ( $\geq$ Rank 4), the Rank 4 upper lip on the Hoyme North American Lip/Philtrum Guide was outlined using the facial software’s circularity tool. The video clip in Figure 2C demonstrates this procedure. The circularity of the Hoyme Rank 4 lip was 52.5; equivalent to the 4-Digit Rank 2 lip (defined by the circularity range 42.5 to 57.4). Thus all patients with an upper lip circularity of 52.5 or greater met the Hoyme criteria for a thin upper lip (Rank 4 or 5 on the Hoyme North American Lip/Philtrum Guide).

### CNS dysfunction

Based on our findings in Objective 1, all patients with 4-Digit Code CNS Ranks of 2 or 3 (moderate or severe CNS dysfunction) were classified as broadly equivalent to the Hoyme criteria for neurobehavioral impairment (at least 1 domain 1.5 SDs below the mean). All patients with 4-Digit Code CNS Rank 3 (severe dysfunction) were classified as meeting the Canadian and Australian criteria for severe dysfunction (3 or more domains of function, 2 SDs below the mean). All patients with 2 domains of severe dysfunction and microcephaly ( $OFC \leq 3$ rd percentile) also met the Canadian and Australian criteria for severe dysfunction.

### CNS structural abnormalities

Based on our findings in Objective 1, all patients with a 4-Digit Code CNS Rank 4 (structural/neurological abnormalities) were classified as meeting the Hoyme criteria for deficient brain growth, abnormal morphogenesis, or abnormal neurophysiology. In addition, all patients with an  $OFC \leq 10$ th percentile were classified as meeting the Hoyme CNS structural criteria. All patients with a 4-Digit Code CNS Rank 4 (structural/neurological abnormalities) were classified as meeting the Canadian criteria for impairment in neuroanatomy or neurophysiology and the Australian criteria for abnormal brain structure/neurology. In contrast to the Hoyme et al. system, the 4-Digit Code, Canadian and Australian systems use a cut-off of  $\leq 3$ rd percentile for microcephaly. The WHO [27] OFC charts for children 0-5 years of age and the Nellhaus [31] OFC growth charts for children 5-18 years of age were used for all four systems

### Statistical analyses

Descriptive statistics (valid percentages) were used to profile the study population. Chi-square tests were used to compare groups and linear trends across groups for outcomes measured on nominal or ordinal scales. One-way analysis of variance (ANOVA) was used to compare means and detect linear trends across three or more groups when outcomes were measured on a continuous scale. T-tests were used to compare means between two independent groups.

Various measures of performance (validity) were administered

to each system to address Objective 2C. Validity is the degree to which a tool (or diagnostic system) is measuring what it purports to measure. Validity is not determined by a single statistic, but by a body of research that demonstrates the relationship between the diagnostic system and the condition it is intended to measure. There are three overarching forms of validity: content validity, criterion validity, and construct validity. Content Validity is a measure of how well the items in the diagnostic system represent the entire range of possible items the diagnostic system should cover. Criterion validity is a measure of a diagnostic tool's accuracy relative to a gold standard. Construct validity refers to the degree to which a test measures what it claims, or purports, to be measuring. It refers to the ability of a measurement tool to measure the physiological concept being assessed. Convergent and discriminant validity are two subtypes of construct validity. Convergent validity refers to the degree to which two measures of constructs that theoretically should be related are in fact related. In contrast, discriminant validity tests whether concepts or measurements that are supposed to be unrelated are in fact unrelated. An important aspect of clinical research is the inference that an association represents a cause-effect relationship. Features of associations that support causation include: the strength of the association; the consistency of observed evidence; specificity of the relationship; temporality of the relationship; the biological gradient of dose-response, biological plausibility; and experimental confirmation. Predictive validity refers to a tool's ability to predict something it should theoretically be able to predict. Statistical measures used to assess these constructs include linear correlation coefficients and tests for trends. Fundamental measures of diagnostic accuracy include sensitivity and specificity. The sensitivity of a test is the proportion of people with the condition who test positive for it (the true positive rate). The specificity of a test is the proportion of people who do not have the condition who test negative for it (the true-negative rate).

## Results

### Study population

The clinical and socio-demographic profile of the study population (N=1,392) is presented in Table 3. The population spanned the entire age range from newborn to adult with 57% Caucasian and 44% female. Eighty-five percent had confirmed PAE; 15% had unknown PAE. Patients with unknown PAE were included because all four diagnostic systems allow a diagnosis of FAS when PAE is unknown. Since the publication of the 2017 study comparing the 4-Digit Code to the Hoyme system [10], updated information became available on 2 of the 1,392 patients, impacting the distribution of diagnoses generated by the two systems by a fraction of a percent in this study relative to the 2017 study.

### Objective 2a: Compare the prevalence of FASD diagnostic outcomes generated by the four systems

The distribution of diagnoses varied substantially across the 4 systems as illustrated in Table 4 and Figures 5A and 5B.

The proportion of patients diagnosed with FAS and FASD varied significantly across the systems (4-Digit 2.1% and ≤79%; Australian 1.8% and 29%; Canadian 1.8% and 16%; and Hoyme 6.4% and 44% (Figure 5A). Even though the proportion of patients

diagnosed with FAS (1.8%-2.1%) by the 4-Digit, Canadian and Australian systems was comparable, the patients that made up the 2% within each system were different (see Objective 2b). The distribution of diagnoses also varied substantially across the four systems among the subset of patients <6 years of age at the time of diagnosis (Figure 5B). Key factors contributing to the diagnostic variability include:

- 1) The Canadian and Australian systems exclude moderate dysfunction as an outcome caused by PAE. This resulted in the greatest magnitude of diagnostic variability between the 4 systems. Exclusion of moderate dysfunction prevented a Canadian diagnosis of FASD in 666 patients with moderate dysfunction and confirmed PAE (48% of whom had confirmed high PAE). 76% had 1 or 2 (but not 3) domains of severe dysfunction and all had multiple domains of moderate dysfunction. Exclusion of moderate dysfunction prevented an Australian diagnosis of FASD in 642 patients with moderate dysfunction and confirmed PAE (50% of whom had confirmed high PAE). 74% had 1 or 2 (but not 3) domains of severe dysfunction and all had multiple domains of moderate dysfunction. Primate research confirms moderate dysfunction (ND/AE) is the most prevalent outcome caused by PAE (5% FAS/PFAS, 31% SE/AE; 59% ND/AE, 5% not FASD) (Figure 6). Only the 4-Digit Code replicated this distribution of diagnoses observed in the primate model of FASD as discussed more fully below. Fifty-three percent of the 1,177 patients with confirmed PAE received a diagnosis of ND/AE using the 4-Digit Code (65% were over 6 years of age).
- 2) The Canadian and Hoyme systems require confirmed high exposure to alcohol in the absence of the FAS facial phenotype. This prevented 47% of 1,155 patients with confirmed PAE, but without the FAS facial phenotype from receiving a FASD diagnosis using the Canadian system. Forty-three percent of these 548 patients had 1 to 2 domains of severe dysfunction; 34% had 3 or more domains of severe dysfunction. The requirement for high PAE also prevented 59% of 664 patients with confirmed PAE, but without the Hoyme FAS facial phenotype, from receiving a FASD diagnosis using the Hoyme system. Forty-two percent of these 389 patients had 1 to 2 domains of severe dysfunction; 33% had 3 or more domains of severe dysfunction. For reference, the proportion of patients with confirmed high PAE (Alcohol Rank 4) within each FASD diagnosis rendered by the 4-Digit Code is marked by yellow lines (Figures 5A and 5B).
- 3) The Australian and Canadian systems excluded growth deficiency as a criterion for FASD. This prevented the early identification of 70% of children <8 years old with confirmed PAE and growth deficiency as especially high risk for severe brain dysfunction (3 or more domains of function 2 or more SDs below the mean) later in childhood [9]. More specifically, of the 770 patients classified as Not FASD by the Australian system, 559 had confirmed PAE. Of the 559 with confirmed PAE, 221 were under 8 years of age and 69 presented with Growth Rank 2, 3 or 4 (height and/or weight at or below the 10th percentile). Seventy-percent of these children with confirmed PAE and growth deficiency will present with severe CNS dysfunction later

**Table 3.** Sociodemographic and 4-Digit Code clinical profile of the study population (n=1,392).

Characteristic	N	Valid %
<b>Gender</b>		
female	608	44
male	784	56
<b>Race/ethnicity</b>		
Caucasian	788	57
Native American	126	9
Hispanic	37	3
African American	0	0
Other (including mixed race)	434	31
<b>Age at FASD diagnostic evaluation (years)</b>		
0-2	141	10
3-5	314	23
6-7	234	16
8-12	411	30
13-19	241	17
20-49	51	4
<b>4-Digit Code Diagnoses (and Categories)</b>		
FAS/AE or A? (A,B)	29	2
PFAS/AE (C)	53	4
SE/AE (E,F)	388	28
ND/AE (G,H)	624	45
SPF/AE (I)	22	1
Normal/AE (J)	69	5
Not FASD/A? (D, K-V)	207	15
<b>Growth Rank</b>		
Normal (height & weight > 10th percentile): 1	954	68
Mild (height and/or weight ≤ 10th but both > 3rd percentile): 2	176	13
Moderate (height or weight ≤ 3rd percentile): 3	161	12
Severe (height & weight ≤ 3rd percentile): 4	101	7
<b>Face Rank</b>		
Normal (no features): 1	705	51
Mild (1-2 features): 2	530	38
Moderate (2.5 features): 3	103	7
Severe (all 3 features): 4	54	4
<b>CNS Rank</b>		
No structural/functional abnormalities: 1	109	8
Moderate dysfunction (1-2 domains ≤ -2 SDs): 2	739	53
Severe dysfunction (3 or more domains ≤ -2 SDs): 3	307	22
Severe structural/neurological abnormalities: 4	237	17
<b>CNS Functional Rank*</b>		
No dysfunction: 1	171	12
Moderate dysfunction: 2	829	60
Severe dysfunction: 3	392	28
<b>Alcohol Rank</b>		
Prenatal Alcohol Exposure (PAE) confirmed absent: 1	0	0
PAE Unknown: 2	215	15
PAE confirmed: :Level unknown: 3	198	14
PAE confirmed: :Level reported moderate: 3	353	26
PAE confirmed: :Level reported high: 4	626	45

\* Includes all 1,392 subjects including the 236 with CNS Rank 4 structural/neurological abnormalities. Abbreviations: AE: alcohol exposed; A?: alcohol exposure unknown; ND neurobehavioral disorder; SD standard deviations; SE static encephalopathy; SPF sentinel physical features

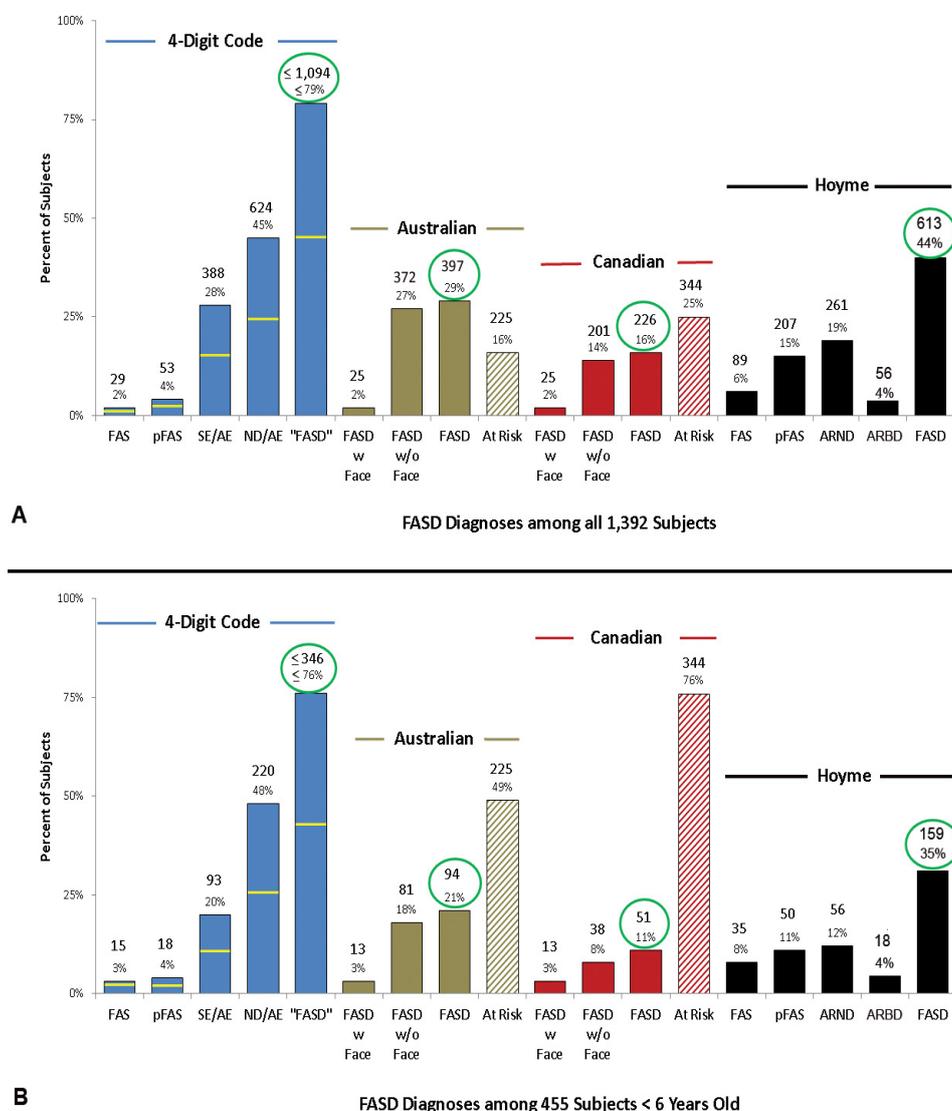
in childhood when they are old enough to participate in a comprehensive neuropsychological assessment. None were identified as “At Risk” by the Australian system. Of the 822 children classified as Not FASD by the Canadian system, 611 had confirmed PAE. Of the 611 with confirmed PAE, 176 were under 8 years of age and 48 presented with Growth Rank 2, 3 or 4. Seventy-one percent of these children with confirmed PAE and growth deficiency will present with severe CNS dysfunction after the age of 8 years. None were identified as “At Risk” by the Canadian system.

- 4) Relaxation of the FAS facial criteria by the Hoyme system (Figure 3), resulted in ten times more patients presenting with the “Hoyme FAS facial phenotype” (552, 40%) than the Rank 4 FAS facial phenotype (54, 4%) used by the 4-Digit Code, Canadian and Australian systems [10]. Seventy-one percent of the 552 patients with the Hoyme FAS facial phenotype had 4-Digit Code Face Ranks 1 and 2. The relaxed Hoyme criteria also resulted in a clinically significant reduction in facial specificity (71% to 75%) [19,32] to alcohol relative to the Rank 4 FAS facial phenotype (>95% specificity) [10]. Of the 552 patients with the Hoyme FAS facial phenotype, almost half (43%) did not receive a diagnosis under the umbrella of FASD using the Hoyme system. In contrast, all 54 patients with the 4-Digit Code Rank 4 FAS face met criteria for a diagnosis under the umbrella of FASD using the 4-Digit Code. More details on these outcomes are presented in Astley et al. [10].
- 5) Switching from the Stromland PFL growth charts to the Clarren PFL growth charts at 6 years of age, as recommended by the Canadian and Australian systems, can result in the FAS facial phenotype appearing to “disappear” at age 6 years. Although the Stromland PFL growth charts [28] span the entire age range from birth to adult, the Canadian and Australian systems recommend use of the Clarren PFL normal growth charts [14] that start at age 6 years and the Stromland PFL charts for children under 6 years of age. This results in an artificial decrease in the prevalence of short PFLs and the FAS facial phenotype in children >6 years of age because the mean PFL for age in the Clarren charts is roughly half a SD larger than the PFL in the Stromland charts in Figure 2 in [13]. A PFL of 23 mm in a 6 year old boy is -2.1 SDs on the Stromland charts, but -1.6 SDs on the Clarren charts. To illustrate the impact this has on diagnostic outcomes, of the 30 patients >6 years of age with the Rank 4 FAS facial phenotype using the Stromland PFL charts, only 21 met the Rank 4 FAS facial phenotype criteria using the Clarren PFL charts. This is a 30% reduction in prevalence of the FAS facial phenotype. The discrepancy between the two charts will also result in the FAS facial phenotype appearing to “disappear” with age. If a child presents with the FAS facial phenotype using the Stromland charts at age 5 years, the child will appear to lose the FAS facial phenotype upon re-evaluation at 6 years of age as a result of switching to the Clarren PFL charts. The 4-Digit Code recommends use of the Stromland PFL

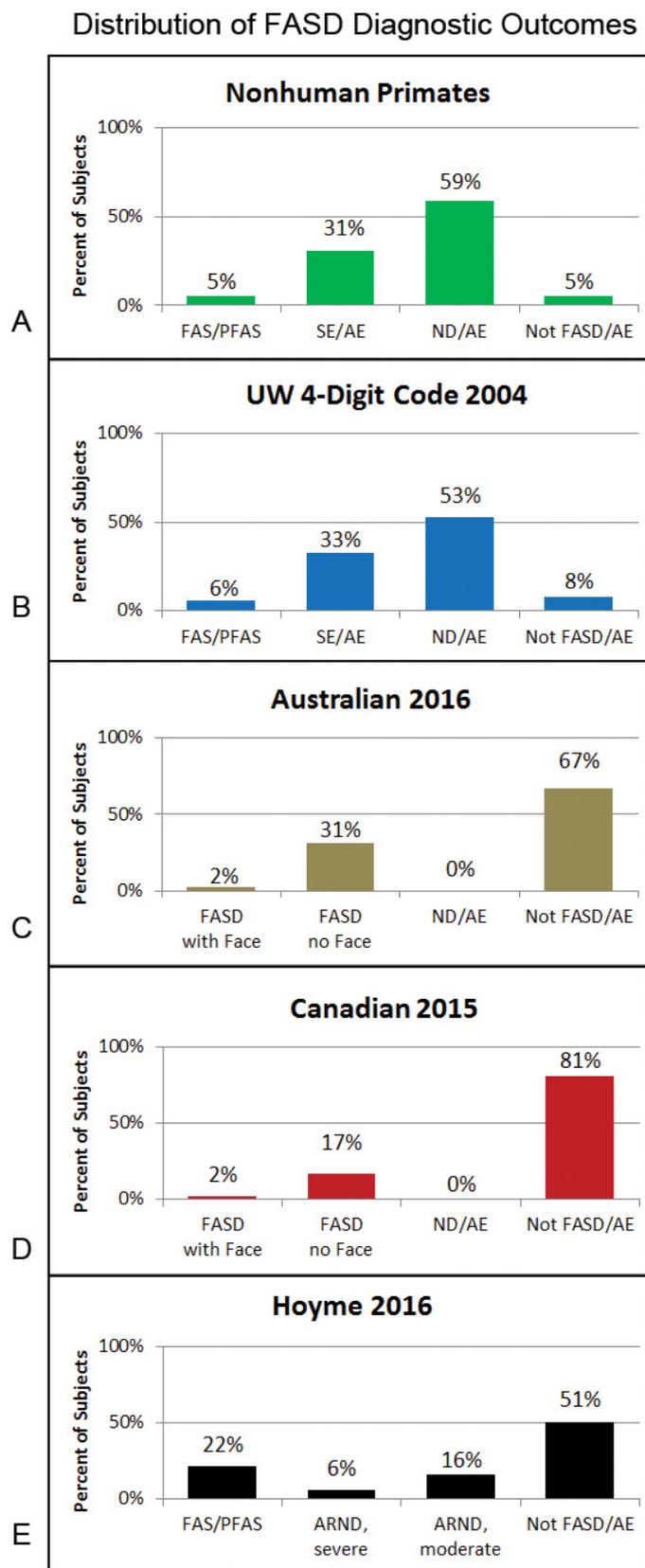
**Table 4.** Prevalence and concordance of FASD diagnoses across the four diagnostic systems.

Variables	Diagnoses generated by the four systems and the various names applied to each													
	FAS, FASD/with Face		PFAS		SE/AE, ARND-severe*, FASD/no Face		ND/AE, ARND-moderate**		ARBD		"FASD" FASD		Not FASD (includes At-Risk)	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
<b>Number diagnosed by each system</b>														
4 Digit	29	2.1	53	3.8	388	27.9	624	44.8			1094	78.6	299	21.4
Australian	25	1.8			372	26.7					397	28.5	995	71.5
Canada	25	1.8			201	14.4					226	16.2	1166	83.8
Hoyme	89	6.4	207	14.9	69	5.0	192	13.8	56	4	613	44	779	56.0
<b>Number diagnosed by at least 1 system</b>														
	107	7.7	241	17.3	430	30.9	624	44.8	56	4	1138	81.8	1240	89.0
<b>Number diagnosed by all 4 systems</b>														
	12	0.9	19	1.4	53	3.8	0	0	0	0	152	10.9	235	16.9

\* ARND-severe: patients with 3 or more functional domains -2 SDs below the mean. \*\* ARND-moderate: patients with 2 or more functional domains -1.5 SDs below the mean, but less than 3 functional domains -2 SDs below the mean. "FASD": 4-Digit Code includes FAS and PFAS under the FASD umbrella, but notes SE/AE and ND/AE are only FASDs if a patient's prenatal alcohol exposure caused their SE or ND.



**Figure 5.** FASD diagnostic outcomes are compared across the four FASD diagnostic systems. (A) Diagnoses across the entire population (n=1,392). (B) Diagnoses across the subset of 455 patients less than 6 years of age at the time of diagnosis. The yellow lines on the blue bars reflect the proportion of patients with confirmed high PAE (4-Digit Code Alcohol Rank 4). The bars labeled FASD for each system reflect the total number of patients diagnosed under the umbrella of FASD by each system. The term "FASD" is in quotes for the 4-Digit Code to denote that the 4-Digit Code defines FASD as including FAS, PFAS and only those individuals whose SE or ND was caused (at least in part) by their prenatal alcohol exposure.



**Figure 6.** Nonhuman-primate study confirms moderate dysfunction is the most prevalent outcome under the umbrella of FASD.

(A) The 4-Digit Code was applied to the outcomes observed in our highly controlled primate model of FASD [4] where PAE was the only risk factor. Moderate dysfunction (ND/AE) was the most prevalent outcome (59%). (B) The 4-Digit Code was the only diagnostic system that replicated the distribution of diagnoses observed in the primate model. (C-D) The Australian and Canadian systems omit moderate dysfunction from FASD. (E) The bar charts reflect the distribution of diagnostic outcomes across the 4 systems among the 1,177 patients with confirmed PAE.

**Abbreviations:** AE: Alcohol Exposed; ARND: Alcohol Related Neurodevelopmental disorder; ND: Neurobehavioral Disorder; SE: Static Encephalopathy. ARND-severe reflect the subset of patients meeting the Hoyme ARND criteria that have 3 or more domains of function  $\leq 2$  SDs below the mean (rendering it comparable to SE/AE and FASD/no Face. ARND-moderate is the remainder of patients meeting the Hoyme ARND criteria that have less than 3 domains  $\leq 2$  SDs below the mean.

charts across the full age span [13] to avoid these artifacts in measurement.

- 6) The Hoyme system includes Alcohol Related Birth Defects (ARBD/AE) under the umbrella of FASD; the other systems do not. Fifty-six individuals met the Hoyme criteria for ARBD/AE (Figure 5A). Of the list of defects that meet the Hoyme criteria for ARBD/AE (Table 2), four types of defects were observed among the 276 patients who met the Hoyme alcohol criteria, but did not meet the Hoyme criteria for FAS, PFAS or ARND. The number of patients presenting with each feature was as follows: strabismus (5), ptosis (36) cardiac anomalies (12) and scoliosis (8). Seven of the 56 patients presented with two of these features. The reported prevalence of these features across the entire study population of 1,392 patients was ptosis (9.0%), cardiac anomalies (3.9%), strabismus (0.5%) and scoliosis (0.4%). Cardiac anomalies were significantly more prevalent among patients receiving a FASD diagnosis (6.5%) using the Hoyme system than among those not receiving a FASD diagnosis (1.9%) ( $\chi^2$  19.2,  $p=0.000$ ). Ptosis was also significantly more prevalent among patients receiving a FASD diagnosis (14.4%) using the Hoyme system than among those not receiving a FASD diagnosis (4.6%) ( $\chi^2$  38.7,  $p=0.000$ ). Cardiac anomalies and ptosis were also significantly more prevalent among patients with FASD than without FASD when the other three systems (4-Digit Code, Australian and Canadian) were used to generate the FASD diagnoses. None of these anomalies were significantly correlated with any measure of PAE available in the FASDPN dataset.

#### Objective 2b: Assess diagnostic discordance/concordance between the four systems

Very little diagnostic concordance was observed across all four diagnostic systems. Of the 1,392 patients, 1,138 (82%) were diagnosed with FASD by at least one of the four systems (Table 4). In contrast, only 152 (11%) were diagnosed with FASD by all four systems. Of the 107 (8%) diagnosed with FAS by at least one of the 4 systems, only 12 (1%) were diagnosed FAS by all four systems.

The patient-by-patient diagnostic outcomes generated by the 4-Digit Code were compared directly with the diagnoses generated by the Hoyme (Figure 7) Canadian (Figure 8) and Australian (Figure 9) systems. The Canadian system was also compared directly with the Australian system (Figure 10) and the Hoyme system in (Figure 11). Of the 1,392 patients, concordant diagnoses (including those being classified as "Not FASD") were as follows: 4-Digit vs Canadian: 31%; 4-Digit vs Hoyme: (38%); 4-Digit vs Australian: (45%); Canadian vs Hoyme (39%) and Canadian vs Australian: (82%). The higher level of concordance between the Canadian and Australian systems is due to the fact that the Australian system adopted the criteria used by the Canadian system, with one important exception. The Canadian system requires confirmed high PAE. The Australian system requires confirmed PAE at any reported level. The higher

level of concordance between the Canadian and Australian systems was due largely to the high proportion (66%, 918/1,392) of patients classified as not under the umbrella of FASD (“At Risk” and “Not FASD”).

The discordance across the systems ranged from subtle differences (e.g., the patient received a diagnosis of FAS by one system and PFAS by another system) to marked contrasts (e.g., the patient received a diagnosis of FAS by one system and no diagnosis under the umbrella of FASD by another system). A few examples of these marked contrasts include the following. Additional contrasts are presented in the legends for Figures 7-11.

1. Of the 21 patients that received a diagnosis of FAS/Alcohol Exposed using the 4-Digit Code, 7 had FASD ruled-out altogether using the Hoyme system (see the 4-Digit Code FAS/AE column in Figure 7). All 7 patients were less than 5 years of age. They presented with CNS structural abnormalities (e.g., microcephaly: OFC $\leq$ 3rd percentile), but early development was broadly within the normal range. All 7 were too young to engage in the necessary level of testing to accurately rule-out moderate or severe CNS dysfunction. The Hoyme system requires both CNS structural abnormalities (e.g., OFC $\leq$ 10th percentile) and evidence of moderate to severe CNS dysfunction for a diagnosis of FAS.
2. Among the 207 patients that were classified “Not FASD” by the 4-Digit Code, 15 received a FAS diagnosis and 23 received a PFAS diagnosis using the Hoyme system (Figure 7). The 4-Digit Code does not render a diagnosis under the umbrella of FASD if: 1) alcohol exposure is unknown and 2) the Rank 4 FAS facial phenotype is absent. If an individual does not have a confirmed PAE, the 4-Digit Code Rank 4 FAS face can serve as confirmation of exposure because the phenotype is confirmed to be so highly specific to (caused only by) PAE (> 95% specificity) [17]. The Hoyme system allowed these 38 patients with unknown alcohol exposures to receive a diagnosis of FAS or PFAS because they presented with the Hoyme FAS face. But the Hoyme FAS facial criteria are so relaxed (specificity 71% to 75% [19,32]), the facial phenotype does not provide the necessary level of specificity to alcohol to use the facial phenotype to confirm exposure. Among the 38 individuals with unknown PAE and a Hoyme diagnosis of FAS or PFAS, 18 had relaxed PFLs (4th-10th percentile), 16 had relaxed philtrums (4-Digit Philtrum Ranks 2 and 3), 22 had relaxed lips (4-Digit Lip Ranks 1-3); 4 had no FAS facial features (4-Digit Face Rank 1); and 19 had only 1 FAS facial feature (4-Digit Face Rank 2).
3. Among the 779 patients that were classified “Not FASD” using the Hoyme system, 24 received a FAS/PFAS diagnosis using the 4-Digit Code (Figure 7, red bars in the Hoyme “Not FASD” row). All 24 presented with the Hoyme FAS face, but none met the Hoyme FAS or PFAS diagnostic criteria. The Hoyme FAS criteria require the presence of both CNS structural abnormalities (e.g., OFC $\leq$ 10th percentile) and neurobehavioral impairment. Fifteen presented with a small head circumference (OFC $\leq$ 10th percentile), but did not present with neurobehavioral impairment. All 15 were under 6 years of age. Of the 15 infants/toddlers, all were

microcephalic (OFC $\leq$ 3rd percentile), but did not present with developmental delay >1.5 SD below the mean. Nine of the 24 presented with severe CNS dysfunction, but were normocephalic. Of the 22 with confirmed PAE, 7 had levels that were reportedly too low to meet the Hoyme alcohol exposure criteria.

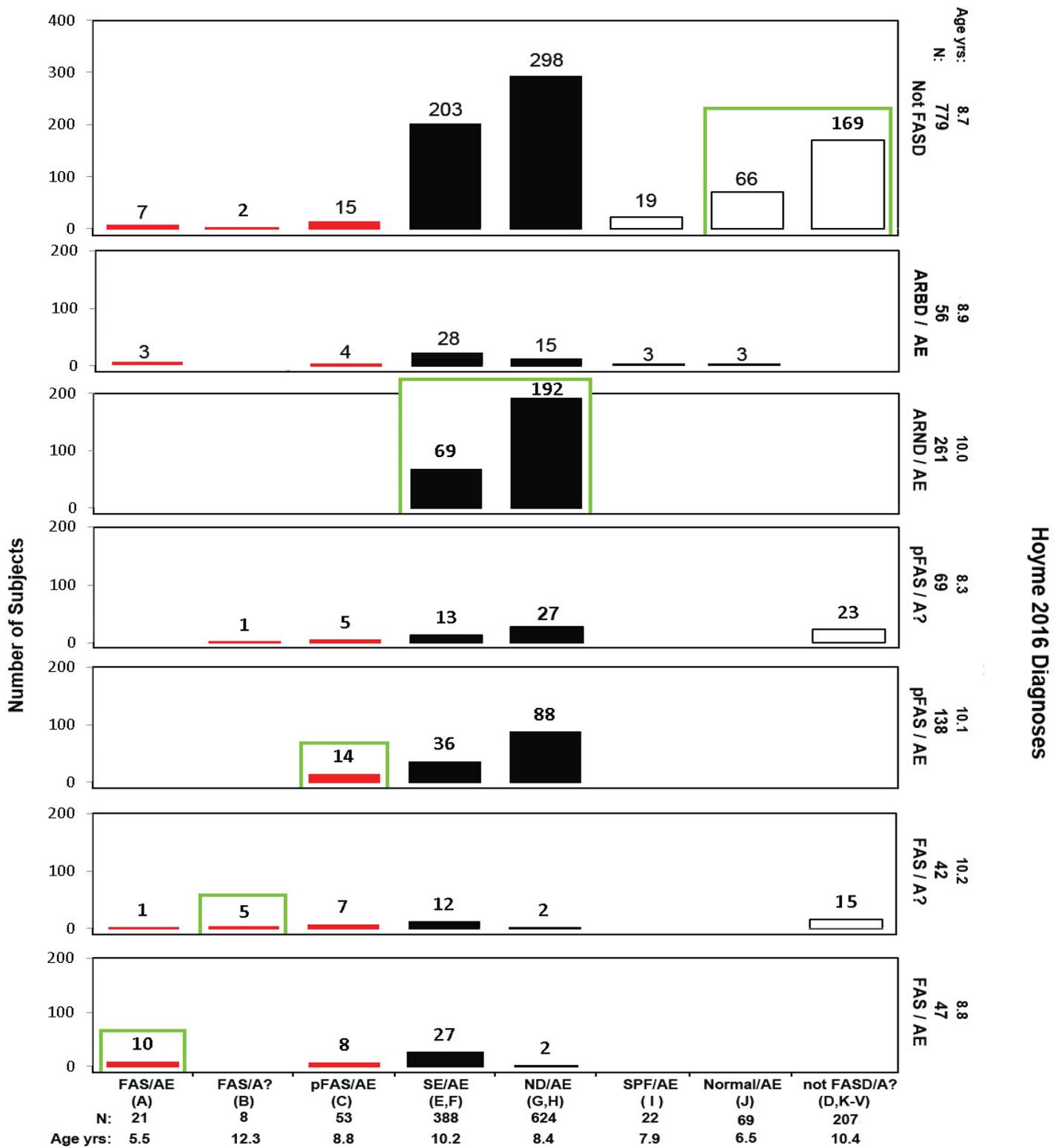
4. Among 82 patients diagnosed FAS/PFAS by the 4-Digit Code, 21 were classified as “Not FASD” by the Canadian system (Figure 8). Of the 4 with FAS/Alcohol unknown, all were > 6 years of age with microcephaly, but 2 with severe CNS dysfunction did not meet the Canadian FAS face criteria (the PFLs were -1.7 SDs on the Clarren PFL charts [14] used by the Canadian system, compared to -2.5 SDs on the Stromland PFL charts [28] used by the 4-Digit Code). The other 2 patients met the Canadian facial criteria, but did not meet the severe CNS criteria, despite their microcephaly. Of the 17 with PFAS/AE, all had Rank 3 facial phenotypes (classified as “normal” by the Canadian system) and 14 had Rank 3 alcohol exposure (not meeting the high PAE required by the Canadian system). The remaining 3 had high PAE and microcephaly, but did not meet the Canadian requirement for severe CNS dysfunction.
5. Among 624 patients diagnosed ND/AE by the 4-Digit Code, 220 received an “At Risk” classification, and 404 received a “Not FASD” classification by both the Australian (Figure 9) and Canadian (Figure 8) diagnostic systems. The 404 patients classified “not FASD” by the two systems were all >6 years of age with confirmed PAE (half with confirmed high PAE). Eighty-seven percent had 1 or 2 (but not 3) domains of severe dysfunction and all had multiple domains of moderate dysfunction. The Australian and Canadian systems do not classify patients with moderate dysfunction under the umbrella of FASD. Primate research documents moderate dysfunction is the most prevalent outcome (59%) caused by prenatal alcohol exposure (Figure 6A).
6. Among the 372 patients diagnosed “FASD without the Face” by the Australian system, only 201 (54%) received the same diagnosis from the Canadian system (Figure 10). The remaining 46% (42+129) had confirmed PAE, but did not receive a FASD diagnosis by the Canadian system because they did not meet the Canadian requirement for high PAE.

### Objective 2c: Assess measures of performance (validation)

Validity is the degree to which a tool (or diagnostic system) is measuring what it purports to measure. Space does not permit a comprehensive assessment of performance across all 4 systems. Below are select examples to demonstrate the impact different measurement scales and criteria can have on the clinical and research performance of the diagnostic systems. The Reader is referred to Astley [17] for a comprehensive assessment of validation of the 4-Digit Code.

#### Correlation between the FAS Facial Phenotype and Prenatal Alcohol Exposure

All four systems allow a diagnosis of FAS to be made in the absence of confirmed PAE because the FAS facial phenotype is so highly specific to (caused only by) PAE, the required presence of the face serves as confirmation of PAE. For this practice to be medically valid, the FAS facial phenotype has to be highly specific to PAE. The

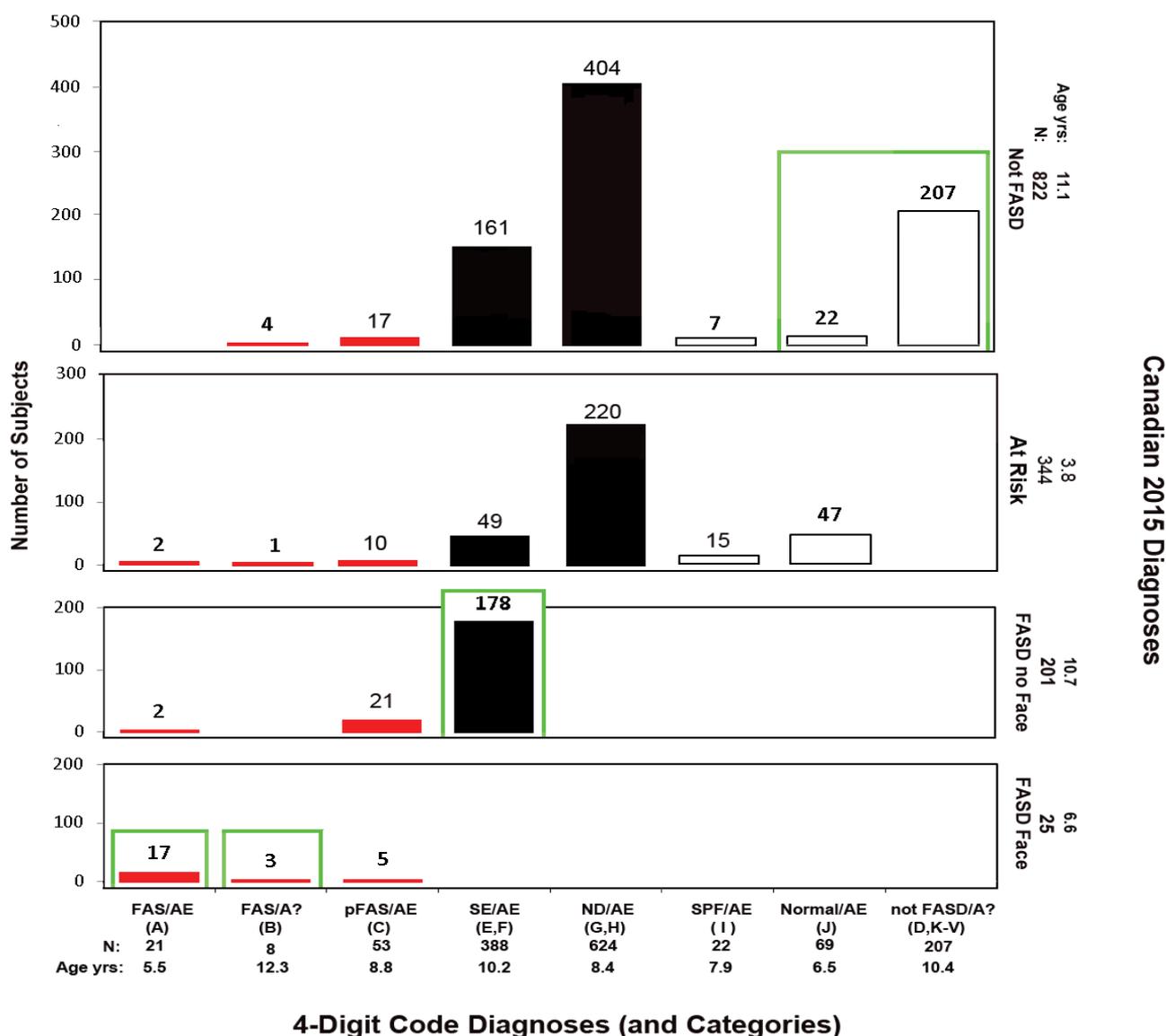


### 4-Digit Code Diagnoses (and Categories)

Figure 7. Cross-tabulation of the 4-Digit Code and Hoyme 2016 FASD Diagnostic Outcomes.

Diagnostic concordance (green boxes) between the 4-Digit Code and Hoyme 2016 systems was observed in 38% (528/1,392) of the patients. Red bars reflect FAS and pFAS diagnoses using the 4-Digit Code. Black bars reflect the rest of the FASD spectrum using the 4-Digit Code. As a demonstration for how to interpret this figure; 21 patients received a 4-Digit Code Diagnosis of FAS/AE. Of the 21 patients, 10 received a FAS/AE diagnosis, 1 received a FAS/A? and 10 did not receive a diagnosis under the umbrella using the Hoyme 2016 diagnostic system.

**Abbreviations:** 4-Digit Code Categories A-V are case-defined in the Diagnostic Guide for FASD [5]. AE: alcohol exposed; A?: alcohol exposure unknown; ND: neurodevelopmental disorder; Not FASD/A?: Individuals who present with or without growth, facial, and/or CNS abnormalities, but are not under the umbrella of FASD because their prenatal alcohol exposure is unknown and they do not meet the criteria for FAS/A?. SE: static encephalopathy; SPF: Sentinel Physical Findings, individuals who present with growth deficiency and/or 1 to 3 FAS facial features, but have normal CNS structure and function; Normal: no evidence of growth, facial, or CNS structural/functional abnormalities. Age yrs; mean age in years at diagnosis.



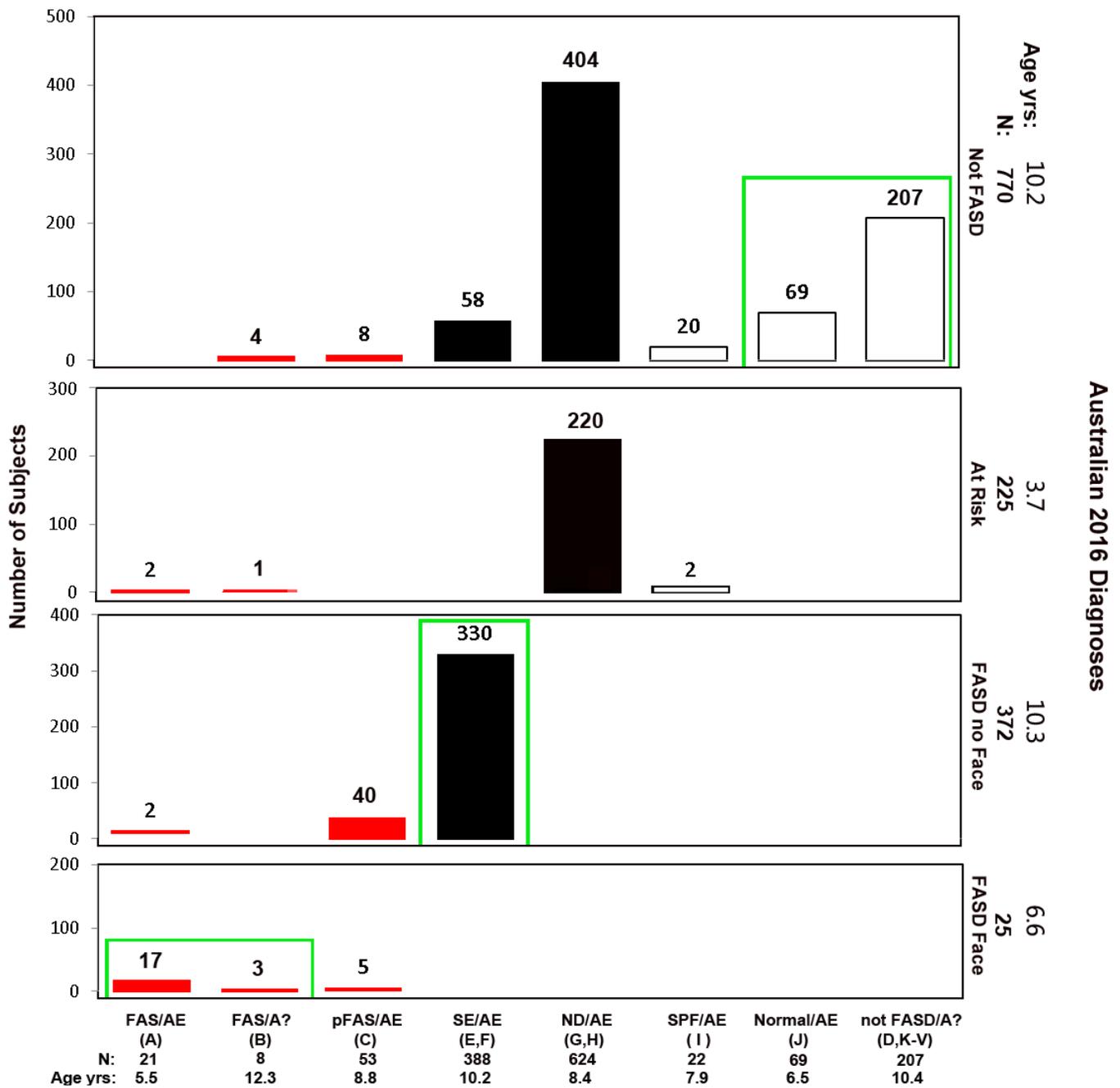
**Figure 8.** Cross-tabulation of the 4-Digit Code and Canadian 2015 FASD Diagnostic Outcomes. Diagnostic concordance (green boxes) between the 4-Digit Code and Canadian 2015 systems was observed in 31% (427/1,392) of the patients. Red bars reflect FAS and pFAS diagnoses using the 4-Digit Code. Black bars reflect the rest of the FASD spectrum using the 4-Digit Code. As a demonstration for how to interpret this figure; 388 patients received a 4-Digit Code Diagnosis of SE/AE (severe CNS abnormalities with confirmed PAE). Of the 388 patients, 178 received a “FASD without the Face” diagnosis, 49 received an “At-Risk” classification and 161 received a “Not FASD” classification using the Canadian diagnostic system. All 49 At-Risk are <6 years with confirmed PAE. Over half have severe dysfunction, but do not meet the high PAE criteria for FASD. The remaining has microcephaly, but do not meet the severe dysfunction criteria for FASD. The 161 classified “Not FASD” have the same profile as those classified “At Risk”, but all are > 6 years of age, thus will not present with both high PAE and severe dysfunction later in childhood as required for a Canadian FASD diagnosis.

**Abbreviations:** 4-Digit Code Categories A-V are case-defined in the Diagnostic Guide for FASD [5]. AE: alcohol exposed; A: alcohol exposure unknown; ND: Neurodevelopmental Disorder; Not FASD/A: Individuals who present with or without growth, facial, and/or CNS abnormalities, but are not under the umbrella of FASD because their prenatal alcohol exposure is unknown and they do not meet the criteria for FAS/A?. SE: static encephalopathy; SPF: Sentinel Physical Findings, individuals who present with growth deficiency and/or 1 to 3 FAS facial features, but have normal CNS structure and function; Normal: no evidence of growth, facial, or CNS structural/functional abnormalities. Age yrs; mean age in years at diagnosis.

Rank 4 FAS facial phenotype, introduced by the 4-Digit Code and adopted by the Canadian and Australian systems, has a specificity of >95% [17,20]. The FAS facial phenotype as defined by the Hoyme system is substantially relaxed relative to the 4-Digit Code Rank 4 facial phenotype (Figure 3) and has a substantially reduced specificity (71% to 75%) [19,32]. If the FAS facial phenotype is specific to PAE, validation studies should confirm the FAS facial phenotype is more prevalent among those with higher exposure and does not occur in individuals with confirmed absence of PAE.

One would also expect that the majority of (if not all) individuals presenting with the FAS facial phenotype would meet criteria for a diagnosis under the umbrella of FASD.

No association was observed between the prevalence of the Hoyme FAS facial phenotype and level of alcohol exposure. The Hoyme FAS facial phenotype was equally prevalent and highly prevalent in the Rank 3 (moderate exposure) and Rank 4 (high exposure) groups when alcohol exposure was classified in accordance with the 4-Digit Code (4-Digit Code Alcohol: Chi<sup>2</sup> 0.95, p=0.33) (Figure

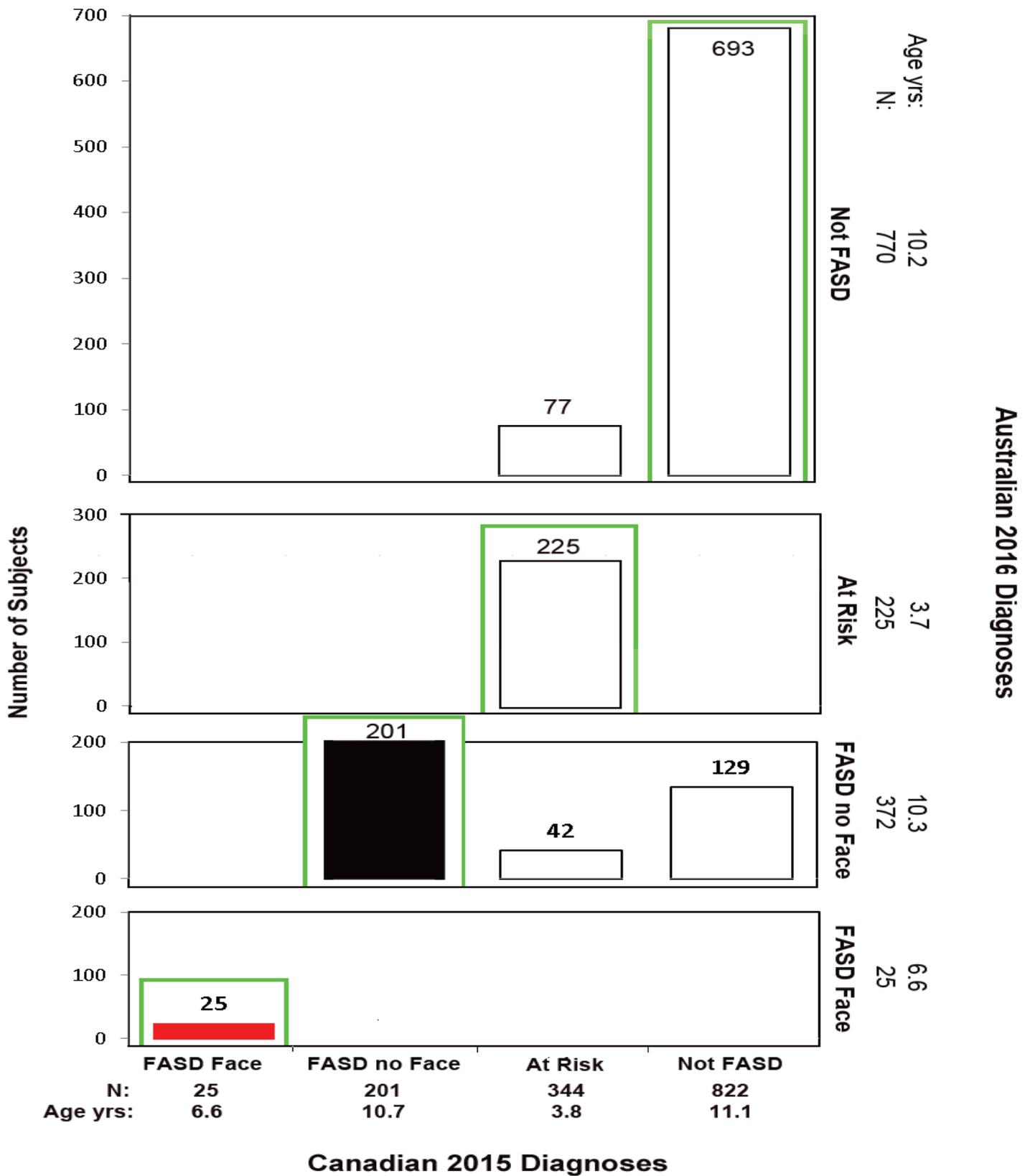


### 4-Digit Code Diagnoses (and Categories)

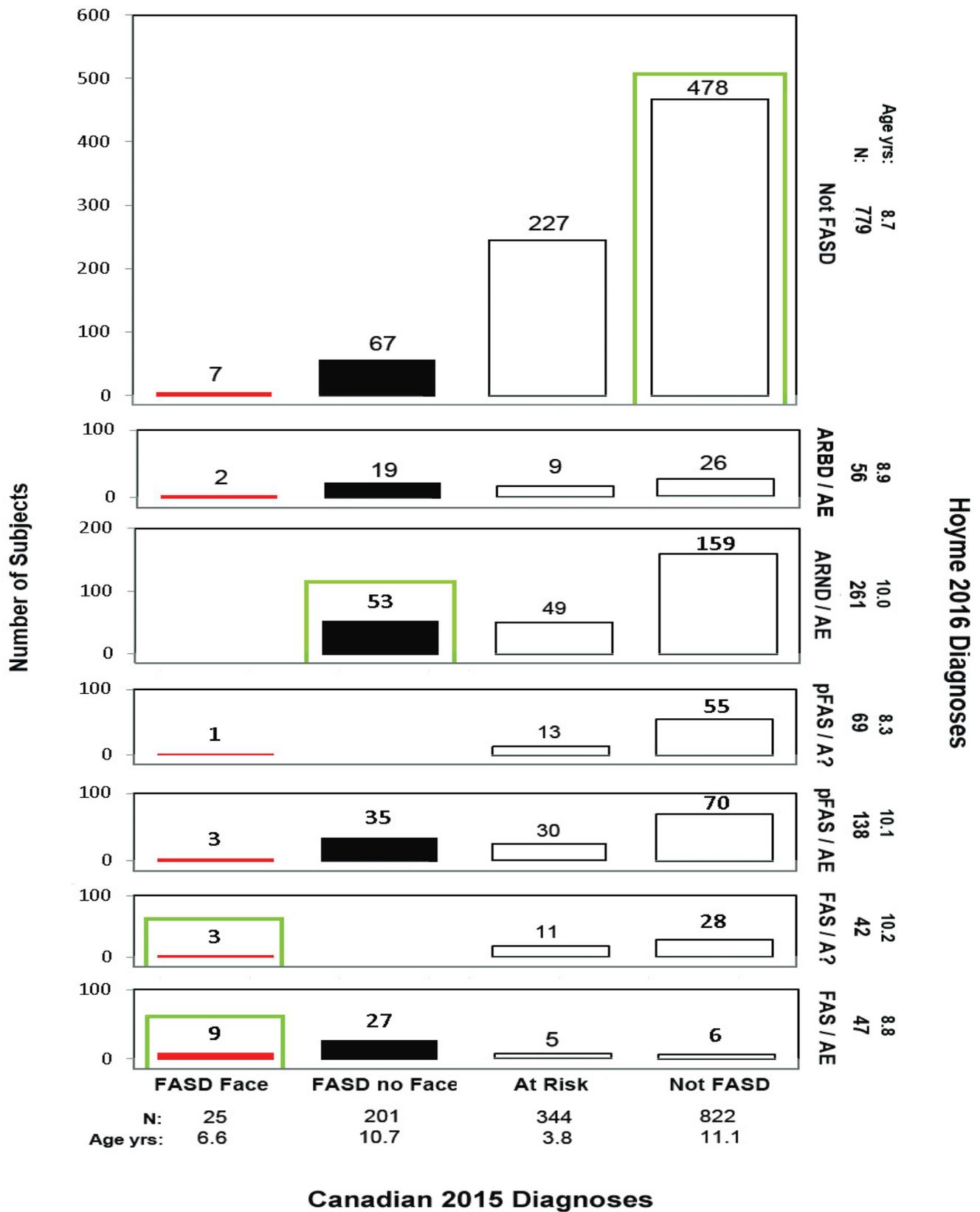
**Figure 9.** Cross-tabulation of the 4-Digit Code and Australian 2016 FASD Diagnostic Outcomes.

Diagnostic concordance (green boxes) between the 4-Digit Code and Australian 2016 systems was observed in 45% (626/1,392) of the patients. Red bars reflect FAS and PFAS diagnoses using the 4-Digit Code. Black bars reflect the rest of the FASD spectrum using the 4-Digit Code. As a demonstration for how to interpret this figure; 624 patients received a 4-Digit Code Diagnosis of ND/AE. Of the 624 patients, 220 received an “At Risk” classification, and 404 received a “Not FASD” classification using the Australian diagnostic system. The 404 patients classified “not FASD” by the Australian system were all > 6 years of age with confirmed PAE (half with confirmed high PAE). 87% had 1 or 2 (but not 3) domains of severe dysfunction and all had multiple domains of moderate dysfunction. The Australian system does not classify patients with this level dysfunction under the umbrella of FASD. Primate research documents moderate dysfunction is the most prevalent outcome caused by prenatal alcohol exposure (Figure 6). The 220 patients classified as “At Risk” by the Australian system have the same exposure and moderate dysfunction profile, but are all <6 years of age. These 220 are identified as “At Risk” because they are at risk of presenting with severe dysfunction later in childhood, and thus still at risk for FASD.

**Abbreviations:** 4-Digit Code Categories A-V are case-defined in the Diagnostic Guide for FASD [5]. AE: alcohol exposed; A?: alcohol exposure unknown; ND: neurodevelopmental disorder; Not FASD/A?: Individuals who present with or without growth, facial, and/or CNS abnormalities, but are not under the umbrella of FASD because their prenatal alcohol exposure is unknown and they do not meet the criteria for FAS/A?. SE: static encephalopathy; SPF: Sentinel Physical Findings, individuals who present with growth deficiency and/or 1 to 3 FAS facial features, but have normal CNS structure and function; Normal: no evidence of growth, facial, or CNS structural/functional abnormalities. Age yrs; mean age in years at diagnosis.



**Figure 10.** Cross-tabulation of the Canadian 2015 and Australian 2016 FASD Diagnostic Outcomes. Diagnostic concordance (green boxes) between the Canadian 2015 and Australian 2016 systems was observed in 82% (1,144/1,392) of the patients with the majority of the concordance due to 693 of the patients receiving a “Not FASD” diagnosis by both systems. This higher level of concordance is due to the fact that the Australian system adopted most of the criteria used by the Canadian system, with one important exception. The Canadian system requires confirmed high PAE. The Australian system requires confirmed PAE at any level. Red bars reflect “FASD with and without the Face” diagnoses using the Canadian system. As a demonstration of how to interpret this figure; 822 patients received a Canadian classification of “Not FASD”. Of the 822 diagnosed “FASD without the Face” and 693 received a “Not FASD” classification using the Australian diagnostic system. The 129 diagnosed “FASD with no Face” by the Australian system all had confirmed PAE, but the level did not meet the Canadian requirement for high exposure. Red bars reflect “FAS with the Face” diagnoses using the Canadian system. Black bars reflect “FASD without the Face” diagnoses using the Canadian system. Abbreviations: Age yrs; mean age in years at diagnosis.



**Figure 11.** Cross-tabulation of the Canadian 2015 and Hoyme 2016 FASD Diagnostic Outcomes. Diagnostic concordance (green boxes) between the Canadian 2015 and Hoyme 2016 systems was observed in 41% (569/1,392) of the patients with the majority of the concordance due to 740 of the patients receiving a “Not FASD” diagnosis by both systems. Red bars reflect “FAS with the Face” diagnoses using the Canadian system. Black bars reflect “FASD without the Face” diagnoses using the Canadian system. As a demonstration for how to interpret this figure; 822 patients received a Canadian classification of “Not FASD”. Of the 822 patients, 28 received a diagnosis of “FAS/A?”, 70 received a diagnosis of pFAS/AE, 55 received a diagnosis of pFAS/A?, 159 received a diagnosis of ARND/AE and 504 received a classification of “Not FASD” using the Hoyme system. Most of the 159 with FAS/PFAS presented with the relaxed Hoyme FAS facial phenotype. Only 8 of the 159 presented with the Canadian FAS face (4-Digit Code Rank 4). The remaining 151 patients with the Hoyme FAS face presented with the following 4-Digit Face Ranks: Rank 1 normal face 15%, Rank 2 mild face 59%, and Rank 3 moderate face 21%. These relaxed FAS facial phenotypes were used by the Hoyme system to overcome the unknown PAE among the 70 patients diagnosed pFAS/A? and the 28 patients diagnosed FAS/A? Abbreviations: Age yrs; mean age in years at diagnosis.

12A). The Hoyme FAS facial phenotype was also equally prevalent and highly prevalent when alcohol exposure was classified in accordance with the Hoyme system ( $\text{Chi}^2$  0.01;  $p=0.92$ ) (Figure 12B). In contrast, the 4-Digit Code Rank 4 FAS face was 5 times more prevalent in the Rank 4 high exposure group than the Rank 3 moderate exposure group ( $\text{Chi}^2$  17.5;  $p=0.000$ ) (Figure 12C). The association between the 4-Digit FAS face and alcohol was weakened substantially when the Hoyme criteria for alcohol exposure were used ( $\text{Chi}^2$  6.1,  $p=0.02$ ). The 4-Digit FAS face was only 2-fold more prevalent in the Hoyme exposed group relative to the Hoyme unknown/too-low exposure group (Figure 12D).

Of the 552 patients with the Hoyme FAS face, 43% did not receive a diagnosis under the umbrella of FASD using the Hoyme system. In contrast, all 54 individuals with the 4-Digit Code Rank 4 FAS face met criteria for a diagnosis under the umbrella of FASD using the 4-Digit Code.

When the Hoyme and 4-Digit Code FAS facial criteria were applied to an adolescent with high function (FSIQ 123) and confirmed absence of PAE (4-Digit Code 1211), she met the Hoyme criteria for the full FAS facial phenotype (Figure 3C). In contrast, her facial phenotype was classified within the normal range by the 4-Digit Code (Face ABC-Score BBC, Face Rank 2).

#### **Should moderate dysfunction be included under the umbrella of FASD? Does PAE cause moderate dysfunction?**

All four diagnostic systems include a diagnosis under the umbrella of FASD for individuals that present with severe dysfunction (3 or more domains of function, 2 or more SDs below the mean). Only the 4-Digit Code and Hoyme systems, however, include diagnostic classifications (ND/AE and ARND respectively) for individuals who present with moderate dysfunction (1 or 2 domains of function 2 or more SDs below the mean). Should moderate dysfunction be included under the umbrella of FASD? Does PAE cause moderate dysfunction? To address this question, the 4-Digit Code was applied to our nonhuman-primate model of FASD [4] to document the distribution of diagnostic (FAS/PFAS, SE/AE, ND/AE and Not FASD/AE) outcomes when the only risk factor present was PAE. The primates had been exposed weekly to binge exposures equivalent to a six-pack of beer for the first 3, 6 or entire 24 weeks of gestation (mean maternal peak plasma ethanol concentrations ranged from 176 to 271 mg/dl). The primate model confirmed PAE causes a spectrum of outcome (FAS/PFAS 5%, SE/AE 31%, ND/AE 59%, and Not FASD/AE 5%) with moderate dysfunction (ND/AE) being the most prevalent outcome (Figure 6A). The 4-Digit Code was the only system that produced a near identical distribution of diagnoses across the full spectrum (including 53% ND/AE) illustrated in Figure 6B. The Australian and Canadian outcomes were in greatest contrast with the primate model due to their exclusion of moderate dysfunction from the spectrum. The Australian system produced a good match to the primate model for the severe end of the spectrum (FASD with and without the Face), whereas the Canadian system's requirement for confirmed high PAE results in a poor match between their diagnostic outcomes and the primate model. The Hoyme criteria produce outcomes across the full spectrum, but the distribution did not match the primate model. The relaxed facial criteria placed far more in the FAS/PFAS category and far less in the moderate

and severe dysfunction categories. The Australian, Canadian and Hoyme systems placed 51% to 81% of patients with PAE in the "Not FASD" category, in contrast to the 5% observed in the primate model.

#### **Does the pattern and magnitude of dysfunction among patients with moderate dysfunction warrant and qualify them for intervention services?**

Of the 402 patients with ND/AE who were 6 years of age or older at the time of their diagnosis, 83% presented with 1-2 domains of severe dysfunction (2 or more SDs below the mean) and 1-6 domains of moderate dysfunction (1 to 1.9 SDs below the mean) (Figure 13). The patterns of moderate dysfunction (1 to 1.9 SDs below the mean) across 9 domains of function (intellect, adaptation, achievement, memory-executive function, language, motor, mental health, behavior and development) is comparable between patients diagnosed with ND/AE, SE/AE and FAS/PFAS using the 4-Digit Code (Figure 14A). The patterns of severe dysfunction (2 or more SDs below the mean) across 9 domains of function is less prevalent among patients with ND/AE than SE/AE and FAS/PFAS (by definition), but present nonetheless (Figure 14B). The magnitude and breadth of dysfunction observed among patients with ND/AE warrant identification and intervention.

#### **The outcomes associated with FASD present along clinically meaningful continuums. Collapsing these continuums to dichotomous (present, absent) scales can hinder clinical practice and research efforts**

The following serves as just one of many examples of how collapsing a continuous outcome into a dichotomous (present, absent) scale can adversely impact clinical practice (e.g., the ability to render an accurate diagnosis and predict those at greatest risk) and research efforts (e.g., the power to detect causal associations). The FAS facial phenotype serves two clinically vital functions in the field of FASD. 1) The Rank 4 FAS facial phenotype is so highly specific to PAE it can be used to confirm PAE when a history of PAE is not available for a patient [20,17]. 2) The phenotype presents along a clinically informative continuum that is highly correlated with (and predictive of) the magnitude of CNS damage in a young patient [9,17]. Linear correlations serve as one of the most powerful metrics for identifying causal associations. Identification of causal and predictive associations serves to validate and inform clinical practice. For example, the causal link between the Rank 4 face and alcohol allows the clinician to render a diagnosis of FAS when PAE is unknown. The ability of the Rank 3 face to predict severe CNS dysfunction later in childhood allows the clinician to identify and provide early intervention to infants/toddlers at high risk. When the continuum of expression of the FAS facial phenotype is collapsed into a dichotomous (present, absent) scale, the clinical utility of the phenotype is diminished or even invalidated. For example, as illustrated in Figure 15, a significant linear correlation between the magnitude of expression of the FAS facial phenotype and the prevalence of severe CNS dysfunction (CNS Rank 3) is identified when the facial phenotype is recorded on the 4-point ordinal scale used by the 4-Digit Code ( $\text{Chi}^2$  linear trend=10.5,  $p=0.001$ ) (Figure 15). It is clear from the pattern of association depicted by the blue line (Figure 15) that the prevalence of CNS dysfunction associated with the Rank 1 and 2 facial phenotypes (32% and 36%) are distinctly lower than the prevalence of dysfunction associated with the Rank 3 and 4 facial phenotypes

(50% and 50%). The 4-point scale preserves the clinician's ability to use the Rank 3 and Rank 4 faces to predict which toddlers are at highest risk of severe CNS dysfunction [9]. The 4-point scale also preserves the clinician's ability to use the Rank 4 FAS facial phenotype to serve as confirmation of PAE when a history of PAE is not available. Both these functions are vital in a clinical setting. When the magnitude of expression of the FAS facial phenotype is collapsed into just two categories (present, absent), as introduced by the Hoyme, Canadian and Australian systems, one or both of these clinical functions are lost. For example, the Canadian and Australian systems collapsed the FAS face into (Present: Rank 4; Absent: Ranks 1, 2 and 3). In so doing, the systems preserved the high specificity of the Rank 4 face and thus the important clinical ability to use the Rank 4 face as confirmation of PAE when a history of PAE is unavailable. But, by collapsing the Rank 3 face with the Rank 1 and 2 faces, the Australian and Canadian systems lost the clinical ability to predict which infants/toddlers (those with the Rank 3 face) will present with severe brain dysfunction later in childhood [9]. The significant linear correlation between the FAS facial phenotype and prevalence of severe dysfunction detected by the 4-Digit Code ordinal scale (blue line in Figure 15) was rendered insignificant by the Australian/Canadian dichotomous facial scale (red line in Figure 14,  $\text{Chi}^2=0.5$ ,  $p=0.46$ ). The Hoyme system also collapsed the FAS face into a dichotomous (present, absent) scale, but used a different cut-point along the 4-Digit Code 4-point Face Rank scale (Present: Ranks 2,3, 4 and half of Rank 1; Absent: the other half of Rank 1) (Figure 3B). By combining Face Ranks 2, 3 and half of Rank 1 with Face Rank 4, the Hoyme system lost the high specificity of the Rank 4 FAS face and thus lost the clinical ability to use the "FAS face" as confirmation of PAE when a history of PAE is unavailable in a patient.

The Hoyme system also lost the clinical ability to predict which infants/toddlers will present with severe brain dysfunction later in childhood because the predictive ability of the Rank 3 and 4 faces are weakened by combining them with the normal Rank 1 and 2 facial phenotypes. The significant linear correlation between the FAS facial phenotype and prevalence of severe dysfunction detected by the 4-Digit Code ordinal scale (blue line in Figure 15) was rendered insignificant ( $\text{Chi}^2=4.7$ ,  $p=0.10$ ) by the Hoyme dichotomous facial scale (black line in Figure 15).

## Discussion

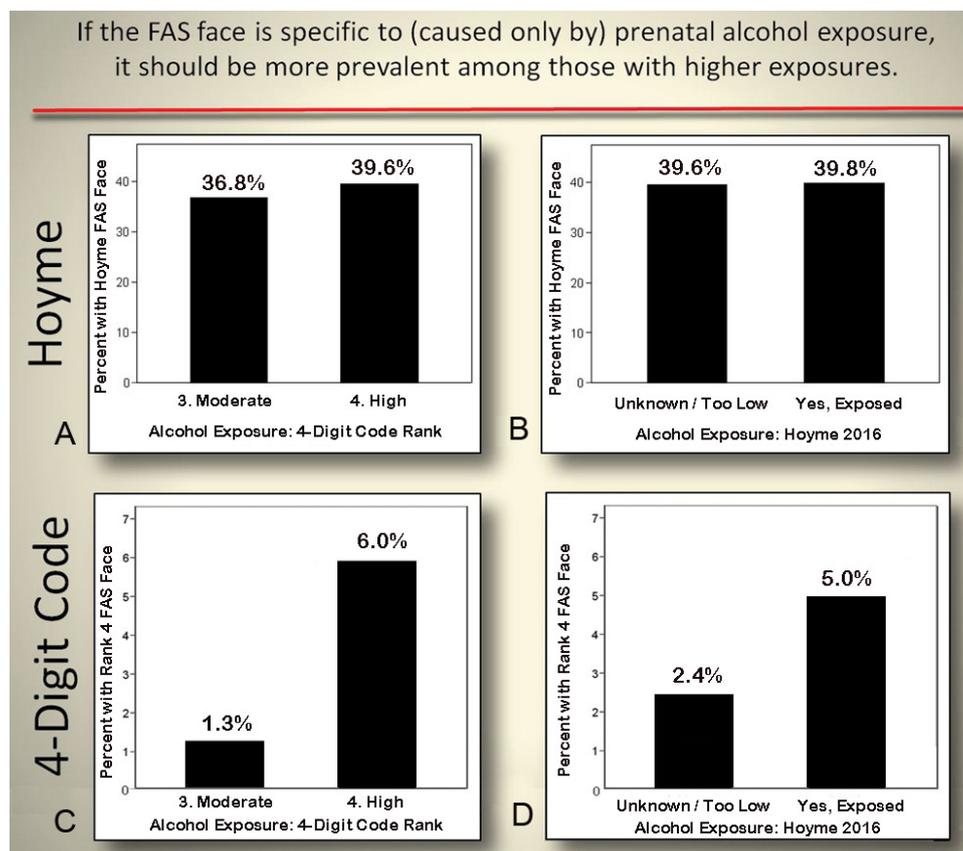
### Contrasts in diagnostic outcomes

The four systems produced markedly different outcomes. Eighty-two percent of patients were diagnosed with FASD by at least one of the four systems, but only 11% of patients were diagnosed with FASD by all four systems. Eight percent of patients were diagnosed with FAS by at least one of the 4 systems, but only 1% was diagnosed with FAS by all four systems. The proportion of patients diagnosed with FAS, severe dysfunction, moderate dysfunction, and FASD overall varied significantly across the systems (4-Digit: 2%, 28%, 45%,  $\leq 79\%$ ; Hoyme: 6%, 5%, 14% 44%; Australian: 2%, 26%, 0%, 29%; and Canadian: 2%, 14%, 0%, 16%) (Table 4).

Five factors accounted for the greatest contrasts in diagnostic outcomes between the four systems.

1. Extensive evidence supports the inclusion of individuals with moderate dysfunction (ND/AE) under the umbrella of FASD. The pattern and magnitude of dysfunction

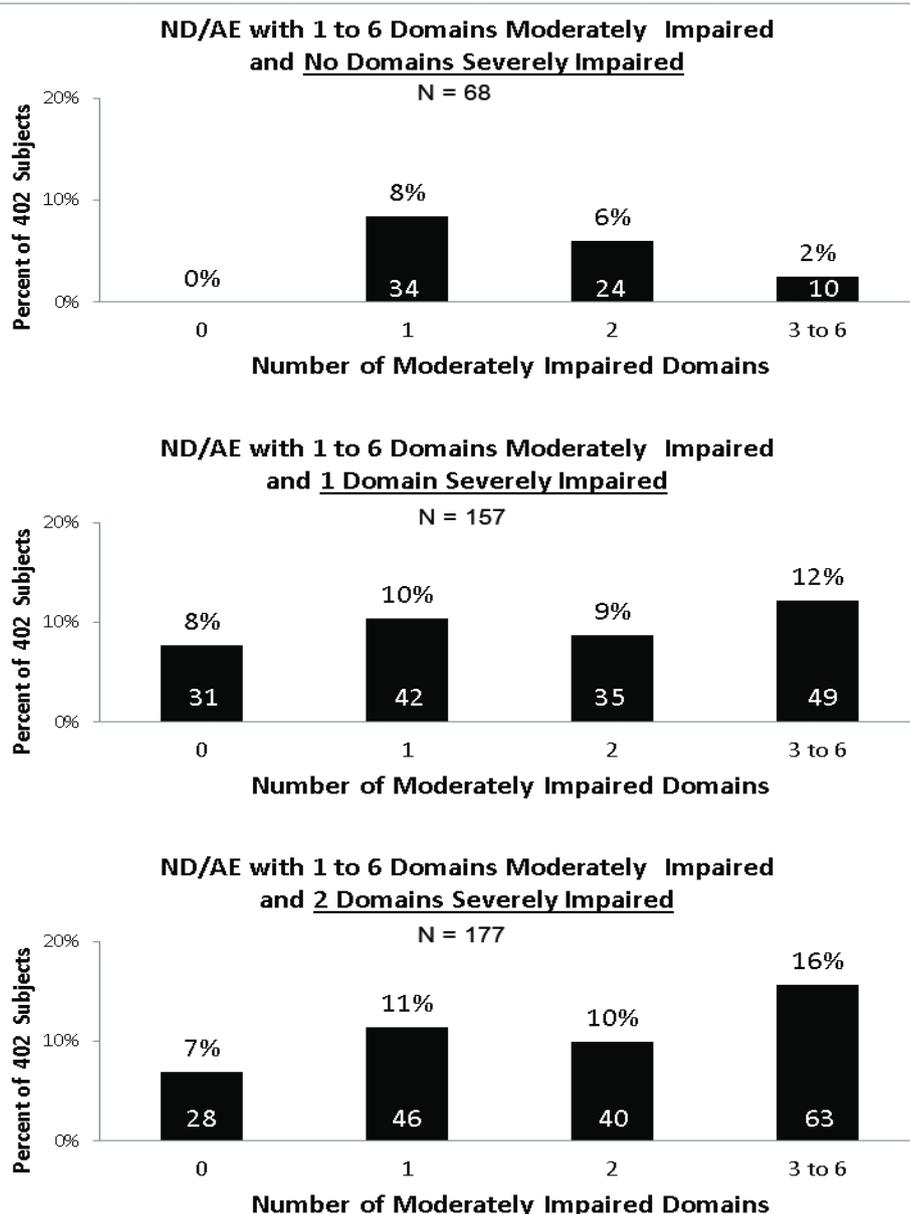
among patients with moderate dysfunction warrant and qualify them for intervention services. Exclusion of moderate dysfunction by the Canadian and Australian systems prevented 53% of patients with confirmed PAE from receiving a FASD diagnosis with the greatest impact on children less than 6 years of age. Individuals with PAE present with the full spectrum of CNS dysfunction from moderate to severe in Table 3 [1,33,34]. The evidence that supports inclusion of moderate dysfunction (ND/AE or moderate ARND) under the umbrella of FASD is as follows. First, and most importantly, hundreds of laboratory-based studies, including our nonhuman-primate studies in Figure 6A [4,35], confirm prenatal alcohol exposure causes moderate dysfunction. Not only does it cause moderate dysfunction, but moderate dysfunction is the most common outcome. In this study population of 1,177 with PAE and the larger population from which it was drawn (2,550 alcohol-exposed patients evaluated at the WA FASDPN clinics over the past 20 years), 45-53% met the criteria for ND/AE [17]. ND/AE was the most common outcome, exceeding the prevalence of FAS/PFAS (6-10%) and SE/AE (24-33%) combined. It is important to note that alcohol is not the only risk factor contributing to adverse outcomes in the FASDPN patient population (see Figure 21 in Astley [17]). So what would the diagnostic distribution look like if alcohol was the only risk factor? To answer that question, we applied the 4-Digit Code to the outcomes observed in our primate model of FASD [4] (Figure 6A). Remarkably, the distribution of FAS/PFAS (5%), SD/AE (31%) and ND/AE (59%) was near identical to that observed in our FASD clinical population, with ND/AE being the most common outcome. And just like in our primate model, individuals with ND/AE have alcohol exposures as high as those with FAS/PFAS and SE/AE (see Figure 22 in Astley [17]). Are these moderate impairments in brain function associated with underlying CNS structural abnormalities? Again, the answer is yes. Our MRI study confirmed at least 43% of individuals with ND/AE have significant CNS structural abnormalities [36] (see also Figure 15C in Astley [17]). Our extensive experience in the WA FASDPN confirms that it is the children with moderate dysfunction that fair the worst and are often in most need of diagnostic identification and intervention. These are the children that too often slip through the cracks. Their disabilities are often not severe enough in the cognitive domain to qualify them for services (only 3% have an IQ less than 70) [17], but severe enough across many other domains (Figures 13 and 14) (see also Figure 23 in Astley [17]) to adversely impact their ability to fully engage in school and live productive, independent lives. Children with ND/AE received as many intervention recommendations as children with FAS/PFAS and SE/AE in our patient population (see Figure 24 in Astley [17] and Table 4 in Jirikowic et al [37]). And perhaps most importantly, the diagnosis of ND/AE provided caregivers with as much access to services as caregivers of children with FAS/PFAS and SE/AE. Caregivers also reported that the interventions worked as well for their children with ND/AE as did caregivers of children with FAS/PFAS and SE/AE (see Figure 31 in Astley [17]).



**Figure 12.** Only the 4-Digit Code FAS Face was Significantly More Prevalent Among Patients with Higher Alcohol Exposure. The Hoyme 2016 [10] FAS face was equally prevalent and highly prevalent in the moderate (4-Digit Code Alcohol Rank 3) and high (4-Digit Code Alcohol Rank 4) alcohol exposure groups ( $\text{Chi}^2$  0.9,  $p=0.33$ ). B) The Hoyme FAS face was also equally prevalent and highly prevalent between those that did and did not meet the Hoyme alcohol exposure criteria ( $\text{Chi}^2$  0.01,  $p=0.92$ ). In contrast, the 4-Digit Code FAS facial phenotype was highly correlated with measures of prenatal alcohol exposure. C) The 4-Digit Code Rank 4 FAS face was 5 times more prevalent in the high exposure group (4-Digit Code Alcohol Rank 4) than the moderate exposure (Digit Code Alcohol Rank 3) group  $\text{Chi}^2$  17.5,  $p=.000$ . D) The association between the 4-Digit Code Rank 4 FAS facial phenotype and alcohol was substantially weakened when the Hoyme 2016 criteria for alcohol exposure were applied ( $\text{Chi}^2$  6.1,  $p=0.02$ ). The 4-Digit FAS face was only 2-fold more prevalent in the Hoyme et al. exposed group relative to the Hoyme et al. unknown/too low exposure group.

It is important to clarify that, when we report above that there is extensive evidence to support inclusion of ND/AE under the umbrella of FASD, we are not stating that all individuals who meet the criteria for ND/AE have FASD. By definition all individuals with Fetal Alcohol Spectrum Disorder have a disorder caused, at least in part, by their prenatal alcohol exposure. But not all individuals with ND/AE necessarily have a FASD. Only the subset of individuals whose neurobehavioral disorder was caused, at least in part, by their prenatal alcohol exposure, have a FASD. This is a current inherent weakness in the umbrella term FASD. In the absence of a biomarker that can causally link an individual's alcohol exposure with their neurodevelopmental disorder, there is no way to identify which individuals with ND/AE have FASD. This same argument applies to the diagnostic classifications of SE/AE, ARND and "FASD without the Face". Not all individuals who meet the criteria for SE/AE, ARND and "FASD without the Face" necessarily have FASD. Only the subset of individuals whose CNS abnormalities were caused, at least in part, by their prenatal alcohol exposure has FASD. And once again the field of FASD currently has no way (no biomarker) to identify this subset. Until such a biomarker is identified, if such a biomarker exists, the 4-Digit Code elects to label these categories with terms that do not imply causality.

2. The more stringent Hoyme and Canadian alcohol exposure criteria prevented 47%-59% of patients with confirmed PAE from receiving a diagnosis of FASD. In a clinical setting, one is not in a position to know how accurate the exposure was recalled and reported. Setting a threshold implies the details of all reported exposures are accurate and no fetus can be harmed by exposures below the threshold. Neither of these statements is true and the latter sends a confusing public health message that lower levels are safe. Recognizing this, the 4-Digit Code requires a confirmed exposure, but does not set a threshold. It is interesting to note that Petryk et al., [38] reported similar findings when they retrospectively assessed the impact of applying the 2016 Canadian guidelines to 119 patients with confirmed PAE (4-Digit Code Alcohol Ranks 3 or 4) and severe structural and/or functional CNS abnormalities (4-Digit Code CNS Ranks 3 and/or 4). In the Petryk study, the more stringent Canadian exposure criteria would have prevented 71% of the individuals from receiving a diagnosis under the umbrella of FASD because the reported exposure would not have met the required threshold.
3. Individuals with FASD are born with FASD, but the Hoyme, Canadian and Australian guidelines prevent most children under 3 or 6 years of age with confirmed PAE and structural or functional CNS abnormalities from receiving a diagnosis under the umbrella of FASD.

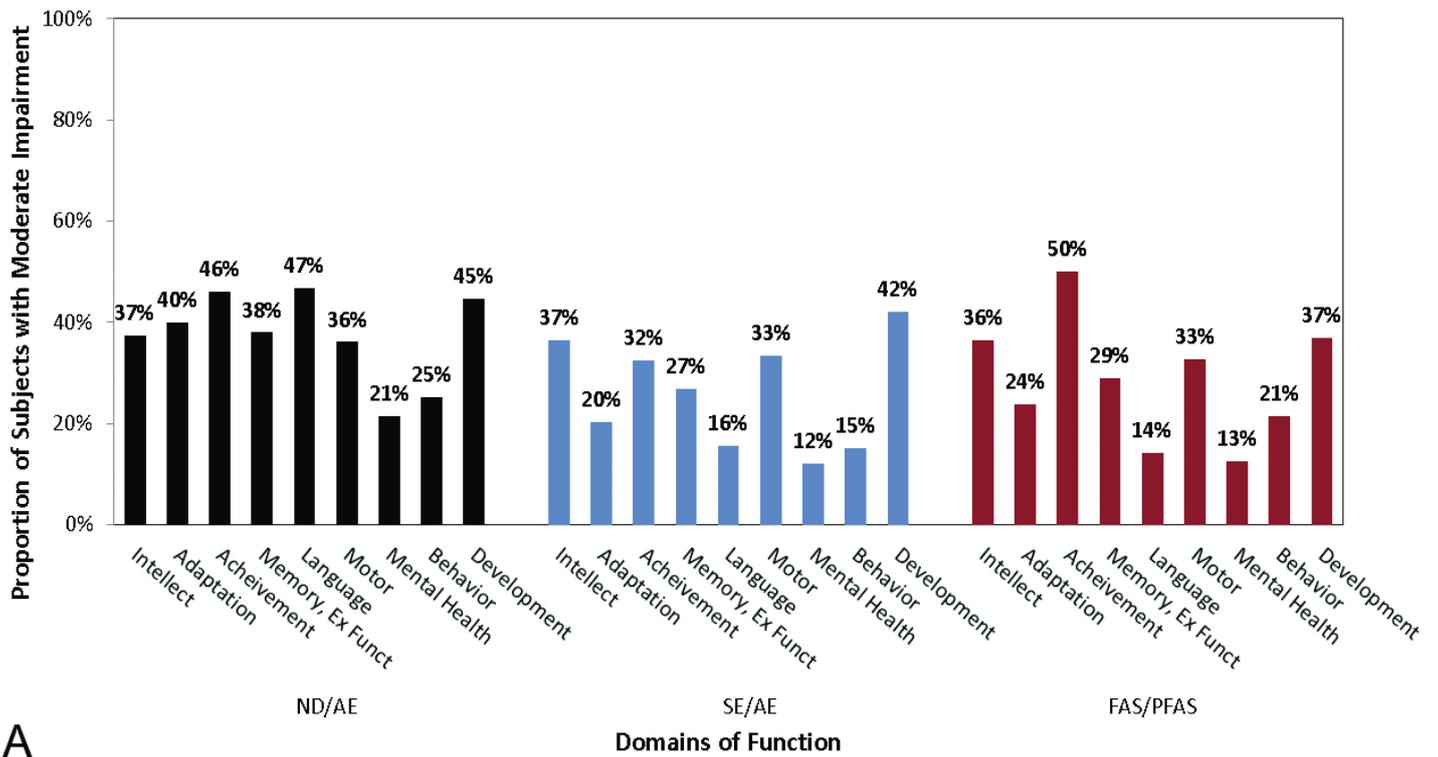


**Figure 13.** Distribution of Moderate and Severe CNS Dysfunction Among Patients Diagnosed with ND/AE by the 4-Digit Code. Of the 402 patients diagnosed with ND/AE by the 4-Digit Code who were 6 years of age or older at the time of their diagnosis, 83% (334/402) presented with 1-2 domains of severe dysfunction (2 or more SDs below the mean) and 1-6 domains of moderate dysfunction (1 to 1.9 SDs below the mean). Domains of function included: intellect, adaptation, achievement, memory-executive function, language, motor, mental health, behavior and development, as illustrated in Figure 14. As a demonstration for how to interpret this figure; the bar on the bottom right documents 63 of the 402 patients (16%) presented with 2 domains of severe dysfunction and 3 to 6 additional domains with moderate dysfunction.

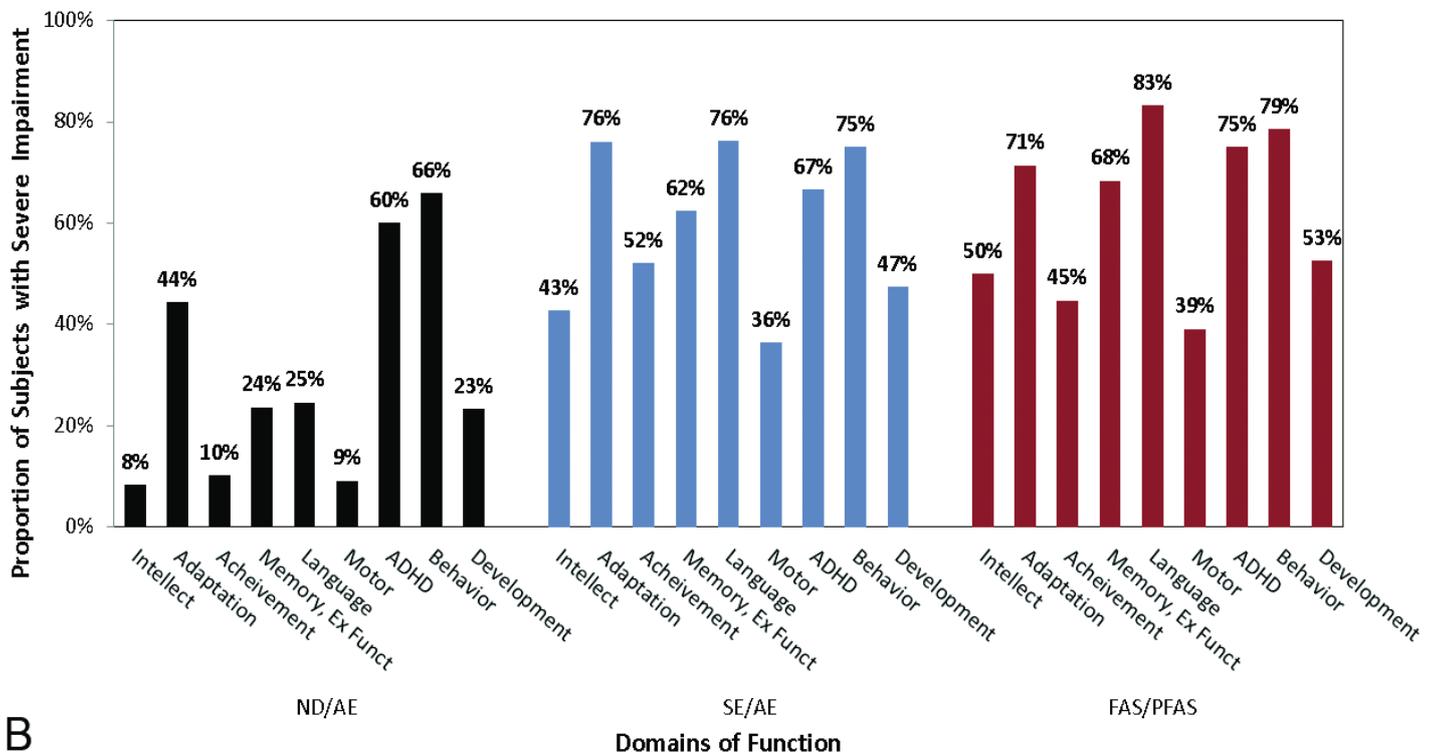
The 4-Digit Code allows a diagnosis of FAS/PFAS at birth based solely on physical abnormalities (growth deficiency, FAS face and microcephaly), having confirmed empirically that over 90% of alcohol-exposed infants and toddlers who present with one or more of the sentinel physical features of FAS as defined by the 4-Digit Code (microcephaly  $\leq$ 3rd percentile, a Rank 4 FAS facial phenotype, or Rank 4 growth deficiency) will present with severe CNS Rank 3 dysfunction later in childhood [9]. In contrast, the Hoyme system requires both reduced head circumference and CNS dysfunction for an FAS/PFAS diagnosis, preventing a diagnosis in infant/toddlers too young to be assessed for CNS dysfunction. In addition, the Hoyme system does not permit a diagnosis of ARND in a child <3 years of age. The Canadian and Australian systems require severe CNS

dysfunction for an FASD diagnosis, preventing all children with PAE who present with microcephaly and/or moderate CNS dysfunction from receiving a diagnosis of FASD, with one exception. If the child with PAE presents with microcephaly and the FAS facial phenotype, a diagnosis of "FASD with the Face" can be made in the absence of CNS dysfunction, based on the finding of the 4-Digit Code that microcephaly and the Rank 4 FAS facial phenotype are highly predictive of severe CNS dysfunction later in childhood. Growth deficiency was as strong a predictor of severe brain dysfunction in infants with PAE as the FAS facial phenotype and microcephaly, but the Canadian and Australian systems excluded growth deficiency from their FASD criteria. While this one exception (microcephaly and the FAS facial phenotype) allowed a diagnosis of "FASD with the Face" in a small number of children (n=6) <6 years

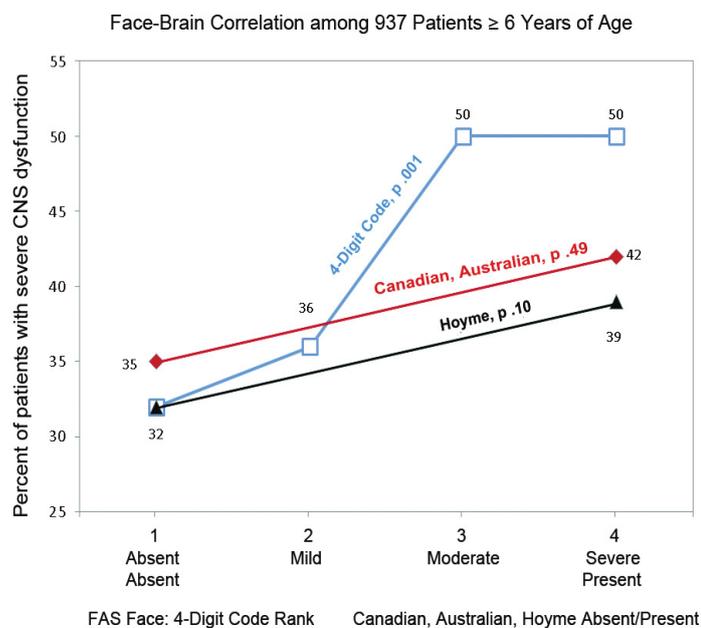
**Pattern of moderate impairment (1 to 1.9 SDs below the mean) among subjects  $\geq$  6 yrs old**



**Pattern of severe impairment (2 or more SDs below the mean) among subjects  $\geq$  6 yrs old**



**Figure 14.** Patterns of Dysfunction Among Patients > 6-years-old with 4-Digit Code Diagnoses ND/AE, SE/AE or FAS/PFAS. The proportion of patients presenting with moderate dysfunction (1 to 1.9 SDs below the mean) across 9 domains of function (intellect, adaptation, achievement, memory-executive function, language, motor, mental health, behavior and development) is comparable between patients diagnosed with ND/AE, SE/AE and FAS/PFAS using the 4-Digit Code. B) The proportion of patients presenting with severe dysfunction (2 or more SDs below the mean) across 9 domains of function is less prevalent among patients with ND/AE than SE/AE and FAS/PFAS (by definition), but present nonetheless.



**Figure 15.** Significant Correlation Between Face and Brain Lost when the Facial Phenotype is Reduced to Present/Absent.

The FAS facial phenotype presents along a clinically meaningful continuum. A significant correlation between the magnitude of expression of the FAS facial phenotype and the prevalence of severe CNS dysfunction (CNS Rank 3) is identified when the facial phenotype is recorded on the 4-point ordinal scale used by the 4-Digit Code (blue line:  $\chi^2$  linear trend=10.5,  $p=0.001$ ). Linear trends serve as one of the most powerful metrics for identifying causal associations. When the magnitude of expression of the FAS facial phenotype is collapsed into just two categories (present, absent), as introduced by the Hoyme, Canadian and Australian systems, the significant correlation between face and brain is lost. Not only does a dichotomous scale have less statistical power to identify real associations, but where the ordinal scale is bisected impacts the validity of the dichotomous scale. It is clear from the pattern of association depicted by the blue line that the prevalence of CNS dysfunction associated with the Rank 1 and 2 faces (32% and 36%) are distinct from the prevalence of dysfunction associated with the Rank 3 and 4 faces (50% and 50%). The most clinically valid cut-point to bisect the ordinal scale would be between Ranks 2 and 3. The Canadian and Australian systems used a cut-point between Ranks 3 and 4 to dichotomize the FAS facial phenotype (Present=Rank 4; Absent=Ranks 1, 2 and 3). The Hoyme system used a cut-point halfway through Rank 1 to dichotomize the face (Present=Ranks 2, 3, 4 and half of Rank 1; Absent=the other half of Rank 1) (Figure 3B). Both dichotomous scales (red and black lines) failed to identify the significant correlation that exists between the severity of the FAS facial phenotype and prevalence of severe brain dysfunction (Canadian and Australian:  $\chi^2=0.5$ ,  $p=0.46$ ; Hoyme:  $\chi^2=4.7$ ,  $p=0.10$ ).

old, the majority of the 407 children with PAE under 6 years of age ( $n=238$ ) failed to receive a diagnosis under the umbrella of FASD because they presented with moderate dysfunction (ND/AE); a diagnosis excluded from the Canadian and Australian systems. Failure to identify and diagnose FAS/D in children <6 years of age will prevent these high-risk children from receiving the benefits of early intervention.

4. Growth deficiency is significantly associated with PAE, is as prevalent as the FAS facial features and CNS abnormalities, and is a highly predictive of severe CNS dysfunction among infants/toddlers. The Canadian and Australian systems removed growth deficiency as a criterion for FASD, yet growth deficiency was as strong a predictor of severe brain dysfunction in infants with PAE as the FAS facial phenotype

and microcephaly. Decades of laboratory and clinical-based studies unequivocally confirm that PAE causes GD [39-43]. While many factors can impact growth, an empirical study conducted by Astley et al., [9] confirmed that postnatal short stature is significantly correlated with PAE (while low birth weight is significantly correlated with prenatal tobacco exposure). The study found growth deficiency was as prevalent as the other core diagnostic features of FASD (FAS facial phenotype and CNS structural abnormalities). Most importantly, growth deficiency among children with PAE is highly predictive of who will present with severe CNS dysfunction. This is especially important in children <8 years of age. Astley et al. [9] found that among children under 8 years of age with PAE who present with height and/or weight at or below the 10th percentile (Growth Rank 2, 3 or 4); 57% with Growth Rank 2; 67% with Growth Rank 3 and 100% with Growth Rank 4 presented with severe CNS dysfunction after 8 years of age when they were old enough to participate in more sophisticated neuropsychological assessments. Of the 844 children classified as Not FASD, 633 had confirmed PAE. Of the 633 with confirmed PAE, 192 were under 8 years of age and 64 presented with Growth Rank 2, 3 or 4. Roughly 70% of these 64 children with confirmed PAE will likely present with severe CNS dysfunction after the age of 8 years, but are not identified as "At Risk" by the Canadian system. Of the 983 children classified as Not FASD by the Australian system, 773 had confirmed PAE. Of the 773 with confirmed PAE, 435 were under 8 years of age and 126 presented with Growth Rank 2, 3 or 4. Roughly 70% of these 126 children with confirmed PAE will likely present with severe CNS dysfunction after the age of 8 years, but are not identified as "At Risk" by the Australian system.

5. The relaxation of the Hoyme FAS facial phenotype criteria greatly increased the prevalence of FAS and PFAS diagnoses and jeopardized the validity of these FAS and PFAS diagnoses.

- The Hoyme system classified 10 times more individuals with the FAS facial phenotype ( $n=552$ ) than the 4-Digit Code ( $n=54$ ) [10].
- The Hoyme system produced 14 times more FAS/PFAS diagnoses with unknown alcohol exposure ( $n=111$ ) than the 4-Digit Code ( $n=6$ ) [10]. This is particularly concerning because 68 (61%) of these patients had 4-Digit Code Rank 1 or Rank 2 facial phenotypes that are, by our definition, clinically "normal". The Rank 1 and 2 phenotypes have no specificity to PAE [33]. The only reason FASD diagnostic systems permit a diagnosis of FAS to be made when PAE is unknown is because the facial phenotype is so highly specific to (caused only by) PAE, the face serves to confirm the exposure. If the facial phenotype defined by the diagnostic system is not confirmed to be highly specific to alcohol, then: 1) the diagnosis cannot be validly labeled FAS, PFAS or FASD because a causal link cannot be confirmed between the patient's alcohol exposure and their adverse outcomes, and 2) the facial phenotype cannot be validly used to confirm PAE when the history of exposure is unknown.

The 4-Digit Code allows a diagnosis of FAS to be made when PAE is unknown because the 4-Digit Code Rank 4 FAS facial phenotype is confirmed to be >95% specific to PAE [17, 20]. The 4-Digit Code does not allow a diagnosis of PFAS to be made when alcohol exposure is unknown, because the facial criteria for PFAS is relaxed to a Face Rank 3 (2.5 of the 3 features must be present), resulting in a subtle reduction in specificity. To err on the conservative side, the 4-Digit Code requires a confirmed exposure for PFAS

- In our previous study the relaxed Hoyme FAS facial phenotype demonstrated no association with PAE [10]. In contrast, the 4-Digit Code FAS facial phenotype demonstrated a strong, significant, linear association with PAE
- 70% of the 296 Hoyme FAS/PFAS cases had “normal” 4-Digit Code Face Ranks 1 or 2.
- 43% of the 552 patients with the Hoyme FAS face did not receive a diagnosis under the umbrella of FASD using the Hoyme system. In contrast, all 54 individuals with the 4-Digit Code Rank 4 FAS face met criteria for a diagnosis under the umbrella of FASD using the 4-Digit Code.
- Hoyme et al. [7] reports the relaxation of their facial criteria was to improve sensitivity and greater inclusion of children in the complete continuum of FASD. But, as demonstrated in this study, one need not sacrifice specificity for sensitivity to achieve greater inclusion across the full continuum of FASD. By documenting the FAS facial phenotype across its full continuum of expression (4-Digit Code Face Ranks 1, 2, 3 and 4), the 4-Digit Code preserves: 1) the high specificity of the Rank 4 FAS facial phenotype, 2) the clinically vital function of the Rank 3 face to predict severe brain dysfunction, and 3) the increased sensitivity to capture the full spectrum of FASD by inclusion of the Rank 2 face.

### Contrasts in diagnostic tools

In addition to the contrasts in diagnostic criteria, the methods and tools used to measure the facial features are also markedly different. The authors of the Hoyme system promote the use of direct examination of facial features over the use of facial photographic software. The 4-Digit Code advises measuring the facial features from 2D digital photos using the FAS Facial Photographic Analysis Software [12]. Empirical studies have confirmed the superior accuracy of the photo versus direct method of facial measurement [13,15]. Significant contrasts also exist between the 4-Digit Code Lip-Philtrum Guide 1 and the Hoyme North American Lip/Philtrum Guide. As illustrated in Figure 3, although the Hoyme North American Lip/Philtrum Guide looks similar in appearance to the 4 Digit Code Lip-Philtrum Guide 1, these are not interchangeable tools. The lips ranked 1 through 5 on the Hoyme Guide do not match the lips ranked 1 through 5 on the 4-Digit Code Guide. The lips on the 4-Digit Code Guide become progressively thinner as Rank increases from 1 to 5. The lips on the Hoyme guide do not become progressively thinner as Rank increases (e.g., the Hoyme Rank 4 lip is thicker than the Hoyme Rank 3 lip). The images used to depict lip thinness for each Rank do not match between the two

guides. When the Hoyme lips are mapped onto the 4-Digit Guide based on the objective measure of thinness (circularity), the Hoyme Rank 1, 2, 3, 4, and 5 lips are equivalent to the 4-Digit Code Lip Ranks 2, 2, 3, 2, and rank unknown, respectively. Both systems define the thin upper lip of FAS as Rank 4 or thinner. But the Hoyme Rank 4 lip is substantially thicker than the 4-Digit Rank 4 lip (it is equivalent to the 4-Digit Rank 2 lip).

The introduction of the Hoyme North American Lip/Philtrum Guide serves to further relax the Hoyme FAS facial phenotype. Only 2 of the 3 cardinal features are required and 2 of the 3 features are relaxed relative to the 4-Digit Code. The PFL is relaxed from the 3rd percentile to the 10th percentile and lip thinness is relaxed from Rank 4 to Rank 2 on the 4-Digit Code Lip-Philtrum Guide 1. An individual presenting with PFLs at the 10th percentile, a Rank 1 deeply grooved philtrum, and a 4-Digit Code Rank 2 moderately thick upper lip would meet the Hoyme criteria for the full FAS facial phenotype. The presence of a single, very minor anomaly (PFL at the 10th percentile) does not constitute a dysmorphic facial phenotype. In fact, it would be difficult to justify classifying any of these three features as minor anomalies outside the normal range. Yet, this facial phenotype is used by the Hoyme system to confirm PAE when PAE is unknown. Of the 102 patients with unknown PAE and the Hoyme FAS facial phenotype, 70% had a 4-Digit Code Rank 1 or Rank 2 facial phenotype. By definition, 4-Digit Face Ranks 1 and 2 reflect normal phenotypes with no specificity to PAE. This was clearly illustrated in our FASD MRI study [33]. Sixteen high-functioning adolescents with confirmed absence of PAE were enrolled as controls in that study. Ten presented with Rank 1 facial phenotypes and 6 presented with Rank 2 facial phenotypes (one of which illustrated in Figure 3C). Based on our previously published findings [10], the Hoyme North American Lip/Philtrum Guide is not a valid tool for use with the 4-Digit Code.

### The quintessential role of the FAS facial phenotype

Why are the criteria used to define the FAS facial phenotype so important to the medical validity of all diagnoses under the umbrella of FASD, not just the diagnosis of FAS (or FASD with the Face)? When one makes a diagnosis of FAS, one is stating explicitly that the individual has a syndrome caused by PAE [17]. One is also stating explicitly that the biological mother drank alcohol during pregnancy and, as a result, harmed her child. These are bold conclusions to draw and are not without medical, ethical, and even legal consequences. When the FAS face is not specific to FAS and PAE, the validity of the entire FASD diagnostic system collapses. Here is why:

- The terms FAS and “FASD with the Face” are rendered invalid. If the face is NOT specific to (caused only by) alcohol, one can no longer label the condition fetal alcohol syndrome or fetal alcohol spectrum disorder. One can no longer confirm alcohol is causally linked to any of the outcomes (growth, brain, or face) in an individual patient.
- The diagnosis “FAS/Alcohol Unknown” is also rendered invalid. The FAS face can no longer serve as the confirmation of alcohol exposure when the exposure history is unknown.
- The terms “ARND” and “FASD without the Face” remain problematic. Since the CNS structural and functional abnormalities that define ARND and “FASD without the Face” are not specific to (caused only by) prenatal

alcohol, one is in no position to declare an individual's Neurodevelopmental Disorder is "Alcohol-Related" (ARND) or their Spectrum Disorder is caused by Fetal Alcohol (FASD).

With terms like ARND and "FASD without the Face", one feels compelled to require a significant exposure to alcohol to increase the odds that the individual's impairments may be caused, at least in part, by their alcohol exposure. This is a dangerous road to go down.

- Setting a threshold of significant exposure for Alcohol-Related Neurodevelopmental Disorder (ARND) or FASD does not confirm the patient's alcohol exposure caused their disorder.
- Alcohol is never the only risk factor contributing to the disorder.
- One is sending a potentially harmful message that lower levels of alcohol exposure are safe. As we illustrated in our previous publication (Figure 9) [10], individuals with reported PAE below the Hoyme or Canadian thresholds do present with full FAS. Either this individual was particularly vulnerable to the teratogenic insult of alcohol, or the reported exposure was not accurate. In a clinical setting, one is never in a position to know how accurate the exposure is recalled and reported. Setting a threshold implies the details of all reported exposures are accurate and no fetus can be harmed by exposures below the threshold.
- And one is blaming a woman for harming her child, when they have limited ability to make/defend such a claim.

The 4-Digit Code introduced the terms ND/AE and SE/AE back in 1997 [23]. These terms state the verifiable facts; the individual presents with a disorder and the individual was exposed to alcohol in utero. The terminology does not explicitly state their disorder is related to their alcohol exposure. In fact, the 4-Digit Code formally Ranks all other prenatal and postnatal risks factors to make clear that alcohol is never the only risk factor contributing to an individual's neurobehavioral disorder or static encephalopathy. In 2013, the DSM5 [44] took a similar nosological approach when it introduced the new term "Neurodevelopmental disorder / prenatal alcohol exposure" (ND/PAE) as a condition for further study. "ND/PAE is characterized by a range of developmental disabilities following exposure to alcohol in utero." ND/PAE is an example of "Other Specified Neurodevelopmental Disorder (315.8 (F88)).

### When is it a FASD?

Fetal Alcohol Spectrum Disorders are, by definition, adverse outcomes caused by PAE. In the absence of an outcome that is specific to (caused only by) PAE (like the Rank 4 FAS facial phenotype), one cannot confirm or rule-out the role PAE played in an individual's CNS dysfunction.

- **Do all individuals with SE/AE, ND/AE, and ARND or "FASD without the Face" have FASD?** Not necessarily. Only the subset of individuals whose CNS dysfunction was caused (in whole or in part) by their alcohol exposure has FASD.
- **Which subset is that?** We currently have no way of knowing.

This is why the 4-Digit Code refers to SE/AE and ND/AE as "broadly" under the umbrella of FASD. Those with SE and ND caused by their alcohol exposure have FASD. Those with SE and ND that was not caused by their alcohol exposure do not have FASD.

- **But if they are exposed to high alcohol levels, can't we just assume alcohol caused their disability?** No. Not everyone exposed to high levels of alcohol presents with adverse outcomes. Among 2,576 alcohol-exposed individuals evaluated in the UW FASDPN Clinic to date, 26 with high exposures presented with full FAS (4-Digit Codes 4444) while 41 with high exposures presented with normal growth, face, and brain development (4-Digit Codes 1114) [17]. We also see discordant outcomes among fraternal twins. Among 20 twin pairs with identical high exposures, 5 had normal CNS function while their twin had moderate to severe CNS dysfunction [18].

When an individual presents with high alcohol exposure and severe CNS dysfunction, but no FAS facial phenotype, as depicted in the diagnosis SE/AE (4-Digit Code 2134):

- o If their CNS dysfunction is caused (at least in part) by their alcohol exposure, then their SE/AE is an FASD.
- o If their CNS dysfunction was caused by other risk factors, not their alcohol exposure, then their SE/AE is NOT an FASD.
- o The only way we can currently link alcohol to an individual's CNS dysfunction is if they present with a highly specific Rank 4 FAS face (FAS 2434).
- **If we cannot confirm alcohol caused a patient's disabilities, does this impact our ability to provide the patient with appropriate intervention?** No. Intervention recommendations and a patient's access to services and supports are based on their disabilities, not on what caused their disabilities. Twenty years of patient surveys [45] confirmed patients with a diagnosis of ND/AE and SE/AE were as likely to access and benefit from interventions as patients with FAS/PFAS. We did not have to label their disorder FAS or PFAS to qualify them for intervention and support services in Washington State.
- **If we cannot confirm a causal link between PAE and adverse outcomes in an individual patient, does this impact our ability to prevent FASDs?** No. To prevent FASD one must prevent PAE. To confirm efforts to prevent PAE are working, one needs to document PAE in a patient's medical record (regardless of outcome) and track the prevalence of PAE by birth cohort annually [46]. If one is reducing the prevalence of PAE, one is reducing the prevalence of FASD.

### The four diagnostic systems produce different outcomes, but which one, if any, is correct?

Validation studies are required to confirm the accuracy, reproducibility, and medical validity of a diagnostic system. Validity is the degree to which a tool (or diagnostic system) is measuring what it purports to measure [47]. When the 4-Digit Code was introduced in 1997 [23,25], it was published as an empirical study confirming its superior performance to the gestalt [48,49] approach it was designed to replace. Since then, two decades of more extensive

laboratory, clinical, and public health empirical studies have comprehensively affirmed the validity of the FASD 4-Digit Code [17]. A clinician's guide for how to fully assess the performance of FASD diagnostic systems was presented in 2013 [17] and replicated with revision below in Table 5. The guide proposes 12 questions clinicians should ask to assess the performance of FASD diagnostic systems. The 4-Digit Code's performance meets all 12 criteria.

## CONCLUSION

The needs of individuals and families impacted by FASD are best met when FASD diagnostic systems provide accurate diagnoses: 1) across the lifespan; 2) across the full spectrum of outcome (FAS, SE and ND); 3) across the full continuum of alcohol exposure; and 4) utilize diagnostic nomenclature that accurately reflects the association between outcome and alcohol exposure. These conclusions are supported by the current, published evidence base. In summary:

1. FASD is characterized by a spectrum of outcomes, not just severe outcomes.
  - a. As illustrated in a primate model of FASD (Figure 6), PAE causes a full spectrum of outcome with moderate dysfunction (ND/AE) being the most prevalent outcome (59%).
    - i) The vast majority (83%) of individuals with ND/ AE have 1 or 2 domains of severe dysfunction and multiple domains of moderate dysfunction. All require and benefit from intervention.
2. FASD is caused by the full continuum of PAE, not just high exposure.
  - a. There is no known safe level of alcohol use during pregnancy.
  - b. Requiring high PAE implies reported levels of PAE are reliably accurate. They are not.
  - c. When high PAE is required for diagnosis, over half of individuals with confirmed PAE and severe CNS abnormalities do not receive a diagnosis of FASD.
3. FASD is present at birth and should be diagnosed as early as possible, not after 3 or 6 years of age.
  - a. Requiring severe CNS dysfunction prevents a diagnosis of FASD in a child too young to be fully assessed for CNS dysfunction.
  - b. Excluding moderate CNS dysfunction from the umbrella of FASD prevents the early identification and intervention of children with confirmed PAE and moderate dysfunction (ND/AE).
  - c. Excluding growth deficiency prevents the early identification of children who are at especially high risk for severe CNS dysfunction later in childhood.
  - d. Children under 6 years of age with confirmed PAE and moderate dysfunction are not "At Risk" for FASD. Their alcohol exposures and moderate dysfunction have already occurred and warrant a diagnosis that documents their disability and qualifies them for early intervention.
4. FASD is characterized by growth deficiency, FAS facial features, and CNS structural/neurological/and functional abnormalities. Each present along clinically meaningful continuums and each are significantly correlated with PAE.
5. Growth deficiency is a core component of FASD.
  - a. Growth deficiency ( $\leq 10$ th percentile) is as prevalent or more prevalent among individuals with PAE (32%) than the FAS facial phenotype (4%) and severe CNS abnormalities (39%).

**Table 5.** As clinicians assess the performance of FASD diagnostic guidelines, clinicians should ask the following questions [17].

1. Have properly designed studies been published to confirm the case definition for the FAS facial phenotype is highly specific (>95%) to FAS and alcohol (e.g., observed only among individuals with prenatal alcohol exposure and FAS)?
2. Was data used to empirically derive the diagnostic guidelines? Was the data drawn from a large, representative, and population-base?
3. Has the performance of the guidelines been empirically assessed (validated)?
4. Individuals are born with FAS/D. Can the diagnostic system identify FAS/D at birth and across the lifespan?
5. Growth deficiency, the FAS facial phenotype, CNS abnormalities, and alcohol exposure all present along clinically meaningful continuums. The FAS facial phenotype is not just present or absent. The brain is not just normal or abnormal. Do the Guidelines recognize/incorporate these important continuums?
6. Do the guidelines produce clinically distinct subgroups across the full spectrum (FAS, PFAS, SE/AE, ND/AE)?
A. Do brain imaging studies identify statistically significant contrasts between the FASD subgroups?
B. Individuals with FAS have more severe CNS dysfunction than individuals with "ARND". Do the Guidelines generate FAS and "ARND" groups that demonstrate this important contrast?
C. Do individuals who meet the criteria for FAS actually have FAS?
7. Can the guidelines detect unique alcohol exposure patterns between the FASD subgroups?
8. Can the diagnostic system be effectively and efficiently taught to interdisciplinary teams?
9. Are the guidelines confirmed to be reproducible? If two clinics use the guidelines, do they render the same diagnoses?
10. Do families report high satisfaction/confidence with the diagnostic process and outcome?
11. Are the names of the diagnoses (FAS, PFAS, SE/AE, ND/AE, ARND, ARBD, FASD with the Face, FASD without the Face) medically valid? Do they imply causality between alcohol and outcome that cannot be confirmed in the individual patient?
12. Do diagnoses under the umbrella of FASD qualify patients for intervention services that lead to improved outcomes?

- b. The 4-Digit Code method for ranking growth deficiency successfully differentiates growth deficiency (postnatal short stature) associated with PAE from growth deficiency (low birth weight) associated with other risk factors like tobacco [9].
  - c. Growth deficiency ( $\leq 10$ th percentile) in infants/toddlers with PAE is as predictive of severe CNS dysfunction later in childhood as the Rank 4 FAS facial phenotype and microcephaly.
6. The 4-Digit Code Rank 4 FAS facial phenotype is the only outcome confirmed to date that is highly specific to (caused only by) PAE. This high specificity is required:
    - a. To render a diagnosis of FAS when PAE is unknown.
    - b. To confirm PAE is causally associated with outcomes in an individual patient.
    - c. To validly label the disorder FAS or FASD.
  7. Diagnostic nomenclature (e.g., ARND, FASD without the Face) should not infer a causal association between a patient's PAE and adverse outcomes when there is no evidence to validate such an inference.
    - a. Inferring causation may erringly impugn some birth mothers.
    - b. Effective intervention and prevention does not require confirmation of causation.

## ETHICAL APPROVAL

This study was approved by the University of Washington Human Subjects Division.

## AUTHOR CONTRIBUTION

All authors are members of the interdisciplinary FASD diagnostic team and participated in the interpretation and reporting of the study's outcomes. SJAH conducted the statistical analyses.

## ACKNOWLEDGEMENTS

Special thanks are extended to the patients and their families for their benevolent contributions to the University of Washington FASDPN dataset.

## COMPETING INTERESTS

The authors do not have any competing interests

## SOURCES OF FUNDING

The Washington State Fetal Alcohol Syndrome Diagnostic & Prevention Network (FASDPN) has been supported over the past two decades by the following organizations: Centers for Disease Control and Prevention (1992-1997); Western Washington Chapter of the National March of Dimes Birth Defects Foundation (1995); Washington State Department of Social and Health Services through the passage of Senate Bill SB5688 (1997-present); and the Chavez Memorial Fund. Support has also been received from the Center on Human Development and Disability, University of Washington since 1993 (National Institute of Child Health and Human Development: P30 HD02274). The brain imaging data used in this study was obtained through support by the National Institute on Alcohol Abuse and Alcoholism R01-AA12915-01. The primate data used in this study was obtained through support by the National Institute on Alcohol Abuse and Alcoholism R01-AA05616-07.

## REFERENCES

1. Astley S. Profile of the first 1,400 patients receiving diagnostic evaluations for fetal alcohol spectrum disorder at the Washington State Fetal Alcohol Syndrome Diagnostic & Prevention Network. *Can J Clin Pharmacol*. 2010;17(1):e132-164.
2. Kodituwakku P, Kodituwakku E. Cognitive and behavioral profiles of children with fetal alcohol spectrum disorders. *Curr Dev Disord Rep*. 2014;1:149-160.
3. 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.
4. Clarren SK, Astley SJ, Gunderson VM, Spellman D. Cognitive and behavioral deficits in nonhuman primates associated with very early embryonic binge exposures to ethanol. *J Ped*. 1992;121(5):789-796.
5. Astley SJ. Diagnostic guide for fetal alcohol spectrum disorders: The 4-digit diagnostic code, 3rd edition. University of Washington publication services, Seattle, WA. [accessed 2019 September 1] 2004.
6. 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. *Can Med Assoc J*. 2015.
7. 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;138(2):e20154256.
8. Bower C and Elliott EJ. 2016, on behalf of the Steering Group. Report to the Australian Government Department of Health: "Australian Guide to the diagnosis of Fetal Alcohol Spectrum Disorder (FASD)".
9. Astley SJ, Bledsoe JM, Davies JK. The essential role of growth deficiency in the diagnosis of fetal alcohol spectrum disorder. *Advances in Pediatric Research* 3:9. doi:10.12715/apr.2016.3.9.
10. Astley SJ, Bledsoe JM, Davies JK, Thorne JC. Comparison of the FASD 4-Digit Code and Hoyme et al. 2016 FASD diagnostic guidelines. *Adv Ped Res*. 2017;4:13.
11. Hoyme HE, Hoyme DB, Elliot AJ, Blankenship J, Kalberg WO, Buckley D, et al. A south african mixed race lip/philtrum guide for diagnosis of fetal alcohol spectrum disorders. *Am J Med Genet A*. 2015;167A(4):e752-755.
12. Astley SJ. FAS facial photographic analysis software manual V2.1.0, 2016. [accessed 2019 September 1]. [http://depts.washington.edu/fasdpn/pdfs/FAS\\_Instruction\\_Manual\\_v2.1.0-050616.pdf](http://depts.washington.edu/fasdpn/pdfs/FAS_Instruction_Manual_v2.1.0-050616.pdf).
13. Astley SJ. Canadian palpebral fissure length growth charts reflect a good fit for two school and FASD clinic-based U.S. populations. *J Popul Ther Clin Pharmacol*. 2011;18(2):e231-241.
14. Clarren SK, Chudley AE, Wong L, Friesen J, Brant R. Normal distribution of palpebral fissure lengths in Canadian school age children. *Can J Clin Pharmacol*. 2010;17(1):e67-78.
15. Astley SJ. Palpebral fissure length measurement: Accuracy of the FAS facial photographic analysis software and inaccuracy of the ruler. *J Popul Ther Clin Pharmacol*. 2015;22(1):e9-26.
16. Astley SJ, Clarren SK. Measuring the facial phenotype of individuals with prenatal alcohol exposure: Correlations with brain dysfunction. *Alcohol & Alcoholism*. 2001;36(2):147-159.
17. Astley SJ. Validation of the fetal alcohol spectrum disorder (FASD) 4-Digit Diagnostic Code. *J Popul Ther Clin Pharmacol*. 2013;20(3):e416-467.
18. Astley Hemingway SJ, Bledsoe JM, Davies JK, Brooks A, Tracy J, Olson EM, et al. Twin study confirms virtually identical prenatal alcohol exposures can lead to markedly different fetal alcohol spectrum disorder outcomes - fetal genetics influences fetal vulnerability. *Adv Ped Res*. 2019;5:23.

19. May PA, Kalberg WO, Hoyme HE. Practical and accurate methods for diagnosing the full spectrum of FASD for both clinical and research applications: Employing the IOM recommended criteria. 5th international conference on FASD: Session E4. Feb 27-March 2, 2013 Vancouver BC, Canada. [accessed 2019 September 1] [https://interprofessional.ubc.ca/files/2016/07/FASD2013\\_Brochure.pdf](https://interprofessional.ubc.ca/files/2016/07/FASD2013_Brochure.pdf).
20. Astley SJ, Clarren SK. A case definition and photographic screening tool for the facial phenotype of fetal alcohol syndrome. *J Ped.* 1996;129:33-41.
21. Astley SJ. Diagnosing fetal alcohol spectrum disorders (FASD). In: Aduabato SA and Cohen DE (eds.) Prenatal alcohol use and fetal alcohol spectrum disorders: Diagnosis, assessment and new directions in research and multimodal treatment, Bentham Science Publishers Ltd. 2011;pp. 3-29.
22. Coles CD, Gailey AR, Mulle JG, Kable JA, Lynch ME, Jones KL. A comparison among 5 methods for the clinical diagnosis of fetal alcohol spectrum disorders. *Alcohol Clin Exp Res.* 2016;40(5):e1000-1009.
23. Astley SJ, Clarren SK. Diagnostic guide to FAS and related conditions: The 4-digit diagnostic code 1st ed. Seattle: university of washington publication services 1997.
24. Astley S, Clarren S. Diagnostic guide for fetal alcohol syndrome and related conditions: the 4-digit diagnostic code. 2nd ed. Seattle: University of washington publication services 1999.
25. Astley SJ, Clarren SK. Diagnosing the full spectrum of fetal alcohol exposed individuals: Introducing the 4-digit diagnostic code. *alcohol and alcoholism.* 2000;35(4):400-410.
26. CDC.gov. CDC Growth Charts [accessed 2019 September 1]. (boys) <https://www.cdc.gov/growthcharts/data/set2clinical/cj41c071.pdf> (girls) <https://www.cdc.gov/growthcharts/data/set2clinical/cj41c072.pdf>.
27. WHO.int [Internet]. The WHO child growth standards [accessed 2019 September 1]. <http://www.who.int/childgrowth/en>.
28. Stromland K, Chen Y, Norberg T, Wennerstrom K, Michael G. Reference values of facial features in Scandinavian children measured with a range-camera technique. *Scand J Plast Reconstr Hand Surg.* 1999;33:59-65.
29. Hall JG, Froster-Iskenius UG, Allanson JE. Handbook of normal physical measurements. New York, USA: Oxford university press 1989.
30. Ross laboratories. Ross growth & development program. Parent-specific adjustments for evaluation of length and stature- boys and girls, 1983, Ross Products Division, Abbott Laboratories, Columbus Ohio, 43216.
31. Nellhaus G. Composite international and interracial graphs, head circumference, girls and boys, birth to 18 years. *Pediatrics.* 1968;41:106.
32. Astley SJ. Comparison of the 4-digit diagnostic code and the hoyme diagnostic guidelines for fetal alcohol spectrum disorders. *Pediatrics.* 2006;118(4):1532-1545.
33. Astley SJ, Olson HC, Kerns K, Brooks A, Aylward E, Coggins T, et al. Neuropsychological and behavioral outcomes from a comprehensive magnetic resonance study of children with fetal alcohol spectrum disorders. *Can J Clin Pharmacol.* 2009;16(1):e178-e201.
34. Astley S, Grant T. Another perspective on "The effect of different alcohol drinking patterns in early to mid pregnancy on the child's intelligence, attention, and executive function". *Br J Obstet Gynaecol.* 2012;1672.
35. Clarren S, Astley S, Bowden D. Physical anomalies and developmental delays in nonhuman primate infants exposed to weekly doses of ethanol during gestation. *Teratology.* 1988;37:561-569.
36. Astley SJ, Aylward EH, Olson HC, Kerns K, Brooks A, Coggins TE et al. Magnetic resonance imaging outcomes from a comprehensive magnetic resonance study of children with fetal alcohol spectrum disorders. *Alcohol Clin Exp Res.* 2009;33(10):1-19.
37. Jirikowic T, Gelo J, Astley S. Children and youth with fetal alcohol spectrum disorders: Summary of intervention recommendations after clinical diagnosis. *Intellectual and Developmental Disabilities.* 2010;48(5):330-344.
38. Petryk SC, Ekeh J, Pandey M, The prenatal alcohol history – it is hard to get and it matters how we define it. 7th international conference on FASD research: Results and relevance: Session D3. March 1-4, 2017 Vancouver BC, Canada. [accessed 2019 September 1] <http://interprofessional.ubc.ca/files/2017/03/D3a-Petryk.pdf>.
39. Abel EL, Dintcheff BA. Effects of prenatal alcohol exposure on growth and development in rats. *J Pharmacol Exp Ther.* 1978;207(3):916-21.
40. Behnke M, Smith VC. Prenatal substance abuse: short and long-term effects on the exposed fetus. *Pediatrics.* 2013;131(3):e1009-24.
41. Gavin CE, Kates B, Gerken LA, Rodier PM. Patterns of growth deficiency in rats exposed in utero to undernutrition, ethanol, or the neuroteratogen methylazoxymethanol (MAM). *Teratology.* 1994;49(2):113-21.
42. Middaugh LD, Boggan WO. Postnatal growth deficits in prenatal ethanol-exposed mice: characteristics and critical periods. *Alcohol Clin Exp Res.* 1991;15(6):919-926.
43. Nordstrom-Klee B, Delaney-Black V, Covington C, Ager J, Sokol R. Growth from birth onwards of children prenatally exposed to drugs. A literature review. *Neurotoxicol Teratol.* 2002;24:481-488.
44. American psychiatric association. Diagnostic and statistical manual of mental disorders: DSM-5. Washington, D.C: American psychiatric association 2013.
45. Astley SJ. Twenty years of patient surveys confirm a FASD 4-Digit-Code interdisciplinary diagnosis afforded substantial access to interventions that met patients' needs. *J Popul Ther Clin Pharmacol.* 2014;6:21(1):e81-105.
46. Astley SJ. Fetal alcohol syndrome prevention in Washington State: Evidence of success. *Paediatric and Perinatal Epidemiology.* 2004;18:344-351.
47. Litwin M. How to measure survey reliability and validity. Thousand Oaks: Sage Publications 1995.
48. Sokol RJ, Clarren SK. Guidelines for use of terminology describing the impact of prenatal alcohol on the offspring. *Alcoholism: Clinical and Experimental Research.* 1989;13(4):597-598.
49. Stratton K, Howe C, Battaglia F. Fetal alcohol syndrome: Diagnosis epidemiology prevention and treatment. Institute of medicine. Washington D C National Academy Press 1996.