A Decision Tree to Identify Children Affected by Prenatal Alcohol Exposure

Patrick K. Goh, BS[1], Lauren R. Doyle, BS[1], Leila Glass, MS[1], Kenneth L. Jones, MD[2], Edward P. Riley, PhD[1], Claire D. Coles, PhD[3,4], H. Eugene Hoyme, MD[5], Julie A. Kable, PhD[4], Philip A. May, PhD[6,7], Wendy O. Kalberg, PhD[7], Elizabeth, R. Sowell, PhD[8,9], Jeffrey R. Wozniak, PhD[10], and Sarah N. Mattson, PhD[1]

Objective To develop and validate a hierarchical decision tree model that combines neurobehavioral and physical measures to identify children affected by prenatal alcohol exposure even when facial dysmorphology is not present.

Study design Data were collected as part of a multisite study across the US. The model was developed after we evaluated more than 1000 neurobehavioral and dysmorphology variables collected from 434 children (8-16 years of age) with prenatal alcohol exposure, with and without fetal alcohol syndrome, and nonexposed control subjects, with and without other clinically-relevant behavioral or cognitive concerns. The model subsequently was validated in an independent sample of 454 children in 2 age ranges (5-7 years or 10-16 years). In all analyses, the discriminatory ability of each model step was tested with logistic regression. Classification accuracies and positive and negative predictive values were calculated.

Results The model consisted of variables from 4 measures (2 parent questionnaires, an IQ score, and a physical examination). Overall accuracy rates for both the development and validation samples met or exceeded our goal of 80% overall accuracy.

Conclusions The decision tree model distinguished children affected by prenatal alcohol exposure from nonexposed control subjects, including those with other behavioral concerns or conditions. Improving identification of this population will streamline access to clinical services, including multidisciplinary evaluation and treatment. *(J Pediatr 2016;177:121-7).*

Fetal alcohol spectrum disorders (FASD) are characterized by a range of physical and neurobehavioral changes caused by prenatal alcohol exposure[1,2] and occur in as many as 4.8% of school-age children.[3] Although fetal alcohol syndrome (FAS) is readily identifiable by trained clinicians,[4] the majority of children affected by prenatal alcohol exposure present with significant neurobehavioral impairments without sufficient dysmorphology to merit the diagnosis of FAS. This heterogeneity of physical presentation coupled with an overlap in clinical presentation with other neurodevelopmental disorders, such as attention-deficit/hyperactivity disorder (ADHD), results in underidentification, misdiagnosis, and misunderstanding of the etiology of difficulties faced by children with FASD.[2,5] Failure to identify alcohol-affected children can adversely affect prognosis and access to tailored services.[6]

Although efforts to identify affected individuals have been moderately successful,[7-10] they have not translated easily to clinical settings, perhaps because of the difficulty in differentiating children affected by alcohol from those with similar behavioral concerns without prenatal alcohol exposure and the extensive neuropsychological testing required. Thus, there is a need for an efficient but specific primary screening tool for use by pediatricians to identify children with neurobehavioral impairments that may be attributable to prenatal alcohol exposure. Positive screens would compel referral for additional medical or developmental evaluations or to an FASD clinic for full evaluation. We aimed to develop and validate a clinically relevant, feasible, and accurate (>80%)[11,12] decision tree

ADHD	Attention-deficit/hyperactivity disorder
AE	Alcohol-exposed
CIFASD	Collaborative Initiative on Fetal Alcohol Spectrum Disorders
FAS	Fetal alcohol syndrome
FASD	Fetal alcohol spectrum disorders
FSIQ	Full-scale intelligence quotient
GCA	General conceptual ability
M	Mean
Non-AE	Non-alcohol-exposed

From the [1]Center for Behavioral Teratology, Department of Psychology, San Diego State University, San Diego, CA; [2]Department of Pediatrics, University of California, San Diego, School of Medicine, San Diego, CA; [3]Department of Psychiatry and Behavior Sciences, Emory University School of Medicine, Atlanta, GA; [4]Department of Pediatrics, Emory University School of Medicine, Atlanta, GA; [5]Sanford Research and Sanford School of Medicine of the University of South Dakota, Sioux Falls, SD; [6]Department of Nutrition, Gillings School of Global Public Health, University of North Carolina Nutrition Research Institute, Kannapolis, NC; [7]Center on Alcoholism, Substance Abuse, and Addictions, The University of New Mexico, Albuquerque, NM; [8]Developmental Cognitive Neuroimaging Laboratory, Department of Pediatrics, Keck School of Medicine, University of Southern California, Los Angeles, CA; [9]Division of Research on Children, Youth, and Families, Department of Pediatrics, Children's Hospital Los Angeles, Los Angeles, CA; and [10]Department of Psychiatry, University of Minnesota, Minneapolis, MN

Supported by the Collaborative Initiative on Fetal Alcohol Spectrum Disorders, which is funded by National Institute on Alcohol and Alcohol Abuse (U01 AA014834, U01 AA014815, U01 AA014811, F31 AA022261, and T32 AA013525). Additional information about CIFASD can be found at www.cifasd.org. The authors declare no conflicts of interest.

0022-3476/$ - see front matter. © 2016 Elsevier Inc. All rights reserved.

http://dx.doi.org10.1016/j.jpeds.2016.06.047

Table I. Demographic information for AE and non-AE groups in child (5-7 y) and adolescent (10-16 y) samples

	CIFASD II*†‡§		CIFASD III (child)*		CIFASD III (adolescent)*¶	
	AE	Non-AE	AE	Non-AE	AE	Non-AE
	n = 146	n = 288	n = 55	n = 110	n = 98	n = 191
Age, M (SD)	12.5 (2.4)	12.2 (2.6)	6.8 (0.9)	6.6 (0.9)	13.0 (2.0)	13.7 (2.0)
Sex, n (% female)	63 (43.2)	109 (37.8)	32 (58.2)	48 (43.6)	43 (43.9)	90 (47.1)
FSIQ/GCA, M (SD)	83.9 (16.8)	100.2 (17.5)	86.8 (13.4)	99.1 (13.4)	88.1 (12.5)	101.2 (16.4)
Race, n (%white)	78 (53.4)	186 (64.6)	24 (43.6)	59 (53.6)	53 (54.1)	103 (53.9)
Ethnicity, n (% Hispanic)	19 (13.0)	67 (23.3)	6 (10.9)	14 (12.7)	16 (16.3)	42 (22.0)
Handedness, n (% right)	128 (87.7)	266 (92.4)	48 (87.3)	103 (93.6)	89 (90.8)	168 (88.0)
Site, n (%)						
San Diego	55 (37.7)	124 (43.1)	11 (20.0)	31 (28.2)	32 (32.7)	66 (34.6)
Atlanta	30 (20.5)	52 (18.1)	22 (40.0)	34 (30.9)	21 (21.4)	55 (28.8)
Los Angeles	28 (19.2)	30 (10.4)	0 (0.0)	1 (0.9)	13 (13.3)	20 (10.5)
Northern Plains	22 (15.1)	34 (11.8)	-	-	-	-
Albuquerque	11 (7.5)	48 (16.7)	-	-	-	-
Minneapolis	-	-	22 (40.0)	44 (40.0)	32 (32.7)	50 (26.2)

*Significant differences in FSIQ/GCA between AE and non-AE groups.
†Significant differences in race between AE and non-AE groups.
‡Significant differences in ethnicity between AE and non-AE groups.
§Significant differences in distribution across site between AE and non-AE groups.
¶Significant differences in age between AE and non-AE groups.

model that could differentiate children affected by prenatal alcohol exposure from nonexposed children with and without other clinically relevant behavioral or cognitive concerns or conditions.

Methods

Data were collected as part of the Collaborative Initiative on Fetal Alcohol Spectrum Disorders (CIFASD), a multisite, multimethod study of FASD. Data for model development and model validation samples were collected in the second (CIFASD II; 2007-2012) and third (CIFASD III, ongoing) phases of the study, respectively. Both phases included standardized neuropsychological assessment and dysmorphology evaluations.[13] Many of the same measures were included in both phases, although different assessments of general cognitive ability were used (CIFASD II: Wechsler Intelligence Scale for Children-Fourth Edition;[14] CIFASD III: Differential Ability Scales-Second Edition[15]). Both phases included a heterogeneous comparison group (non-alcohol-exposed [non-AE] group) comprising nonexposed (<1 drink/week, never >2 drinks/occasion) typically developing children and nonexposed children with clinical conditions or concerns (approximately 30%-50% of comparison subjects). The most common concerns in the comparison group were related to ADHD, other internalizing and externalizing problems, low IQ, and learning disorders, which are relatively representative of the populations of an outpatient pediatric psychology or psychiatry clinic.

Prenatal alcohol exposure histories were obtained through retrospective maternal report or social service, legal, or medical records. Subjects were included in the alcohol-exposed (AE) group if a history of heavy prenatal alcohol exposure (>13 drinks/week or >4 drinks/occasion during pregnancy) was known or when such exposure was suspected in a child with a diagnosis of FAS. In many cases, an accurate exposure history was unattainable; children were included in the AE group if mothers were known to be "alcoholic" or abusing alcohol during pregnancy. All subjects were evaluated with a standardized dysmorphology examination conducted by a member of the CIFASD Dysmorphology Core.[4,13] Informed consent and assent were obtained from caregivers and subjects before testing. Financial incentive for participation and written feedback was provided after evaluation. The Institutional Review Boards at San Diego State University and other CIFASD sites approved this study. Demographic information is presented in Table I.

Model Development Sample

Data from 434 subjects aged 8-16 years (mean [M] = 12.3) were used in the development of the decision tree model. Thirty-nine (26.7%) subjects in the AE group met criteria for FAS.[4,13] One hundred forty-seven (51.0%) subjects in the non-AE group had clinically concerning behaviors or conditions, including 93 (32.3%) subjects who met criteria for ADHD based on the Computerized Diagnostic Interview Schedule for Children - Fourth Edition.[16-18]

Model Validation Sample

Data from 454 subjects aged either 5-7 years (child; M = 6.6, n = 165) or 10-16 years (adolescent; M = 13.4, n = 289) were used for model validation. Seven (12.7%) subjects in the child age range and 11 (11.2%) subjects in the adolescent age range met criteria for FAS. Forty-four (40.0%) subjects in the child age range and 73 (38.2%) subjects in the adolescent age range had clinically concerning behaviors or conditions, including 33 (30.0%) children and 38 (19.9%) adolescents who met criteria for ADHD on the basis of the Computerized Diagnostic Interview Schedule for Children - Fourth Edition.

Table II. Variables and cut-off criteria for each measure included in final decision tree

Decision points	Description	Selected variables	Cut-off score for decision point
Child Behavior Checklist (CBCL)[20]	Parent-reported problem behaviors	Thought Problems, Attention Problems, Social Problems, Aggressive Behavior, Rule-Breaking Behavior, Somatic Complaints	>0 Subscales with T-scores >65
IQ test (WISC-IV/DAS-II)[14,15]	Direct measures of general cognitive ability	FSIQ or GCA	FSIQ/GCA <92
Physical exam for key features of FAS	Physical measurements from dysmorphology exam consistent with a diagnosis of FAS[21-23]	Palpebral fissure length ≤10 percentile, vermilion border lipometer score = 4 or 5, philtrum lipometer score = 4 or 5, head circumference ≤10th percentile, height and/or weight ≤10th percentile	Meeting criteria for FAS* or >0 key facial features
Physical exam for extended features	Physical measurements from dysmorphology exam included in the expanded range of effects of prenatal alcohol exposure[21]	Ptosis, incomplete extension of digits	>0 criteria
VABS [1] and VABS [2]†[24]	Parent-reported adaptive functioning	Socialization, Communication, Daily Living Skills	>1 Domain with standard scores <86

DAS-II, Differential Ability Scales - Second Edition; *VABS*, Vineland Adaptive Behavior Scales-II; *WISC-IV*, Wechsler Intelligence Scale for Children-Fourth Edition.
*FAS diagnoses requires at least 2 of 3 key facial features (palpebral fissure length ≤10th percentile; philtrum lipometer score = 4 or 5; vermilion border lipometer score = 4 or 5), and presence of head circumference ≤10th percentile or height and/or weight ≤10th percentile.[21-23]
†VABS (1) and VABS (2) refer to the same variables from the Vineland Adaptive Behavior Scale but are used at 2 different decision points in the decision tree model.

Decision Tree Development

More than 1000 composite and subtest scores from measures that assessed verbal and nonverbal memory, executive functioning, verbal comprehension, perceptual reasoning, working memory, processing speed, psychopathology, adaptive behavior, and dysmorphology were considered for inclusion in the model. For continuous variables (ie, neuropsychological measures), cut-off scores based on normed clinical levels of impairment were refined sequentially into binary scores that maximized each measure's ability to differentiate between AE and non-AE groups. Categorical variables (ie, dysmorphology measures) were retained as binary values indicating the presence or absence of a feature. In addition, composite variables that included multiple subscales from a single measure were created (eg, combining the binary indicators from multiple indexes in the Wechsler Intelligence Scale for Children-Fourth Edition into one variable). These variables were retained as binary values indicating whether criteria was met on a specific number of included subscales, with cut-offs determined through logistic regression.

Between-subject ANOVAs were conducted to obtain the effect size (partial η^2) corresponding to the group differences (AE vs non-AE) for each variable. Variables with partial $\eta^2 < .1$ were eliminated, resulting in 40 retained variables for further consideration. These variables were added sequentially to the decision tree, beginning with the largest effect sizes. Variables' discriminatory abilities were retested through logistic regression to obtain OR values and significance levels (P values) for each binary indicator at its position within the tree.[19] Variables were added to paths within the decision tree model until sample size prohibited further discrimination (<10 subjects in a single group) or no variables were found that could significantly differentiate between AE and non-AE subjects.

Once we established the initial version of the tree, we tested substitutions of other variables that discriminated between AE and non-AE groups with a medium-to-large effect size. More than 150 different orders and combinations of variables were examined to produce the most efficient and parsimonious tree while maximizing overall accuracy rates (correctly identified subjects/total subjects). Cut-off scores, ORs, and other statistical results, as well as a list of the measures and number of variables per measure that were considered for inclusion in the decision tree, are included in **Tables II-V** (**Tables III-V**; available at www.jpeds.com).

Subject group membership (AE or non-AE) was determined on enrollment into the CIFASD project on the basis of knowledge of exposure history. Comparison of actual group membership with proposed group membership enabled calculation of overall accuracy, sensitivity, specificity, positive predictive value, and negative predictive value of the final decision tree model. For all statistical tests, an alpha level of $P < .05$ indicated significance and P values between .05 and .08 indicated marginal significance. Two routes through the decision tree were tested: (1) a child who presents to a pediatrician with suspicion of prenatal alcohol exposure or with concerning physical or behavioral features (pediatrician route); and (2) a child who presents to a psychologist with concerns about neurobehavioral impairment (psychologist route) and subsequently is referred to the pediatrician on the basis of results from psychopathology and IQ tests. The proposed decision tree is presented in the **Figure**.

Results

Demographic data were analyzed by χ^2 or ANOVA. In the development sample (CIFASD II), the AE and non-AE groups differed on ethnicity ($P = .01$), race ($P = .02$), site ($P = .01$), full-scale IQ (FSIQ; $P < .001$), and rate of ADHD diagnosis ($P < .001$), but not on sex, handedness, or age ($P > .11$). In the validation sample (CIFASD III), the AE and non-AE groups in the child sample differed on general conceptual ability (GCA; $P < .001$) and rate of ADHD diagnosis ($P < .001$), but not on sex, ethnicity, race, handedness, age, or site ($P > .08$). In the adolescent sample, the AE and non-AE groups differed on age

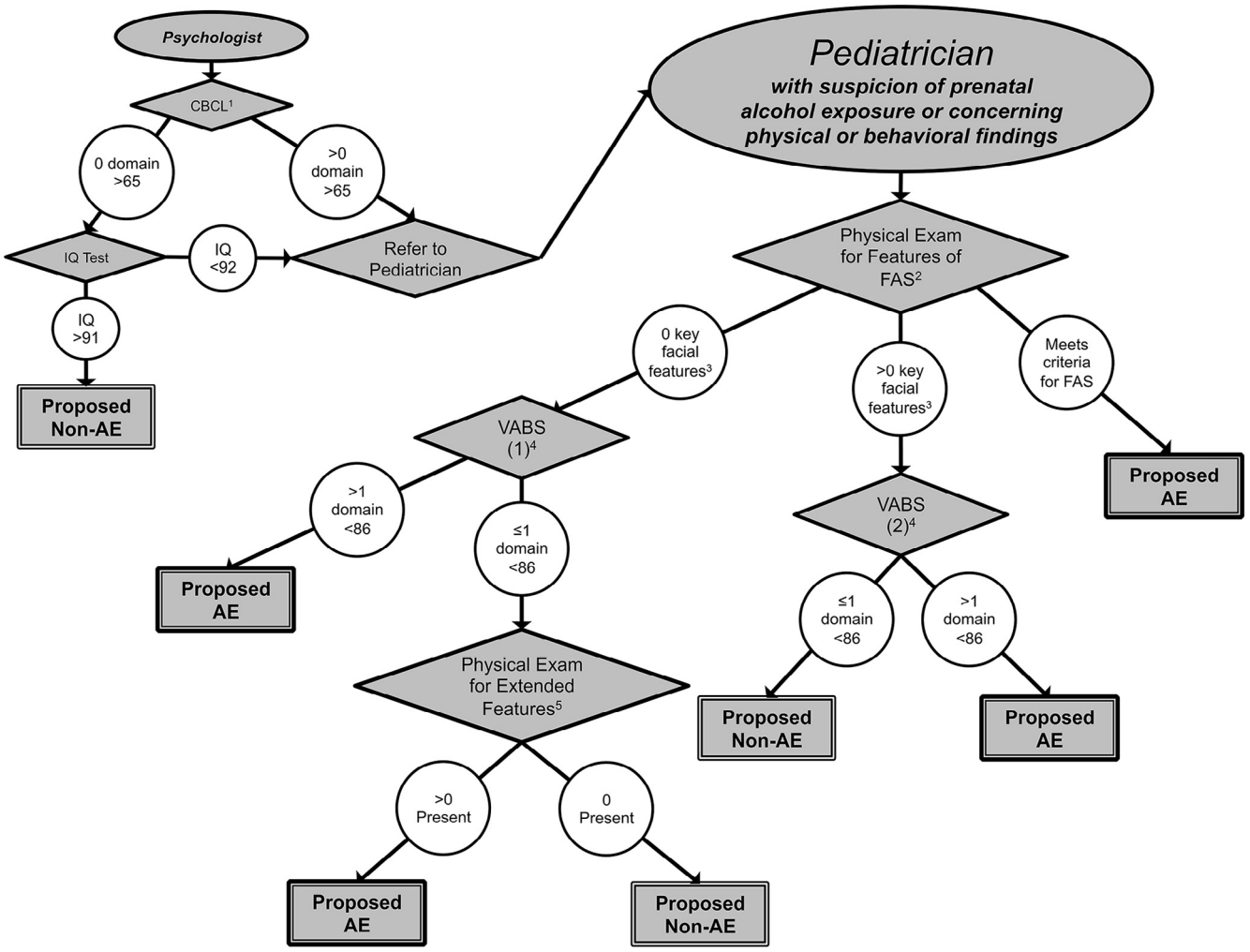

Figure. Decision tree to identify youth as AE or non-AE. Ovals, actors; diamonds, action steps; circles, criteria for action step; and boxes, ending points in the decision tree. If a child presents directly to a pediatrician with a suspicion of prenatal alcohol exposure or concerning physical or behavioral findings, the pediatrician would work through the decision tree model by assessing for dysmorphology and adaptive behavior using the Vineland Adaptive Behavior Scales - Second Edition (VABS-II). A child assessed by a psychologist would need to be referred to a pediatrician only if IQ was less than 92 or there were Child Behavior Checklist (CBCL) scores in the clinical range. [1]CBCL domains included Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior, and Aggressive Behavior. [2]FAS diagnoses require at least 2 of 3 key facial features (palpebral fissure length ≤10th percentile; philtrum lipometer score = 4 or 5; vermilion border lipometer score = 4 or 5), and either presence of head circumference ≤10th percentile or height and/or weight ≤10th percentile.[21-23] [3]If a child was not diagnosed with FAS, he or she would be assessed for the number of key facial features present. [4]VABS (1) and VABS (2) refer to the same variables from the VABS-II (Communication, Socialization, and Daily Living Skills), but used at 2 different decision points in the decision tree model. [5]Extended features are specified as ptosis and incomplete extension of one or more digits.[21]

($P = .01$), GCA ($P < .001$), and rate of ADHD diagnosis ($P < .001$), but not on sex, ethnicity, race, handedness, or site ($P > .26$).

Accuracy Rate Analyses

Overall accuracy rates across model development and validation samples were 79.5%-84.7% (**Table VI**). *Z*-test analyses were conducted to examine whether the proportion of children correctly identified by the decision tree in each sample was significantly greater than chance (50% based on the inclusion of 2 groups). In the model development sample, overall classification accuracies when the psychologist route was used ($z = 14.4$, $P < .001$, 95% CI 80.9-87.9) and pediatrician route ($z = 12.5$, $P < .001$, 95% CI 76.0-83.8) were greater than chance. Logistic regression analyses revealed that the discriminatory ability of all measures was statistically significant across routes ($P \leq .02$). In the model validation sample, overall classification accuracies when we used the psychologist route (child:

Table VI. Accuracy percentages in the decision tree model for development (CIFASD II) and validation (CIFASD III) samples

	Psychologist route			Pediatrician route		
Accuracy measures (%)	CIFASD II	CIFASD III [child]	CIFASD III [adolescent]	CIFASD II	CIFASD III [child]	CIFASD III [adolescent]
Overall accuracy	84.6	84.5	84.7	80.1	82.1	79.5
Sensitivity	74.2	70.7	79.3	79.2	63.8	81.3
Specificity	89.9	93.5	87.6	80.6	93.4	78.3
Positive predictive value	78.6	87.9	77.4	70.7	85.7	71.4
Negative predictive value	87.4	82.9	88.7	86.7	80.7	86.2

$z = 8.9$, $P < .001$, 95% CI 78.1-89.7; adolescent: $z = 11.8$, $P < .001$, 95% CI 80.0-88.7) and pediatrician route (child: $z = 8.2$, $P < .001$, 95% CI 75.4-87.6; adolescent: $z = 10.0$, $P < .001$, 95% CI 74.4-84.0) were greater than chance in both age ranges. In both age ranges, logistic regression analyses revealed that the discriminatory ability of all individual measures in the tree was statistically significant ($P \leq .05$), except for GCA ($P \geq .10$).

Misclassified Subjects

Misclassified subjects were compared with correctly classified subjects in the development and validation samples to examine potential systematic sources of error. Age and FSIQ/GCA were tested with ANOVA. Sex, handedness, race, and ethnicity were tested with the χ^2 or Fisher exact test (when n in ≥1 cells <5). Compared with correctly identified AE subjects, misclassified AE subjects (ie, exposed subjects misclassified as non-AE) were younger ($P \leq .033$), had a greater FSIQ/GCA score ($P \leq .003$), a lower rate of ADHD diagnosis ($P \leq .023$), and were less likely to be Hispanic ($P \leq .039$). Compared with correctly identified non-AE subjects, misclassified non-AE subjects (ie, non-AE subjects misclassified as AE) were older ($P \leq .019$), had a lower FSIQ/GCA score ($P \leq .012$), a greater rate of ADHD diagnosis ($P \leq .019$) and left-handedness ($P \leq .047$), and were less likely to be white ($P \leq .001$).

Post-Hoc Analyses

To follow up our a priori analyses and to improve differential diagnosis ability, we examined the overall accuracy rates for various clinically relevant scenarios, as follows: (1) limiting the nonexposed comparison group to include only subjects without known behavior/cognitive problems (psychologist route: 89.2%, pediatrician route: 83.3%); (2) limiting the nonexposed comparison group to include only subjects with known behavior/cognitive problems (psychologist route: 67.2%, pediatrician route: 75.1%); (3) limiting the exposed group by excluding subjects with FAS (psychologist route: 84.7%, pediatrician route: 77.4%); (4) omitting dysmorphology measures from the model (psychologist route: 80.2%, pediatrician route: 79.0%); and (5) omitting neurobehavioral measures from the model (pediatrician route: 68.6%). To further ensure the robustness of our findings, we calculated accuracy rates for 10 random samples of 100 children from the model development sample (10 AE, 90 non-AE).

Although the tree model considers each child individually and is not influenced by the number of affected children included, using this ratio of exposed to nonexposed subjects (10%) is closer to the rate of affected children in the general population than the ratio in the larger study (34%). Using this sample, we found that application of the decision tree continued until sample sizes became too small for reliable comparison (n < 5). The average overall accuracy rate from these 10 random samples was identical to the original analyses (M = 84.6%).

Discussion

A diagnostic challenge arises in situations in which a clinician is asked to evaluate a child with neurodevelopmental problems who lacks physical features of FAS and has, at best, collateral reports of a history of prenatal alcohol exposure. The decision tree developed in this study differentiates children affected by prenatal alcohol exposure from nonexposed children by the use of physical and neurobehavioral measures, even in the absence of a diagnosis of FAS. The model requires a limited number of measures that can be obtained easily as part of standard clinical practice.

Our model resulted in accuracy rates (79.5%-84.7%) similar to previous studies, which achieved up to 84.7% accuracy in differentiating children with FASD from nonexposed control subjects.[7,8] There are several distinctions, however, that support significant incremental validity and improved clinical utility of our model. Post-hoc analyses revealed high accuracy rates (85%) even after we removed children with FAS, who are already reliably identified through dysmorphology exams. Exposed individuals without physical markers display cognitive and behavioral impairments similar to those found in FAS; thus, accurate identification of these individuals could facilitate access to intervention and tailored treatment. Our post-hoc analyses also indicated that the combination of both neurobehavioral and dysmorphology measures resulted in the greatest accuracy rates; reliance solely on dysmorphology measures resulted in reduced overall accuracy, because many children affected by prenatal alcohol exposure present without dysmorphology.

The accuracy of identification of individuals who are AE was achieved even when we compared them with a heterogeneous comparison group that included individuals with other clinical behavior conditions or concerns. Distinguishing children with alcohol exposure from the larger population of children with concerning behavior problems is more clinically

relevant than the potentially easier comparison of differentiating children who are AE from typically developing children and represents a more traditional outpatient practice. Therefore, even though the inclusion of this diverse comparison group increased the difficulty of an accurate diagnosis because of greater variance, it provided additional ecological validity to ensure that the model is well-suited for differential diagnosis or screening within an outpatient clinic.

Our model used a hierarchical, rather than concurrent, application of a very small number of measures, promoting parsimony and clinical feasibility. By design, not all children require all measures, which reduces time requirements for both families and clinicians. Alcohol effects may be ruled out quickly by the psychologist if no parent-reported behavior problems exist and IQ is >92 (step 1 and 2 using the psychologist route). In addition, the psychologist route ensures economy of assessment, because children would only be referred to a pediatrician for clinical behavior problems or IQ impairment on the basis of assessment or record review. By design, children who initially are seen by the pediatrician (ie, the pediatrician route) do not require the same testing as those who begin with the psychologist, further ensuring efficiency of decision making. Moreover, clinical utility is enhanced through the use of cut-off scores, which provide easy application of the tree in clinical settings. Refinement of these cut-off scores may improve accuracy rates of the decision tree while still maintaining ease of use. Importantly, the decision points within the proposed tree resemble the algorithm for evaluation of FASD provided by American Academy of Pediatrics and Centers for Disease Control and Prevention (http://www.cdc.gov/ncbddd/fasd/curriculum/fasdguide_web.pdf), although neurobehavioral impairments serve a larger role in our model, reflecting our multidisciplinary view of FASD. Moreover, the variables retained in the decision tree map on to the criteria for neurobehavioral disorder associated with prenatal alcohol exposure recently specified in the *Diagnostic and Statistical Manual of Mental Disorders - Fifth Edition*,[18,31] providing validation for the criteria for neurobehavioral disorder associated with prenatal alcohol exposure and clinical relevance of our decision tree model.

In addition to overall accuracy rates, positive predictive value and negative predictive value percentages were significantly greater than chance, indicating increased likelihood that a child is actually AE when classified as such by the decision tree. However, the decision tree correctly identified subjects in the nonexposed group at a greater rate than the AE group across all comparisons, which suggests stronger specificity than sensitivity, reducing the risk of false-positive identification, and promoting effective allocation of resources. Of note, our post-hoc analyses confirmed that differentiation of children who are alcohol-affected from nonexposed children with other clinical conditions or concerns is the most challenging. Although our accuracy rates were still significant and considerably greater than identification rates of children with FASD in a high-risk population,[5] further refinement of the decision tree model may improve accuracy rates.

Although our sample is diverse in its racial and ethnic composition, accuracy was not tested in every possible combination of demographic factors, and clinical judgment should be used when interpreting these results and implementing the decision tree in clinical settings. Another potential limitation is that the proportion of subjects with prenatal alcohol exposure in our sample likely does not reflect rates found in the population, given our targeted recruitment of this population. Although precise population estimates of alcohol consumption in pregnancy are not known, approximately one-third of our sample had histories of heavy prenatal alcohol exposure, which is likely greater than in the general population. Because each child's classification is independent of other children, composition of the subject group should not affect accuracy rates. Furthermore, our post hoc analyses with a more conservative proportion of exposed to nonexposed subjects yielded comparable results.

An additional limitation is the use of specific neurobehavioral tests in the decision tree. We used 2 different measures of general cognitive function with similar overall accuracy levels; however, other measures of behavior problems or adaptive function may lead to similar results. Although not tested in this study, it may be possible to use existing scores (eg, IQ scores from school records) to complete the decision tree. Similar accuracy rates using measures that are widely available and require a shorter administration time would promote flexibility and efficiency of the tree in clinical settings.

Although overall accuracies were high, there was variation of accuracy rates at each stopping point in the tree because of small sample sizes. This finding supports the need for further research extending those nodes to potentially improve accuracy rates. Regardless, the order of measures in the final model produced the greatest overall accuracy rates compared with more than 150 other iterations tested in the current study, and was validated in two age ranges in the independent CIFASD III sample.

Finally, as part of model development, we assessed the differential ability of more than 1000 variables separately, raising the possibility of potential increases in type 1 error. Even though the measures included in the tree had the greatest differential ability and accuracy rates were retained in an independent validation sample, further refinement through replication studies using other independent samples would increase the robustness of the proposed model.

Pediatricians represent a critical contact point for families of children with known or suspected prenatal alcohol exposure and, thus, in our model, are responsible for screening children for prenatal alcohol exposure, carrying out the physical examination, assessing adaptive behavior, and ultimately classifying a child as AE. This role is consistent with that proposed in the American Academy of Pediatrics FASD Toolkit (https://www.aap.org/en-us/advocacy-and-policy/aap-health-initiatives/fetal-alcohol-spectrum-disorders-toolkit/Pages/default.aspx) as well as the screening for other diagnoses such as autism spectrum disorders (http://www.aafp.org/afp/2008/1201/p1301.html). Our model provides a means to determine which children with neurodevelopmental problems are likely to be affected by prenatal alcohol exposure and require clinical follow-up. When a child is identified as "proposed AE"

via the use of this model, prenatal alcohol exposure should be considered as an etiological factor and additional evaluation through a multidisciplinary team with expertise in FASD may be appropriate.

The ability to accurately distinguish individuals affected by prenatal alcohol exposure is of great clinical importance, given the high rates of missed diagnosis and misdiagnosis.[5,32] Improving identification also would provide a basis for future research aimed at identifying developmental trajectories and refining intervention strategies. Currently there are very few studies in this area, because many children affected by prenatal alcohol exposure are not identified accurately. Thus, the information gained from the proposed model could inform new interventions specifically tailored toward the population with FASD. Replication studies, including deployment in diverse clinical settings, will further enhance the clinical utility of this decision tree. ■

We thank the families and children who participated in our studies and to the members of the Center for Behavioral Teratology for ongoing assistance and support. We also acknowledge the recruitment, data collection, and administrative efforts of Benjamin Deweese, Amy Flink, and Jill Vander Velde (San Diego); Trinh Luu and Alexy Andrade (Los Angeles); Sharron Paige-Whitaker (Atlanta); and Julia Tang, Birgit Fink, and Christopher J. Boys (Minneapolis).

Submitted for publication Jan 27, 2016; last revision received May 6, 2016; accepted Jun 13, 2016

Reprint requests: Sarah N. Mattson, PhD, 6330 Alvarado Court, Suite 100, San Diego, CA 92120. E-mail: sarah.mattson@sdsu.edu

References

1. Bertrand J, Floyd RL, Weber MK, O'Connor M, Riley EP, Johnson KA, et al. National Task Force on FAS/FAE: guidelines for referral and diagnosis. Atlanta (GA): Centers for Disease Control and Prevention; 2004.
2. Mattson SN, Crocker N, Nguyen TT. Fetal alcohol spectrum disorders: neuropsychological and behavioral features. Neuropsychol Rev 2011;21:81-101.
3. May PA, Keaster C, Bozeman R, Goodover J, Blankenship J, Kalberg WO, et al. Prevalence and characteristics of fetal alcohol syndrome and partial fetal alcohol syndrome in a Rocky Mountain Region City. Drug Alcohol Depend 2015;155:118-27.
4. Jones KL, Robinson LK, Bakhireva LN, Marintcheva G, Storojev V, Strahova A, et al. Accuracy of the diagnosis of physical features of fetal alcohol syndrome by pediatricians after specialized training. Pediatrics 2006;118:E1734-8.
5. Chasnoff IJ, Wells AM, King L. Misdiagnosis and missed diagnoses in foster and adopted children with prenatal alcohol exposure. Pediatrics 2015;135:264-70.
6. Streissguth AP, Barr HM, Kogan J, Bookstein FL. Final report: Understanding the occurrence of secondary disabilities in clients with fetal alcohol syndrome (FAS) and fetal alcohol effects (FAE). Seattle (WA): University of Washington Publication Services; 1996.
7. Mattson SN, Roesch SC, Fagerlund A, Autti-Ramo I, Jones KL, May PA, et al. Toward a neurobehavioral profile of fetal alcohol spectrum disorders. Alcohol Clin Exp Res 2010;34:1640-50.
8. Mattson SN, Roesch SC, Glass L, Deweese BN, Coles CD, Kable JA, et al. Further development of a neurobehavioral profile of fetal alcohol spectrum disorders. Alcohol Clin Exp Res 2013;37:517-28.
9. Mattson SN, Riley EP. The quest for a neurobehavioral profile of heavy prenatal alcohol exposure. Alcohol Res Health 2011;34:51-5.
10. Kalberg WO, May PA, Blankenship J, Buckley D, Gossage JP, Adnams CM. A practical testing battery to measure neurobehavioral ability among children with fetal alcohol spectrum disorder. Int J Alcohol Drug Res 2013;2:51-60.
11. Podgorelec V, Kokol P, Stiglic B, Rozman I. Decision trees: an overview and their use in medicine. J Med Syst 2002;26:445-63.
12. Falkmer T, Anderson K, Falkmer M, Horlin C. Diagnostic procedures in autism spectrum disorders: a systematic literature review. Eur Child Adolesc Psychiatry 2013;22:329-40.
13. Mattson SN, Foroud T, Sowell ER, Jones KL, Coles CD, Fagerlund A, et al. Collaborative initiative on fetal alcohol spectrum disorders: methodology of clinical projects. Alcohol 2010;44:635-41.
14. Wechsler D. Manual for the Wechsler intelligence scale for children – fourth edition integrated. 4th ed. San Antonio (TX): Pearson; 2004.
15. Elliott CD. Differential ability scales – second edition (DAS-II). San Antonio (TX): Harcourt Assessment; 2007.
16. Shaffer D, Fisher P, Lucas CP, Dulcan MK, Schwab-Stone ME. NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. J Am Acad Child Adolesc Psychiatry 2000;39:28-38.
17. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, fourth edition (DSM-IV). 4th ed. Washington (DC): American Psychiatric Association; 1994.
18. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, fifth edition (DSM-5). 5th ed. Arlington (VA): American Psychiatric Publishing, Inc; 2013.
19. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PMM. The diagnostic odds ratio: a single indicator of test performance. J Clin Epidemiol 2003;56:1129-35.
20. Achenbach TM, Rescorla LA. Manual for the ASEBA school-age forms & profiles. Burlington (VT): University of Vermont, Research Center for Children, Youth, & Families; 2001.
21. Jones KL, Hoyme HE, Robinson LK, del Campo M, Manning MA, Prewitt LM, et al. Fetal alcohol spectrum disorders: extending the range of structural defects. Am J Med Genet A 2010;152A:2731-5.
22. Hoyme HE, May PA, Kalberg WO, Kodituwakku P, Gossage JP, Trujillo PM, et al. A practical clinical approach to diagnosis of fetal alcohol spectrum disorders: clarification of the 1996 Institute of Medicine criteria. Pediatrics 2005;115:39-47.
23. Astley SJ. Diagnostic guide for fetal alcohol spectrum disorders: the 4-digit diagnostic code. 3rd ed. Seattle (WA): University of Washington; 2004.
24. Sparrow SS, Cicchetti DV, Balla DA. Vineland adaptive behavior scales, second edition: Survey forms manual. 2nd ed. Circle Pines (MN): AGS; 2005.
25. Cambridge Cognition. CANTABeclipse Version 3.0.0: Test Administration Guide. Cambridge, UK: Cambridge Cognition Limited; 2006.
26. Gioia GA, Isquith PK, Guy SC, Kenworthy L. Behavior rating inventory of executive function: professional manual. Odessa (FL): Psychological Assessment Resources, Inc; 2000.
27. Delis DC, Kaplan E, Kramer JH. The delis-kaplan executive function system: examiner's manual. San Antonio (TX): The Psychological Corporation; 2001.
28. Pelham WE Jr, Gnagy EM, Greenslade KE, Milich R. Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders. J Am Acad Child Adolesc Psychiatry 1992;31:210-8.
29. McBurnett K, Pfiffner LJ, Frick PJ. Symptom properties as a function of ADHD type: an argument for continued study of sluggish cognitive tempo. J Abnorm Child Psychol 2001;29:207-13.
30. McBurnett K, Pfiffner L. Sluggish cognitive tempo (SCT) scale. University of California, San Francisco. 2005.
31. Doyle LR, Mattson SN. Neurobehavioral disorder associated with prenatal alcohol exposure (ND-PAE):Review of evidence and guidelines for assessment. Curr Dev Disord Rep 2015;2:175-86.
32. Peadon E, Elliott EJ. Distinguishing between attention-deficit hyperactivity and fetal alcohol spectrum disorders in children: clinical guidelines. Neuropsychiatr Dis Treat 2010;6:509-15.

Table III. Logistic regression analyses for psychologist/pediatrician route in CIFASD II, CIFASD III child (5-7 y), and CIFASD III adolescent (10-16 y) samples

Decision points	OR	χ^2	Accuracy, %	*P* value
CIFASD II				
CBCL	31.6	149.6	83.8	<.001
FSIQ	6.5	28.2	75.9	<.001
Key facial features	2.1	5.7	60.0	.017
VABS (1)*	4.7	13.8	69.3	<.001
VABS (2)	11.3	14.6	77.2	<.001
Extended features	5.9	9.2	75.3	.002
CIFASD III (child)				
CBCL	32.7	35.0	82.0	<.001
GCA	2.7	2.1	67.9	.143
Key facial features	9.8	10.9	75.5	.001
VABS (1)†	-	-	-	-
VABS (2)	7.2	5.2	77.8	.022
Extended features†	-	-	-	-
CIFASD III (adolescent)				
CBCL	42.1	90.2	83.8	<.001
GCA	2.1	2.7	69.2	.101
Key facial features	2.2	4.0	59.5	.045
VABS (1)	4.8	5.6	65.5	.018
VABS (2)	12.2	12.9	78.2	<.001
Extended features†	-	-	-	-

Key facial features, excluding those with fetal alcohol syndrome (palpebral fissure length ≤10th percentile, vermilion border lipometer score = 4 or 5 [ie, thin vermilion border, see Astley[23]], philtrum lipometer score = 4 or 5 [ie, smooth philtrum, see ref. 23]); *extended features*, physical exam for extended features (ptosis, incomplete extension of >0 digits).
*VABS (1) and VABS (2) refer to the same variables from the Vineland Adaptive Behavior Scales-II, but are used at 2 different decision points in the decision tree model.
†Samples sizes too small to run analyses.

Table IV. Logistic regression analyses for pediatrician route in CIFASD II, CIFASD III child (5-7 y) and CIFASD III adolescent (10-16 y) samples

Decision points	OR	χ^2	Accuracy (%)	*P* value
CIFASD II				
Key facial features	2.4	11.3	67.2	.001
VABS (1)*	9.8	33.5	80.8	<.001
VABS (2)	20.8	27.1	82.4	<.001
Extended features	3.5	7.3	78.0	.007
CIFASD III (child)				
Key facial features	8.7	18.6	69.0	<.001
VABS (1)†	-	-	-	-
VABS (2)	10.5	14.7	75.0	<.001
Extended features†	-	-	-	-
CIFASD III (adolescent)				
Key facial features	3.3	14.8	65.8	<.001
VABS (1)	18.9	24.2	78.4	<.001
VABS (2)	14.1	21.7	78.2	<.001
Extended features†	-	-	-	-

*VABS (1) and VABS (2) refer to the same variables from the Vineland Adaptive Behavior Scales-II, but are used at 2 different decision points in the decision tree model.
†Samples sizes too small to run analyses.

Table V. Measures from which variables were obtained for consideration in the proposed decision tree model

Measures	Number of variables assessed
Cambridge Neuropsychological Test Automated Battery[25]	235
Behavior Rating Inventory of Executive Function[26]	123
Child Behavior Checklist[20]	24
Computerized Diagnostic Interview Schedule for Children-Fourth Edition[16]	14
Delis-Kaplan Executive Function System[27]	448
Disruptive Behavior Disorder Rating Scale[28]	130
Dysmorphology[21-23]	67
Sluggish Cognitive Tempo Rating Scale[29,30]	21
Vineland Adaptive Behavior Scales II[24]	219
Wechsler Intelligence Scale for Children-IV[14]	73
Youth Self-Report[20]	14