ELSEVIER

Alcohol 44 (2010) 605–614

ALCOHOL

# Diagnosis of fetal alcohol spectrum disorders: a validity study of the fetal alcohol syndrome checklist

Larry Burd[a,*], Marilyn G. Klug[a], Qing Li[a], Jacob Kerbeshian[b], John T. Martsolf[a]

[a]*Department of Pediatrics, University of North Dakota School of Medicine and Health Sciences, Grand Forks, ND, USA*
[b]*Department of Neuroscience, University of North Dakota School of Medicine and Health Sciences, Grand Forks, ND, USA*

Received 7 October 2008; received in revised form 19 May 2009; accepted 12 August 2009

**Abstract**

Fetal alcohol spectrum disorders (FASD) are a common cause of developmental disability, birth defects, and mortality. The performance characteristics of current diagnostic tools for FASD are not adequately reported. This study examines the performance characteristics of the Fetal Alcohol Syndrome Diagnostic Checklist (FASDC). In a population of 658 subjects from North Dakota, we used the FASDC score to examine the agreement between FASDC score, clinical diagnosis, and the Institute of Medicine criteria for FASD. All subjects were seen for evaluation in the genetic/dysmorphology clinics, which are funded by the state to provide genetic diagnostic services for residents of North Dakota. We compared the clinical diagnosis and the FASDC scores to determine the performance characteristics of the FASDC in the categorical diagnosis of fetal alcohol spectrum (FAS), other-FASD, and a group with No-FASD. Comparisons were made using univariate and logistic models of outcomes using both the presence and the absence of alcohol exposure or FASDC phenotype data. The FASDC performance characteristics for differentiation of the FAS group from non-FASD were excellent (accuracy 99%, sensitivity 99%, and specificity 99%). Logistic models for subjects with scores in the FASD range were differentiated with an accuracy of 82%, sensitivity 85%, and specificity 80% using the data on phenotype and exposure. We were able to delineate subjects with scores in the No-FASD range with an accuracy of 78%, sensitivity 64%, and specificity 81% without including the exposure and phenotype data by use of the other descriptive data (maternal characteristics, birth records, and demographic data) from the FASDC. All diagnostic tools should have performance characteristics assessed and available before adoption for use in clinical settings. The FASDC scores produce diagnostic groupings that approximate expert clinical judgment. The tool may be useful in other clinical settings for the diagnosis of FASD or as an FASD registry or research database. 

*Keywords:* Fetal alcohol spectrum disorders; Diagnosis; Sensitivity; Specificity; Accuracy; Children

## Introduction

Of the 4 million annual pregnancies in the United States (US), about 40% of women drink some alcohol during pregnancy and about 3–5% of women drink heavily throughout pregnancy (Folyd and Sidhu, 2004). The results of this exposure range from mortality to no observable adverse effects (Abel, 1998). Between these extremes is a large population of people with highly variable patterns of outcomes from prenatal exposure (Sampson et al., 1997). Much of the variance in prevalence estimates likely results from actual variation in prevalence, differing strategies for ascertainment, and variance in exposure (Burd and Moffatt, 1994). Differing diagnostic strategies and inconsistent application of these strategies may also account for a substantial portion of the variance in prevalence rates.

At a recent consensus meeting, numerous groups selected the term fetal alcohol spectrum disorders (FASD) as a conceptual term to describe the range of adverse outcomes resulting from prenatal alcohol exposure. Current prevalence estimates for FASD range from a rate of 0.30 (fetal alcohol syndrome only) to 9.1 cases per 1,000 live births (FASD) (Abel, 1998; Sampson et al., 1997). This suggests that the

This study was supported by Northern Plains Prenatal and Infant Health Consortium project (Department of Health and Human Services [HHS] from National Institutes of Health/National Institute of Child Health and Human Development, National Institute on Alcohol Abuse and Alcoholism #5 U01 HD45935-02 [NIH]); Fetal Alcohol Syndrome Prevention Project (Department of HHS from Centers for Disease Control and Prevention #U84/CCU823298-02); and Consortium on Fetal Alcohol Syndrome Project (Center for Substance Abuse Prevention from Substance Abuse and Mental Health Services Administration 1SP10834-01).

* Corresponding author. North Dakota Fetal Alcohol Syndrome Center, University of North Dakota School of Medicine and Health Sciences, 501 N Columbia Road, Grand Forks, ND 58203, USA. Tel.: +1-701-777-3683; fax: +1-701-777-4474.

*E-mail address*: laburd@medicine.nodak.edu (L. Burd).

0741-8329/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi: 10.1016/j.alcohol.2009.08.010

annual number of affected pregnancies in the US would range from 1,200 (4.0 million pregnancies × rate of 0.30 per 1,000 live births) to 36,400 (4.0 million × rate of 9.1 per 1,000 live births).

Because several thousand affected children are born each year in the US alone, the need for reliable and valid diagnostic systems is clear. In addition, the epidemiological performance characteristics of diagnostic schema (sensitivity, specificity, accuracy, and validity studies in clinical practice) are crucial for patient care, prevalence estimates, and outcome studies. Although multiple diagnostic schemas for FASD are currently recommended for use, few have available published epidemiological performance criteria from population-based studies (Astley, 2006; Bertrand et al., 2004; Burd et al., 2003c; Hoyme et al., 2005; Stratton et al., 1996).

The classification schema proposed by the Institute of Medicine (IOM) uses several diagnostic categories. In our experience, we have found that the categories may not be mutually exclusive (Burd et al., 2003c). The lack of categorical exclusivity would not be unique to the diagnosis of FASD. Thus, as it currently stands, the IOM criteria are valuable but lack the advantage of validity studies from clinical populations.

In previous publications, we have described the development of the Fetal Alcohol Syndrome Diagnostic Checklist (FASDC) (Burd et al., 2003c). The tool has been in use in North Dakota since 1984 and is a key component of the North Dakota fetal alcohol spectrum (FAS) registry (Burd and Martsolf, 1989). The FAS registry was established in 1984 and consists of consecutive referrals to the North Dakota Genetics–Dysmorphology Program with a referral question of FAS or subjects who had a diagnosis of FAS, partial FAS, or alcohol-related neurodevelopmental disorder (ARND). The North Dakota Genetics–Dysmorphology Program is statewide and has multiple outreach clinics across the state, including clinics on three of the four tribal nations.

In a previous publication, we presented data comparing the FASDC with the IOM diagnostic criteria on FAS (Burd et al., 2003c). Hoyme et al. (2005) then independently, published diagnostic criteria comparing their diagnostic system to modified IOM criteria. In Table 1, we list the criteria for both the FASDC and the Hoyme et al. tools.

In this article, we have used a statewide clinical population to compare the accuracy of the FASDC performance in a cohort of subjects from North Dakota. The FASDC score is a composite of the alcohol exposure score and the phenotype score, which produces an overall estimate of severity (range 0–205, Table 1). The phenotype data consist of major criteria and minor criteria. The major criteria consist of the detailed dysmorphology exam variables commonly used in the diagnosis of FASD, and minor criteria which are infrequently reported in FASD but which may be useful in identification of other syndromes commonly seen in a birth defects or developmental disorders diagnostic clinic. The minor criteria are scored but are not included in the calculation of the total FASDC score. To determine the validity of this diagnostic test by comparison against the IOM criteria in an appropriate spectrum of subjects, we examined the agreement between subjects with a clinical diagnosis of FAS, subjects with other-FASD, and a group where FASD had been excluded (Greenhalgh, 1997).

We had three specific questions:

1. What combination of alcohol exposure and phenotype scores provides the best cutoff for distinguishing between FAS and No-FASD?
2. What distinguishes those children with other-FASD from those with FAS when they have similar but overlapping FASDC scores? This is important because diagnostic tools nearly always include some subjects in whom the scores overlap and are compatible with two or more options for a categorical diagnosis. Which of these factors comprise the most useful criteria for categorical diagnosis when exposure and phenotype are considered and when they are not? Are other data available that may be useful in the diagnosis of FASD that are not included in a diagnostic assessment? These might be conceptualized as modifying variables history of FASD in siblings, maternal mortality, race, maternal smoking, maternal age, age at evaluation, child's placement, for example, foster care or adopted, and other population or demographic variables that might add important information to the diagnostic formulation.
3. What factors distinguish children with other-FASD from those with No-FASD when they have similar FASDC scores? Which of these factors are most useful when exposure and phenotype are considered and when they are not used and other data are considered that may be useful in clinical delineation of the other-FASD categorical diagnoses?

## Materials and methods

### *FASD registry inclusion criteria and study participants*

The North Dakota Fetal Alcohol Syndrome Registry database contains information on consecutive subjects seen for diagnostic evaluation at genetic and dysmorphology clinics across the state of North Dakota from 1984 to 2003. North Dakota has a population of 632,000 and has about 8,000 births each year. The registry data were developed from a statewide medical genetics and dysmorphology service funded to provide services for the entire state of North Dakota. The clinics are held across the state at six locations at least annually and at two locations monthly or by appointment. Referrals are made by pediatricians, family practice physicians, obstetricians, Indian Health Service physicians, parents, foster care providers, social services agencies, and multiple other referral sources. The registry also contains all cases of FASD seen at the North

Table 1
Comparison of diagnostic criteria for the modified Institute of Medicine criteria by Hoyme et al. (2005) and the criteria for the Fetal Alcohol Syndrome Diagnostic Checklist (FASDC) (Burd et al., 2003a)

| Feature | Hoyme et al. (2005) Points | Burd and Martsolf (2003) Points |
|---|---|---|
| Current height | 1 (<10) | 4 (<5) |
| Current weight | 2 (<10) | 6 (<5) |
| Occipitofrontal circumference | 3 (<10) | 10 (<2) |
| Birth weight | — | 4 (<10) |
| Birth length | — | 6 (<10) |
| Birth occipitofrontal circumference | — | 10 (<10) |
| Palpebral fissure length | 3 (<10) | 5 (<10) |
| Midfacial hypoplasia | 2 | 4 |
| "Railroad track" ears | 1 | — |
| Strabismus | 0 | 2 |
| Ptosis | 2 | 4 |
| Epicanthal folds (nonracial) | 1 | 1 |
| Protruding helical root | — | 3 |
| Protruding auricle | — | 3 |
| Flat (low) nasal bridge | 1 | 1 |
| Anteverted nares | 2 | 2 |
| Long philtrum | 2 | 3 |
| Smooth philtrum | 3 | 4 |
| Thin vermilion border of upper lip | 3 | 4 |
| Cleft lip/palate | — | 3 |
| Relative prognathism | — | 2 |
| Cardiac murmur | 0 | 2 |
| Cardiac malformation (confirmed) | 1 | 4 |
| Pectus excavatum | — | 2 |
| Hypoplastic nails | 0 | — |
| Klippel–Feil anomaly | — | 3 |
| Meningomyelocele | — | 3 |
| Unable to fully supinate forearm (elbow) | — | 3 |
| Short fifth metacarpel | — | 3 |
| Clinodactyly of fifth digits | 1 | 2 |
| Camptodactyly | 1 | 3 |
| Sharply angulated distal palmar crease | — | 3 |
| Multiple or raised hemangiomas | — | 3 |
| Hirsutism | 1 | 2 |
| Hypoplasia of distal phalanges | — | 4 |
| Bone age 1–2 S.D. below mean | — | 1 |
| Bone age >2 S.D. below mean | — | 3 |
| Attention-deficit hyperactivity disorder | 1 | 6 |
| Fine motor dysfunction | 1 | — |
| Mental retardation (IQ <70) | — | 10 |
| IQ 70–80 | — | 5 |

| Hoyme et al. (2005) | Burd and Martsolf (2003) Prenatal alcohol exposure | | | |
|---|---|---|---|---|
| | Daily | Weekly | Binges per month | Score |
| | – | — | — | 0 |
| | – | 1–2 days per week | 0–2 | 30 |
| | – | Most weeks | 3 or more | 40 |
| | + | Nearly all weeks | Variable | 50 |
| | | Major criteria score 0–155 | | |
| | | Minor criteria score 0–36* | | |
| | | Alcohol exposure score 0–50 | | |
| | | Total score | | |

*For the FASDC, the major criteria have scores from 0 to 155. The alcohol exposure score ranges from 0 to 50. The minor criteria are each scored 1 if present, but the minor criteria are not included in the total FASDC score.

Dakota Fetal Alcohol Syndrome Clinic. The FAS clinic provides developmental and neuropsychiatric evaluations and follow-up for subjects with FASD for the state. The registry included all subjects identified by the two community-level, population-based screening programs in the state, which have been ongoing for many years. To our

knowledge, nearly all cases of FASD diagnosed at the time of this study in the state were contained in this registry.

This data set consisted of 658 subjects. Of the subjects, 152 (23.1%) had a clinical diagnosis of FAS, 167 (25.6%) other-FASD (partial FAS or ARND), and 339 subjects (51.5%) who did not have either FAS or other-FASD (for simplicity, we will refer to this group as No-FASD). All subjects also had the FASDC completed at the time of the clinical evaluation. The No-FASD group comprised 339 subjects referred for evaluation with attention-deficit hyperactivity disorder (ADHD), chromosomal abnormalities, various syndromes, and familial neuropsychiatric disorders and subjects who were evaluated and did not receive a specific clinical diagnosis.

Each child received an evaluation that included assessment by a board-certified medical geneticist who is also a pediatrician with more than two decades of experience with FASD. The genetics and dysmorphology evaluation uses a clinical approach to diagnosis and, as we have described, is consistent with both the IOM and Hoyme criteria (Burd et al., 2003a). The FASDC is the diagnostic format used for the systematic evaluation of patients referred for FASD evaluation in North Dakota and inclusion in the state FAS registry. This evaluation procedure has been described elsewhere (Burd and Kerbeshian, 1988; Burd and Martsolf, 1989; Burd et al., 2001, 2003a, 2003c).

### *Variables*

The diagnostic features of the FASDC and the scores for each are listed in Table 1. Alcohol exposure was measured from the FASDC exposure scale, which collects categorical data on four exposure groups: no alcohol use (score 0), low use (score 30–35), moderate use (score 40–45), and high use (score 50–65). The FASDC uses a best-fit exposure score for each patient based on either binge drinking or a pattern of weekly drinking. Binge drinking has been reported to be 79–83% sensitive and specific for harmful drinking or dependence (Bush et al., 1998). If a history of binge drinking (five or more drinks on an occasion) is present, the score is developed from the binge drink frequency. The four categorical exposure groups (none, exposure score of 0; up to one to two binges per month, score of 30; three binges monthly, score of 40; multiple monthly binges (four or more) and additional daily or weekly drinking below binge levels, score of 50). If no history of binge episodes is obtained or the predominant pattern of drinking is nonbinge drinking, the exposure score is developed using the daily or weekly pattern of drinking (drinks 1–2 days per week but not most weeks, score of 30; drinks most weeks, score of 40; drinks nearly all weeks and often multiple days per week, score of 50). For nonbinge drinkers, the mean cumulative exposure estimates during pregnancy are mild, 160 drinks; moderate, 320 drinks; and heavy, 560 drinks. It is important to emphasize that assessment of exposure is nearly always a challenge in a population of children where the mother's self-report of exposure is often not available and in some cases may not be reliable. Although accurate maternal self-report is ideal, exposure assessment is often dependent on reporting from the woman's partner, relatives who had close contact with her during pregnancy, or those cases where treatment or medical records provide useful data on exposure. Other variables that might influence a diagnosis of FAS, or other-FASD, included gender, race, age at diagnoses, year evaluated, foster care or adopted, intelligence scores (verbal, performance, and full scale), comorbid conditions (ADHD, tics, obsessive–compulsive disorder, mood disorders, anxiety disorders, conduct disorder, oppositional defiant disorder, pervasive developmental disorder, developmental disorders, neurological disorders, other psychiatric disorders, on medications, sleep disorders, anger disorders, stuttering, self-injury, social skills problems, schizophrenia, bipolar disorders, alcohol problems, and drug problems), total number of comorbid conditions, and length of gestation. Meanwhile, we also considered parental influences, including maternal: age, education, marital status, gravida, number of live births now dead, parity, having been in substance abuse treatment or currently in treatment, number of times in treatment, and smoking during pregnancy; and paternal: age, education, marital status, and having been or currently in substance abuse treatment.

### *Statistical analysis*

Univariate analyses of Chi-square, relative risks, and independent *t*-tests were used to determine which of these variables influenced a clinical diagnosis of other-FASD rather than FAS or No-FASD. We estimated the performance characteristics of the FASDC (accuracy, sensitivity, and specificity). Logistic regression was used to model the optimal set of variables for diagnostic categorization. Although the range of scores on the FASDC is considerable (0–205), we do consider it important to allow for inclusion of low to high FASDC scores even if the extreme scores are not common.

## Results

### *Separation of diagnostic categories*

The regression line, phenotype score $= -5/6 \times$ alcohol exposure score $+ 55$, was estimated to best separate the FAS and No-FAS subjects using their FASDC scores. Fig. 1 shows the distribution of the 152 FAS and 339 No-FASD cases by their alcohol exposure and phenotype scores from the FASDC. Table 2 shows the cutoff scores based on this regression equation and the number of cases that were correctly predicted by the separation line. This cutoff correctly classified 148 of the 152 FAS cases (97.4%) as FASD, with 114 (77%) having an alcohol exposure score of 50 and a phenotype score above 13. The No-FASD cases were also well delineated (337 of the 339 No-FASD cases or 99.4%) with 276 (81.4%) having an alcohol exposure score

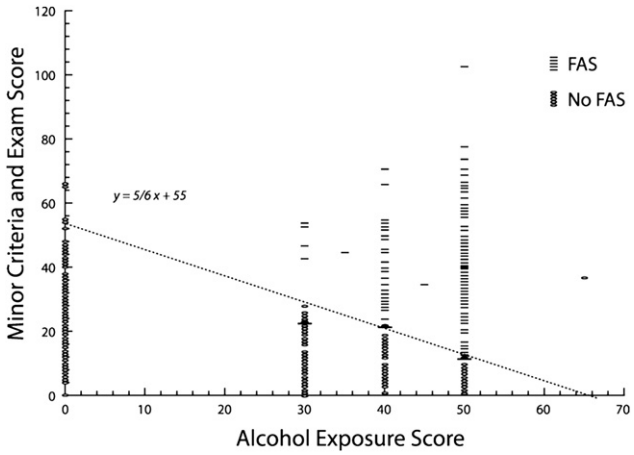


Fig. 1. Regression line for separation of subjects with FAS ($n = 152$) and a diagnosis of No-FASD ($n = 339$).

of 30 or less and phenotype of 30 or higher. The performance characteristics of the FASDC for discrimination of these two groups were accuracy of 98.8%, sensitivity 98.7%, and specificity 98.8%.

The 148 FAS cases and 337 No-FASD cases that were accurately delineated were then used as comparison groups for the 167 other-FASD cases. Using the FASDC scores, the 167 other-FASD cases were partitioned into two groups: those with FASDC scores in FASD diagnosis range (above the regression line in Fig. 1) and those whose FASDC scores were in the No-FASD score range (below the regression line in Fig. 1). We refer to the subjects above the regression line who had a diagnosis of other-FASD as *estimated-FASD* ($n = 105$) and subjects below the regression line with a diagnosis of other-FASD as *estimated-No-FASD* ($n = 62$). In Fig. 2, we present a graphical summary of the classification of cases from Fig. 1.

Table 2 shows the alcohol exposure and phenotype scores for the estimated-FASD and estimated-No-FASD. Of the 105 other-FASD cases above the regression line (estimated-FASD), 50% had exposure scores of 50 compared to 77% of the cases with a diagnosis of FAS ($P \leq .001$). Of the 62 other-FASD cases with FASDC scores below the regression line (estimated-No-FASD), 52% had exposure scores of 30 or lower, compared with 82% of the No-FASD cases ($P \leq .001$).

### *Distinguishing the FASD phenotype*

To identify risk and environmental factors that influence a categorical diagnosis of other-FASD instead of FAS or No-FASD, the 105 other-FASD cases that were estimated-FASD were compared with the 148 FAS cases. Table 3 shows the factors that were found to be statistically significant in differentiating between subjects with a diagnosis of other-FASD who were estimated-FASD and subjects with FAS. Although both groups had similar FASDC scores, the estimated-FASD group was more likely to have received their diagnosis in recent years ($P = .047$), have an anger disorder ($P = .023$), or be on other psychoactive medications ($P = .026$) and to have had a normal duration of gestation ($P = .005$) but not to have a neurological disorder ($P = .035$). The mothers of the other-FASD group were more likely to be younger ($P = .003$) and nonsmokers ($P < .001$).

We then used logistic regression and entered these variables without the FASDC alcohol exposure and phenotype scores to find combinations of variables that significantly associated with an FAS diagnosis versus a diagnosis of other-FASD. Only smoking and calendar year of diagnosis were significant predictors when all variables were considered together (Table 4). The parameters for classification of subjects as FAS or other-FAS in this model were accuracy 60.1%, sensitivity 81.0%, and specificity 45.3%. Adding the FASDC alcohol exposure and phenotype scores to the model increased accuracy from 60.1 to 81.8% (+21.7%), sensitivity from 81.0 to 84.8% (+3.7%), and specificity from 45.3 to 79.7% (+34.4%).

### *Distinguishing No-FASD cases*

In like manner, the 62 other-FASD cases in the estimated-No-FASD group were compared with the 337 No-FASD cases. Table 5 presents a comparison of subjects with FASDC scores below the regression line in Fig. 3 who have a diagnosis of other-FASD (the estimated-No-FASD group) compared with subjects with a diagnosis of No-FASD. Thirteen factors were identified that were associated with low FASDC exposure and phenotype scores and a diagnosis of other-FASD rather than a diagnosis of No-FASD. Subjects were more likely to receive a diagnosis of other-FASD and to have estimated FASDC score in the No-FASD

Table 2
Prevalence of FAS, other-FASD, and No-FASD diagnoses by prenatal alcohol exposure score and phenotype cutoff scores from FASDC

| | Alcohol exposure scores | | | | | | |
|---|---|---|---|---|---|---|---|
| Phenotype | 0 | 30 | 35 | 40 | 45 | 50 | Total |
| FASD area | | | | | | | |
| Phenotype scores | 0 | >30 | >25 | >21 | >17 | >13 | |
| FAS, *n* (%) | | 4 (2.7) | 1 (0.7) | 28 (18.9) | 1 (0.7) | 114 (77.0) | 148 (97.4) |
| Other-FASD, *n* (%) | | 19 (18.1) | 0 | 34 (32.4) | 0 | 52 (49.5) | 105 |
| No-FASD area | | | | | | | |
| Phenotype scores | ≤120 | ≤30 | ≤25 | ≤21 | ≤17 | ≤13 | |
| No-FASD, *n* (%) | 182 (54.0) | 94 (27.9) | 0 | 40 (11.7) | 0 | 21 (6.2) | 337 (99.4) |
| Other-FASD, *n* (%) | 5 (8.1) | 27 (43.5) | 0 | 0 (32.3) | 0 | 10 (16.1) | 62 |

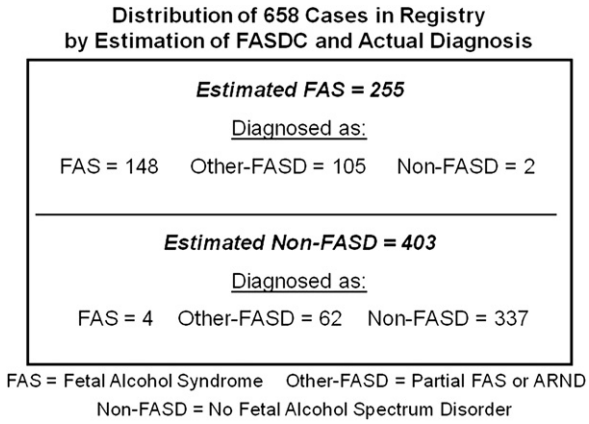


Fig. 2. Table demonstrating the development of the groups in this article.

range if they were seen in earlier years ($P < .001$); were older ($P = .040$); were Native American ($P < .001$); were in foster care or adopted ($P < .001$); had many siblings ($P = .024$); had ADHD ($P = .002$), self-injury ($P = .012$), and alcohol problems ($P = .011$); did not have another diagnosed developmental disorder ($P = .011$); or had neurological problems ($P = .028$); if their parents were not married ($P < .001$); or if their mothers were currently in treatment ($P = .034$).

The variables listed above were entered into a logistic regression model again without the alcohol exposure or the phenotype scores from the FASDC to find the best set for predicting other-FASDC who were estimated-No-FASDC ($n = 62$) compared with those with a diagnosis of No-FASDC ($n = 337$) (Table 6). Five variables, race, ADHD, self-injury, alcohol problems, and calendar year seen, were significant when entered together. The model parameters were accuracy 78.4%, sensitivity 64.5%, and specificity 81.0%. When the FASDC alcohol exposure and phenotype scores were added into the model, the model parameters for accuracy dropped from 78.4 to 74.6% (−3.8%), sensitivity increased from 64.5 to 80.7% (+16.2%), and specificity decreased from 81.0 to 73.5% (−7.5%).

Table 3
Risk factors and environmental variables that significantly differentiate between 105 other-FASD cases who were estimated-FASD and 148 FAS cases

| | FAS | Estimated-FASD | | |
|---|---|---|---|---|
| Variable | *n* (%) | *n* (%) | *P* | RR |
| Year diagnosed | | | | |
| 1978–1991 | 39 (26.35) | 19 (18.10) | .047 | |
| 1992–1994 | 56 (37.84) | 39 (37.14) | | |
| 1995–1997 | 33 (22.30) | 19 (18.10) | | |
| 1998–2003 | 20 (13.51) | 28 (26.67) | | |
| Other diagnoses | | | | |
| Neurological | 9 (6.08) | 1 (0.95) | .035 | 0.23 |
| Anger disorder | 19 (12.84) | 26 (24.76) | .023 | 1.52 |
| Other medications | 6 (4.05) | 13 (12.38) | .26 | 1.74 |
| Gestation | | | | |
| 36 or fewer wk | 31 (28.44) | 9 (11.39) | .005 | |
| 37–39 wk | 27 (24.77) | 18 (22.78) | | |
| 40 wk | 39 (35.78) | 47 (59.49) | | |
| >40 wk | 12 (11.01) | 5 (6.33) | | |
| Mean | 37.95 | 38.68 | .58 | |
| Mother's age | | | | |
| 15–20 | 15 (11.81) | 21 (21.65) | .067 | |
| 21–25 | 25 (19.69) | 23 (23.71) | | |
| 26–30 | 44 (34.65) | 33 (34.02) | | |
| 31–50 | 43 (33.86) | 20 (20.62) | | |
| Mean | 28.06 | 25.79 | .003 | |
| Current smoker | 33 (22.30) | 3 (2.86) | <.001 | 0.18 |

## Discussion

This study examined the performance characteristics of the FASDC and determined the validity of this diagnostic test by comparison against clinical expert diagnosis and the IOM criteria. The FASDC had excellent performance in distinguishing between FAS and the No-FASD group with accuracy 99%, sensitivity 99%, and specificity 99%. For these two groups, the FASDC score was an unambiguous method for separating them demonstrated by the post hoc estimated regression line in Fig. 1. The differential diagnosis of the other-FASD category was far more complex. When formulating this study, we felt that the other-FASD cases were likely primarily "borderline" cases, that is, close to the cutoff between FASD and No-FASD. This would suggest that other-FASD cases have a less distinctive physical phenotype than FAS cases and broader neuropsychiatric phenotypes than No-FASD cases. To test the possibility that phenotype influenced diagnosis more than the other factors described above, additional univariate and logistic analyses were performed.

In Fig. 3, we present data on the number of subjects within 5, 10, or 15 FASDC points from the regression separation line in Fig. 1. This figure demonstrates that the subjects with other-FASD are not primarily those with "borderline" scores but rather represent a group with wide variation in their severity scores. In Fig. 4, we present a graphic of the factors that were significant in the logistic prediction models comparing other-FASD with FAS and No-FASD.

### *FAS and other-FASD*

This analysis indicated that the variable (mother smokes) was a key factor in distinguishing between FAS and a diagnosis of other-FASD (80% correctly classified). The other-FASD cases were associated with nonsmoking mothers. Why is smoking status a risk factor for a diagnosis of FAS as compared with other-FASD? There is a high correlation between cigarette smoking and alcohol abuse. It may be that nonsmoking women consume less alcohol than smoking women. Alternatively, concurrent cigarette smoking and alcohol consumption in pregnancy may produce a synergistic teratogenic effect (Odendaal et al., 2009).

Table 4
Logistic models of significant predictors of 105 other-FASD children who were estimated-FASD compared with the 148 with FAS diagnoses

| Predictor | Coefficient | *P* | OR | Accuracy | Sensitivity | Specificity |
|---|---|---|---|---|---|---|
| Risk/environmental factors | | | | | | |
| Intercept | 0.892 | | | 60.1 | 81.0 | 45.3 |
| Smoking | −2.293 | <.001 | 0.101 | | | |
| Year seen = 1 | −1.102 | .009 | 0.332 | | | |
| Year seen = 2 | −0.661 | .079 | 0.516 | | | |
| Year seen = 3 | −0.842 | .050 | 0.431 | | | |
| All factors | | | | | | |
| Intercept | 12.756 | | | 81.8 | 84.8 | 79.7 |
| Alcohol score | −0.183 | <.001 | 0.833 | | | |
| Phenotype score | −0.124 | <.001 | 0.883 | | | |
| Smoking | −2.417 | <.001 | 0.089 | | | |

Year seen: 1 = 1 and 4 from Table 2; 2 = 2 and 4 from Table 2; and 3 = 3 and 4 from Table 2.

The influence of calendar year of diagnosis may indicate a trend toward reduced severity of FAS or of an increase in awareness of other-FASD (a broader phenotype). The comorbid conditions (neurological disorder, anger, and using medications) were also strongly related to a diagnosis of FAS. For the other-FASD group, the additional variables may be a useful strategy to improve the categorical diagnosis of subjects with decreased severity of their physical phenotype. Older mothers were more likely to be heavier drinkers (higher FASDC exposure scores) and to have other children (some with FAS or related conditions). Children with younger mothers were more likely to receive a diagnosis of other-FASD rather than FAS.

Table 5
Summary table of risk factors and environmental variables that significantly differentiate between 62 other-FASD cases who were estimated-No-FASD and 337 No-FASD cases

| | FAS | Estimated-FASD | | |
|---|---|---|---|---|
| Variable | *n* (%) | *n* (%) | *P* | RR |
| Year diagnosed | | | | |
| 1978–1991 | 91 (27.41) | 9 (14.52) | <.001 | |
| 1992–1994 | 60 (18.07) | 27 (43.55) | | |
| 1995–1997 | 74 (22.29) | 15 (24.19) | | |
| 1998–2003 | 107 (32.23) | 11 (17.74) | | |
| Age diagnosed | | | | |
| Infant to 3 | 71 (22.90) | 5 (8.20) | .040 | |
| 4–6 | 86 (27.74) | 19 (31.15) | | |
| 7–10 | 86 (27.74) | 17 (27.87) | | |
| 11–56 | 67 (21.61) | 20 (32.79) | | |
| Mean | 7.55 | 8.82 | .046 | |
| Race | 138 (40.95) | 44 (70.97) | <.001 | 2.91 |
| Adopted or foster care | 116 (34.42) | 39 (62.90) | <.001 | 2.67 |
| Parity | | | | |
| 0 | 119 (37.30) | 16 (27.59) | .553 | |
| 1 | 73 (22.88) | 13 (22.41) | | |
| 2 | 57 (17.87) | 12 (20.69) | | |
| 3 | 34 (10.66) | 7 (12.07) | | |
| 4–10 | 36 (11.29) | 10 (17.24) | .024 | |
| Mean | 1.4 | 1.8 | .024 | |
| Other diagnoses | | | | |
| ADHD | 149 (44.21) | 41 (66.13) | .002 | 2.15 |
| Developmental disorder | 99 (29.38) | 8 (12.90) | .011 | 0.40 |
| Neurological | 48 (14.24) | 2 (3.23) | .028 | 0.23 |
| Self-injury | 11 (3.26) | 7 (11.29) | .012 | 2.69 |
| Alcohol problem | 5 (1.48) | 5 (8.06) | .011 | 3.41 |
| Parents | | | | |
| Mother married | 141 (45.19) | 11 (20.00) | <.001 | 0.35 |
| Father married | 136 (47.72) | 12 (26.09) | .010 | 0.44 |
| Mother in treatment | 13 (34.21) | 5 (83.33) | .034 | 7.22 |

ADHD = attention-deficit hyperactivity disorder.

As the logistic regression models demonstrated, the FASDC scores were helpful in discrimination of other-FASD from FAS. The percent correctly classified was increased by 22% when the FASDC phenotype and exposure scores were included in the model. Smoking was the only other variable retained in the logistic model. The ability to detect other-FASD cases (sensitivity) was more than 80% for both models. But adding the FASDC exposure and phenotype scores increased the ability to detect FAS cases (specificity) by 34%. When FASDC scores are above the regression separation line in Fig. 1, it may be helpful to consider both the additional factors and the FASDC scores in the formulation of a diagnosis.

### *No-FASD and other-FASD*

We found a trend for a decrease in the number of other-FASD subjects with low FASDC scores. The recognition of less complete phenotypes may indicate an improvement in

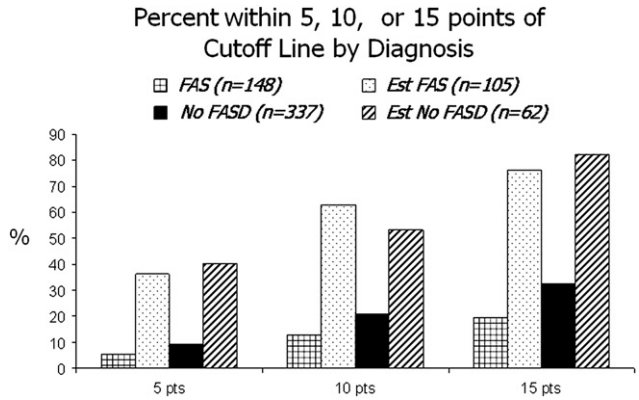


Fig. 3. The percent of subjects in the four groups: FAS, No-FASD, subjects with a diagnosis of other-FASD who were estimated-FASD, and subjects with an other-FASD diagnosis who were estimated-No-FASD.

Table 6
Logistic models of significant predictors of other-FASD who were estimated-No-FASD ($n = 62$) versus those with No-FASD diagnoses ($n = 337$)

| Predictor | Coefficient | P | OR | Accuracy | Sensitivity | Specificity |
|---|---|---|---|---|---|---|
| Risk/environmental factors | | | | | | |
| Intercept | −2.468 | | | 78.4 | 64.5 | 81.0 |
| Race | 1.404 | <.001 | 4.072 | | | |
| ADHD | 0.970 | .002 | 2.639 | | | |
| Self-injury | 1.446 | .018 | 4.247 | | | |
| Alcohol | 1.692 | .037 | 5.430 | | | |
| Year seen = 1 | 0.887 | .095 | 2.428 | | | |
| Year seen = 2 | 1.927 | <.001 | 6.867 | | | |
| Year seen = 3 | 0.911 | .046 | 2.487 | | | |
| All factors | | | | | | |
| Intercept | −5.872 | | | 74.6 | 80.7 | 73.5 |
| Phenotype score | 0.069 | .002 | 1.071 | | | |
| Alcohol score | 0.093 | <.001 | 1.098 | | | |
| Race | 0.811 | .034 | 2.250 | | | |
| ADHD | 1.246 | <.001 | 3.475 | | | |
| Year seen = 1 | 1.598 | .006 | 4.945 | | | |
| Year seen = 2 | 1.974 | <.001 | 7.200 | | | |
| Year seen = 3 | 0.620 | .182 | 1.859 | | | |

ADHD = attention-deficit hyperactivity disorder.

the identification of adverse effects from prenatal alcohol exposure other than FAS (the partial FAS or ARND groups). Alternatively, the increased prevalence of older subjects with a diagnosis of other-FASD may reflect the difficulty in making the diagnosis in older subjects. We have demonstrated this problem in longitudinal follow-up of the growth parameters height and weight for a large cohort (Burd et al., 2003b). As subjects age, their growth impairment tends to decrease. Also, the older the subject, the less likely they are to be with their mothers, and the exposure data available for diagnosis are more difficult to obtain. We have also found FAS to be progressively more severe in the younger siblings (Burd et al., 2003a). These results imply that both phenotype and other factors are important in the clinical diagnoses of other FASD, even when the FASDC phenotype scores are near the average. This is also an important clinical issue in the assessment of infants and young children in whom the neuropsychiatric manifestations of brain damage/dysfunction are difficult to evaluate with confidence especially in the presence of multiple other postnatal environmental influences including abuse, neglect, and poor nutrition and other environmental adversities such as multiple foster home placements.

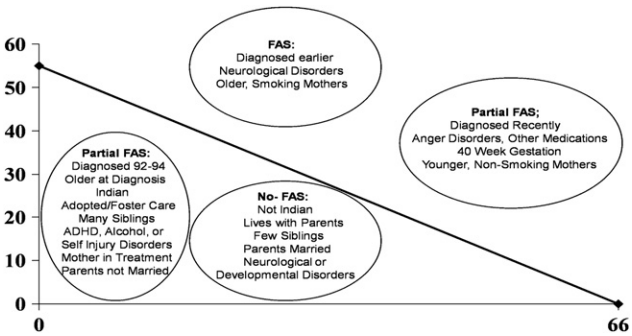


Fig. 4. The variables associated with diagnosis of FAS, other-FASD estimated-FASD, other-FASD estimated-No-FASD, and No-FASD relative to FASDC scores.

The logistic regression models demonstrate that although the FASDC scores were significant in predicting other-FASD diagnoses for subjects with low scores, they produced a variable effect on the overall correct classification of cases. The percent correctly classified dropped from 78 to 75%, although the ability to correctly identify FAS cases (sensitivity) rose 17%. This suggests that the diagnosis of cases with low FASDC scores may be influenced by other factors (i.e., a diagnosis of ADHD). The decrease in the overall FASDC severity score may be influenced by more older subjects entering foster care or adoption and having poorer exposure data (no biological mother to interview) or may reflect decreased distinctiveness of the classic FAS phenotype triad in older subjects. Lastly, does the FASDC score measure severity or the extensiveness of the phenotype based on symptom frequency, which may be more a measure of the range of factors included in the diagnostic assessment? In future studies using an independent severity measure for other behavioral disorders or increased psychosocial impairments to compare with the FASDC may be useful. This may be important to improve our understanding of factors that explain both a wider range of impairments or increase severity of impairments and would likely be very useful in counseling families, caretakers, and service providers.

### *Limitations*

Of the 62 cases with a diagnosis of other-FASD in the No-FASD range, 5 of them did not have information on exposure. The only distinguishing characteristics about them were that four of them were diagnosed from 1992 to 1994, four of them were male and Native Americans,

four of them were in foster care or adopted, and none of the mothers smoked. The problem of classification of subjects with an FAS phenotype or partial phenotype with missing data on exposure is a common problem in the diagnosis of FASD. The primary problem is that a majority of patients are not accompanied by their mothers to the evaluation and that the medical records of the patient or the medical records of their mothers may be uninformative as a source of exposure data (Burd et al., 2006).

The data from this study are from a small state that is largely composed of rural and frontier populations. The assessments were provided by a single group of clinicians, and as a result where error is present, it is likely to be systematic. The assessments were conducted over a long time period and changes in our diagnostic formulation over time are difficult to identify and may result in changes over time that could affect diagnosis especially for cases of other-FASD.

It is important to emphasize that assessment of exposure is nearly always a challenge especially in a population of children in which the mother's self-report of exposure is often not available. As a result, considerable variance in assessment of exposure may be present. It is likely that we may have not have diagnosed some children with an FASD due to the absence of exposure data where exposure did occur and they may have had an FASD and would have been given an FASD diagnosis if this information had been available.

However, some of these factors may increase the value of this study. We have been able to include a large number of subjects from a state with access to diagnostic services for all people across the state including high-risk populations. We have made considerable efforts to include populations for which access to services is difficult due to distance and extreme weather conditions lasting several months each year, and we have provided services on or near each reservation community. We have made ongoing efforts to include health care staff and other relevant referral sources from the four Indian nations in North Dakota over more than 20 consecutive years.

This study also describes some of the variables that may alter the likelihood of a diagnosis of other-FASD, which was first modeled without inclusion of the FASDC phenotype and alcohol exposure score in the prediction model (Fig. 4). It could be argued that FASD is prevalent in many of these cohorts, for example, Native Americans, smokers, and disruptive homes; so it is correct to make diagnoses based on these factors. However, it is also possible that diagnosis of other-FASD was made based, in part, on the presence of these factors, with the result being an increased prevalence of other-FASD. This would have the undesirable effect of increasing the prevalence of a diagnosis of other-FASD in the group of subjects with an increased prevalence of these variables. However, the presence of statistically significant differences in the FASDC phenotype and exposure scores for the majority of other-FASD subjects when compared with the No-FAS group argued against this concept. The presence of numerous statistically significant environmental and parental risk markers does offer support for clinically significant phenotypic differences between FAS, other-FASD, and No-FAS groups.

The associated conditions with higher prevalence among patients with FAS and other-FASD included ADHD, other neurological problems, alcohol abuse, and so on. In relation to FAS and other-FASD, these conditions could be viewed as (1) additional candidate diagnostic criteria for an expanded FAS phenotype; and/or (2) frequent comorbid disorders that have minimal effect on severity but rather reflect the broad range of adverse outcomes from exposure; and/or (3) risk factors for FAS and other-FASD as defined by the FASDC. In the article, we present them as risk factors for FAS and other-FASD as defined by the FASDC. This concept may have validity based on the analytic outcomes from the data. From an etiologic point of view, these conditions may represent pleiotropic endpoints with FAS from the common variable of prenatal alcohol exposure. Additional research will be required to address these conceptual issues.

The option of rigidly adhering to a test score as the only important variable in the formulation of a diagnosis seems problematic. Clinicians often see classic test results in false-positive subjects without the disease or phenotype. Alternatively, inconsistent clinical diagnosis of subjects with classic exposure and phenotype scores suggests a need for improvement in the consistency of diagnostic inquiry. Although both alternatives are undesirable, the solution is far from apparent, because this problem is identified in a setting where the FASDC score was used to provide a standardized examination to produce both reliable and valid exposure and phenotype data. The problem may be more problematic in clinical or research settings where structured diagnostic tools, such as the FASDC, are not used.

Additional research will be required to develop a consistent diagnostic system for the difficult cases where clinical impressions differ from the FASDC scores. However, this is not uncommon in developmental disorders where the clinical dilemma of how to classify atypical, mild, and very severe cases is difficult (Burd and Kerbeshian, 1988). The development of biologic markers may be helpful, but the development of markers will also be dependent on the diagnostic grouping of subjects used to validate the biologic markers (Burd, 2008; Burd and Hofer, 2008; Burd et al., 2001). One needs to be alert to the error of misconstruing a dependent variable as an independent variable. The large cohort size required for development of diagnostic tools for other diagnostic entities suggests that this may be an expensive and time-consuming process. Further study of the FASDC in diverse populations will be helpful in further development of the database on the specificity and sensitivity of the FASDC.

Last, the boundaries of the phenotype from prenatal alcohol exposure have yet to be precisely determined.

The problem is important, and additional detailed studies of diagnostic schema are needed. Adverse outcomes associated with or resulting from prenatal alcohol exposure may be common and would place prenatal alcohol exposure as a common cause of highly variable adverse outcomes. Thus, comparisons of differing diagnostic schema in populations of shared subjects are urgently needed. Alternatively, prenatal alcohol exposure may be a common occurrence in people with developmental disorders but have a minimal causal role in the prevalence or severity of these disorders and phenotypes. Much work remains to be done.

## References


Abel, E. L. (1998). *Fetal Alcohol Abuse Syndrome*. New York, NY: Plenum Press.

Astley, S. J. (2006). Comparison of the 4-digit diagnostic code and the Hoyme diagnostic guidelines for fetal alcohol spectrum disorders. Pediatrics *118*, 1532–1545.

Bertrand, J., Floyd, R. L., Weber, M. K., O'Connor, M., Riley, E. P., Johnson, K. A., et al. (2004). *National Task Force on FAS/FAE. Fetal Alcohol Syndrome: Guidelines for Referral and Diagnosis*. Atlanta, GA: Centers for Disease Control and Prevention.

Burd, L. (2008). Response to letter to the editor. Birth Defects Res. A Clin. Mol. Teratol. (Part A) *85*, 231–232.

Burd, L., Cotsonas-Hassler, T., Martsolf, J., and Kerbeshian, J. (2003a). Recognition and management of fetal alcohol syndrome. Neurotoxicol. Teratol. *25*, 681–688.

Burd, L., and Hofer, R. (2008). Biomarkers for detection of prenatal alcohol exposure: a critical review of fatty acid ethyl esters in meconium. Birth Defects Res. A Clin. Mol. Teratol. *82*, 487–493.

Burd, L., and Kerbeshian, J. (1988). Infantile autism: psychosocial and neurodevelopmental factors. J. Am. Acad. Child Adolesc. Psychiatry *27*, 252–253.

Burd, L., Kerbeshian, J., and Klug, M. G. (2001). Neuropsychiatric genetics: misclassification in linkage studies of phenotype-genotype research. J. Child Neurol. *16*, 499–504.

Burd, L., Klug, M. G., Martsolf, J., and Ebertowski, M. (2003b). Body mass index in fetal alcohol syndrome. Neurotoxicol. Teratol. *25*, 689–696.

Burd, L., Klug, M. G., Martsolf, J. T., Martsolf, C., Deal, E., and Kerbeshian, J. (2006). A staged screening strategy for prenatal alcohol exposure and maternal risk stratification. J.R. Soc. Promot. Health *126*, 86–94.

Burd, L., and Martsolf, J. T. (1989). Fetal alcohol syndrome: diagnosis and syndromal variability. Physiol. Behav. *46*, 39–43.

Burd, L., and Moffatt, M. E. (1994). Epidemiology of fetal alcohol syndrome in American Indians, Alaskan Natives, and Canadian Aboriginal peoples: a review of the literature. Public Health Rep. *109*, 688–693.

Burd, L., Martsolf, J., and Kerbeshian, J. (2003c). Diagnosis of FAS: a comparison of the Fetal Alcohol Syndrome Diagnostic Checklist and the Institute of Medicine Criteria for Fetal Alcohol Syndrome. Neurotoxicol. Teratol. *25*, 719–724.

Bush, K., Kivlahan, D. R., McDonell, M. B., Fihn, S. D., and Bradley, K. A. (1998). The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Arch. Intern. Med. *158*, 1789–1795.

Floyd, R. L., and Sidhu, J. S. (2004). Monitoring prenatal alcohol exposure. Am. J. Med. Genet. C. Semin. Med. Genet. *127C*, 3–9.

Greenhalgh, T. (1997). How to read a paper. Papers that report diagnostic or screening tests. BMJ *315*, 540–543.

Hoyme, H. E., May, P. A., Kalberg, W. O., Kodituwakku, P., Gossage, J. P., Trujillo, P. M., et al. (2005). A practical clinical approach to diagnosis of fetal alcohol spectrum disorders: clarification of the 1996 Institute of Medicine criteria. Pediatrics *115*, 39–47.

Odendaal, H. J., Steyn, D. W., Elliott, A., and Burd, L. (2009). Combined effects of cigarette smoking and alcohol consumption on perinatal outcome. Gynecol. Obstet. Invest. *67*, 1–8.

Sampson, P. D., Streissguth, A. P., Bookstein, F. L., Little, R. E., Clarren, S. K., Dehaene, P., et al. (1997). Incidence of fetal alcohol syndrome and prevalence of alcohol-related neurodevelopmental disorder. Teratology *56*, 317–326.

Stratton, K. R., Howe, C. J., Battaglia, F. C., and Institute of Medicine (1996). *Fetal Alcohol Syndrome-Diagnosis, Epidemiology, Prevention, and Treatment*. Washington, DC: National Academy Press.