Title

Alcohol-related dysmorphic features as predictors of neurodevelopmental delay in infants and preschool-aged children: Results from a birth cohort in Ukraine.

Journal

Alcoholism: Clinical and Experimental Research, 46(12)

Authors

Bandoli, Gretchen

Coles, Claire

Kable, Julie

et al.

Publication Date

2022-12-01

DOI

10.1111/acer.14966

Peer reviewed

HHS Public Access

Author manuscript

Alcohol Clin Exp Res. Author manuscript; available in PMC 2023 December 01.

Published in final edited form as:

Alcohol Clin Exp Res. 2022 December ; 46(12): 2236–2244. doi:10.1111/acer.14966.

Alcohol-related dysmorphic features as predictors of neurodevelopmental delay in infants and preschool-aged children: results from a birth cohort in Ukraine

Gretchen Bandoli, PhD[1], **Claire Coles, PhD**[2], **Julie Kable, PhD**[2], **Kenneth Lyons Jones, MD**[1], **Erin Delker, PhD**[1], **Wladimir Wertelecki, MD**[1,3], **Lyubov Yevtushok, MD**[3,4,6], **Natalya Zymak-Zakutnya, MD**[3,5], **Iryna Granovska, MD**[3,4], **Larysa Plotka, MD**[3,4], **Christina Chambers, PhD**[1], **CIFASD**

[1]Department of Pediatrics, University of California San Diego

[2]Department of Psychiatry, Emory University

[3]OMNI-Net Ukraine Birth Defects Program

[4]Rivne Regional Medical Diagnostic Center, Rivne, Ukraine

[5]Khmelnytsky Perinatal Center, Khmelnytsky, Ukraine

[6]Lviv National Medical University, Lviv, Ukraine

Abstract

Background.—Cardinal and non-cardinal dysmorphic features are associated with prenatal alcohol exposure (PAE); however, their association with neurodevelopment is less clear. The objective of this study was to determine whether alcohol-related dysmorphic features predicted neurodevelopmental delay in infants and toddlers.

Methods.—We analyzed a prospective pregnancy cohort in western Ukraine enrolled between 2008-2014. A dysmorphology exam of body size, 3 cardinal and 14 non-cardinal dysmorphic features was performed at approximately 6-12 months of age. PAE was self-reported and operationalized as absolute ounces of alcohol per day around the time of conception. Neurodevelopment was assessed at 6-12 months with the Bayley Scales of Infant Development-II (BSID-II), and at 3.5-4.5 years of age with the Differential Ability Scales-II, the Child Behavior Checklist, and multiple measures that were used to create an executive functioning factor score. We performed logistic regression to predict children's neurodevelopment from dysmorphic features, growth measures, sex and PAE.

Results.—From an analytic sample of 582 unique children, 566 had BSID-II scores in infancy, and 289 completed the preschool battery. Models with all cardinal and non-cardinal dysmorphic features, growth measures, sex and PAE had the best performance compared to models with subsets of those inputs. In general, models had poor performance classifying delays in infancy (AUC <0.7) and acceptable performance on preschool-aged outcomes (AUC ~0.75). When the

Corresponding author: Gretchen Bandoli, PhD MPH MBA, 9500 Gilman Drive, MC 0828, San Diego CA 92093, gbandoli@ucsd.edu, 858-246-1733.

sample was limited to children with moderate to high PAE, predictive ability improved on preschool-aged outcomes (AUC 0.76-0.89). Sensitivity was relatively low on all models (12-63%), although other metrics of performance were higher.

Conclusion.—Predictive analysis based on dysmorphic features and measures of growth performed modestly in this sample. As these features are reliably measurable at an earlier age than neurodevelopment, their inclusion in predictive models should be further explored and validated in different settings and populations.

Keywords

prenatal alcohol exposure; neurodevelopment; predictive modeling; dysmorphology

Presence of at least two of the three cardinal features (short palpebral fissure, thin vermilion border and smooth philtrum) are consistently used as diagnostic criteria for fetal alcohol syndrome (FAS) among the leading diagnostic systems (Coles et al., 2016). Neuroimaging studies have reported changes in brain morphology associated with the cardinal features (Astley et al., 2009; Yang et al., 2012). There are at least 16 dysmorphic features assessed in standard dysmorphology exams used in the evaluation for fetal alcohol spectrum disorders (FASD), and there is evidence of an increase in prevalence of other physical malformations, outside of the cardinal features, associated with prenatal alcohol exposure, both in the presence and absence of FAS (Bandoli et al., 2020; del Campo and Jones, 2017; Feldman et al., 2011; Hoyme et al., 2016, 2005; Jones et al., 2010; May et al., 2007).

One issue that persists in the identification of children with or at risk for FASD is that cardinal features are not required for alcohol-related neurodevelopmental disorder (ARND) diagnosis. ARND is the most common of FASD categories, constituting 50% (May et al., 2018) to upwards of 80-90% (McQuire et al., 2019; Popova et al., 2019) of FASD cases. Further, while cardinal and non-cardinal features are associated with PAE, it is unclear the extent to which these features associate with neurodevelopmental deficits, which is a key criteria in diagnosing ARND. One study found that while PAE predicted neuropsychological deficits, the association was independent of the dysmorphic features of FAS (Mattson et al., 1998). Conversely, a different study found that FAS and partial FAS; i.e. requiring the presence of cardinal features, were associated with poorer neurodevelopmental functioning as compared to children with an ARND diagnosis (lacking cardinal features) (Chasnoff et al., 2010). Recently, an analysis of a cohort in South Africa found that a dysmorphology scoring system predicted FASD diagnosis; however, neurodevelopment as an outcome was not assessed (Kalberg et al., 2019).

Many studies to date of dysmorphology and neurodevelopment have involved samples discriminating between children with or without an FASD diagnosis. However, PAE is associated with neurodevelopmental deficits that may not rise to the level of FASD diagnosis (Coles et al., 2021, 2020). Given the reported association between prenatal alcohol use and an increase in select dysmorphic features, it is imperative to determine whether these dysmorphic features reliably predict neurocognitive or behavioral outcomes. Importantly, these alcohol-related minor malformations are present at birth, which is on average three years earlier than ARND is diagnosed or neurodevelopmental deficits are typically noticed

(Hoyme et al., 2016), in part because the faculties that are being measured (language, executive functioning) are not yet extant (Olson et al., 2007). A profile based on minor malformations that identifies those at risk for future neurodevelopmental delay presents the possibility of earlier identification and intervention of children prenatally exposed to alcohol. Therefore, our objective was to build and evaluate the optimal model of dysmorphic features and PAE information to predict neurodevelopment in infancy or at 3.5-4.5 years of age. In addition, given that PAE information is often unavailable or difficult to confirm in clinical practice, models would ideally not require PAE as an input. Therefore, we tested a model using only dysmorphology exam information excluding PAE. Finally, in an exploratory analysis conducted in a subset of the cohort who were evaluated for FASD, we evaluated whether the same nested models predicted an FASD diagnosis.

Study design

Data are from a prospective cohort study of pregnant women in western Ukraine conducted as part of the Collaborative Initiative on Fetal Alcohol Spectrum Disorders (CIFASD) supported by the National Institute on Alcohol Abuse and Alcoholism (www.CIFASD.org). This randomized clinical trial of micronutrient supplementation has been described elsewhere in detail (Chambers et al., 2014; Coles et al., 2015). Briefly, all pregnant women who presented to one of two centralized prenatal care facilities in western Ukraine, the Rivne Regional Medical Diagnostic Center and the Khmelnytsky Perinatal Center, between April 2008 through August 2014 were eligible for screening about their alcohol consumption around conception and the most recent month of pregnancy. Women who reported at least weekly binge episodes of 4-5 alcoholic drinks/occasion, at least 5 episodes of 3-4 drinks, or at least 10 episodes of 1-2 drinks in the month around conception and/or in the most recent month of pregnancy were recruited. Following identification of an exposed woman who consented to participate, the next minimally exposed or unexposed woman (no more than 2 drinks per occasion and no more than 2 drinks per week in the month around conception and no alcohol in the most recent month of pregnancy) was recruited for participation. Women were interviewed about demographics, behaviors and pregnancy characteristics using standard questionnaires upon enrollment and again at approximately 32 weeks of gestation. Mothers of live born infants were invited back twice in the first year postpartum (~six and twelve months) at which time neurodevelopmental assessments were conducted with the infant and dysmorphology exams were conducted. At 3.5-4.5 years of age, children returned for neurodevelopment assessments and dysmorphology exams.

The study was approved by the institutional review board at the University of California, San Diego and the Lviv National Medical University, Lviv, Ukraine.

Measures

Prenatal alcohol exposure

After enrollment, women who reported ever using alcohol in their lifetime completed a timeline follow-back assessment of day-by-day alcohol consumption by type, quantity and frequency in a typical week around conception and in the most recent two weeks of pregnancy (Sobell and Sobell, 2000). Quantity and frequency of alcohol consumption in

response to these questions was summarized as the average number of standard drinks per day over the period for which the mother was reporting. This information was then converted into absolute ounces of alcohol per day (ozAA/day). One standard drink is equivalent to 0.5 ozAA. For this analysis, we used the measure of the absolute ounces of alcohol per day at the time of enrollment.

Dysmorphology and anthropometric outcomes

This analysis relied on infant dysmorphology exams. Exams were performed by study pediatricians/geneticists at each of the study sites following specialized training and validation by one of us (KLJ) on recognition and measurement of all dysmorphic features, using a standard protocol and an examination checklist. Exams were completed at a median age of 12 months (range 0-52 months, interquartile range 5 months). Some children had more than one exam performed during infancy. In those situations, we prioritized the exam selected for analysis in the following order: 1) between 9 and 13 months of age (n=398), 2) between 4 and 8 months of age (n=138), 3) between 14 and 18 months of age (n=27), 4) ages from 0-3 months (n=9) or >18 months (n=10), using the first assessment that met this prioritization. Physical exams included assessment of height, weight, head circumference, the three cardinal features (short palpebral fissure length, thin smooth vermilion border, smooth philtrum), inner canthal distance, midface hypoplasia, epicanthal folds, anteverted nose, long philtrum, clinodactyly, camptodactyly, aberrant creases of the palm, railroad track ears, strabismus, ptosis, decreased pronation or supination of the arms, heart murmur, and heart defect. A common set of reference standards for dysmorphic features and growth was established for the CIFASD consortium irrespective of the population in which it was applied. These standards were used to assign percentiles for growth measures, palpebral fissure length, inner canthal distance and philtrum length and can be found at CIFASD.org. Smoothness/thinness of the vermilion border of the upper lip and smoothness of the philtrum were categorized using the Likert scale reference photographs published by Astley and Clarren (Astley and Clarren, 1996). A category of 4 or 5 was considered to indicate presence of the feature. The infants' weight, length and head circumference were converted to sex and age specific percentiles based on U.S. Centers for Disease Control and Prevention growth charts (Moriarty et al., 2009).

For analysis, palpebral fissures were dichotomized at < 10^{th} centile, as were measures for weight, length and head circumference. Thin smooth vermilion border of the upper lip and smoothness of the philtrum were each considered present if the Likert scale ranking was 4 or 5. Inner canthal distance (hypotelorism) was dichotomized at <25^{th} centile, and long philtrum was dichotomized at >90^{th} centile for length. All other features were dichotomized based upon presence/absence.

Neurodevelopmental assessments

Neurodevelopmental assessments were administered by one of two Ukrainian psychologists blinded to the mother's alcohol exposure group. These individuals were trained by psychologists on the study team, with periodic review of test administration and scoring carried out in person and via recordings.

Infant neurodevelopmental assessment

Neurodevelopment was evaluated at six and twelve months with the Bayley Scales for Infant Development, Second Edition (BSID-II) (Bayley, 1993). Participants who missed the six month assessment (n=120) remained eligible to complete the twelve month assessment.

We analyzed the Mental Development Index (MDI) and Psychomotor Development Index (PDI), both of which were standardized to a scale with a mean of 100 and a standard deviation of 15. The MDI assesses early cognitive and language development and the PDI assesses fine and gross motor development. All scores were age standardized and age corrected for births prior to 37 weeks. For analyses, BSID-II scores were dichotomized at ≤ 85, classifying developmental delay as scores falling outside of 1 standard deviation of the normative mean (Bayley, 1993).

Preschool neurodevelopmental assessment

Children were invited back for additional neurodevelopmental assessments between 3.5 and 4.5 years of age. The battery administered to the cohort was designed to be used in a non-English speaking environment while discriminating the effects of alcohol and are detailed elsewhere (Coles et al., 2021). Briefly, priority was given to non-verbal processing ability, visual/spatial processing, and aspects of executive functioning that could be measured in preschool. Language and verbal measures were omitted as a result of difficulties with cross-cultural adaptions of language. Instructions were provided to children in their native language. Children were administered the Nonverbal and Spatial portions of the Differential Ability Scales, 2nd Edition, Upper Preschool version (DAS-II) (Elliott, 2007), the Visual Attention subtest from the NEPSY (Korkman et al., 1998), the Statue and Speeded Naming subtest from the NEPSY-II (Korkman et al., 2007), the Attention Sustained subtest from the Leiter-3 (Korkman et al., 2007), a modification of the “Corsi Block” Task (Berch et al., 1998) for the preschool period, and other measures adapted for use in preschool children in Ukraine from non-standardized neuropsychological tests (Supplemental Table 1).

Coles and colleagues previously assessed the relationship between these measures and PAE (Coles et al., 2021), and reported significant group differences on the DAS-II Nonverbal Summary Scores, as well as a measure of executive functioning (EF), which was identified through principal component factor analysis using all non-DAS-II measures. A four factor solution accounted for 54% of the variance, with the following tests which appeared to represent executive functioning loading on the first factor at greater than 0.5 (Corsi Blocks Forward, Draw-a-line Slowly, Hand Game Imitative, Hand Game, Conflict, AB Game Total Correct, AB Game Preservation Errors, Delayed Alternation Total and number of Errors, and Subject Oriented Pointing). Based on these findings, scores ≤ 85 on the DAS-II Nonverbal Summary standard scores (mean of 100 and standard deviation of 15) and factor score of ≤-1 on the EF factor score (created by Coles et al; (Coles et al., 2021), mean of 0 and standard deviation of 1) were designed as thresholds for neurocognitive risk status.

In addition to the standardized neurocognitive outcomes, the primary caregiver completed a Ukrainian translated version of the Child Behavior Checklist (CBCL) (Achenbach and Edelbrock, 1991) to assess neurobehavioral risk. To assign risk status, we used one standard

deviation from the sample mean (≥60 based on a sample mean of 52.7 and a standard deviation of 8.6); which corresponds to the same value in the US reference provided by the manufacturer (T-score ≥60).

Finally, an FASD evaluation was completed for children with sufficient information, including the preschool-aged neurodevelopmental battery, using the criteria from the CoFASP study (May et al., 2018), which was based on modified Hoyme 2016 criteria (Hoyme et al., 2016). Children were evaluated for FAS, partial FAS, ARND, or no diagnosis. A full description of criteria as they relate to this cohort are available elsewhere (Kable et al., 2021).

Statistical analysis

In order to visualize potential patterns or clustering of dysmorphic features, we first performed a correlation analysis on the full set of dysmorphic and anthropometric features. In order to assess whether the features patterned differently by PAE, we stratified the sample by PAE group at enrollment.

We summarized all information (dysmorphic features, anthropometric measures, infant and preschool aged neurodevelopmental outcomes) by PAE group. Then, we performed logistic regression to examine the ability of nested sets of models to predict individuals with neurodevelopmental scores that deviated from the mean, indicating delay. Models were framed as a classification problem, where the sample was represented as belonging to two different classes, delayed and normal. For each outcome (MDI, PDI, DAS-II, CBCL, and EF factor score), we first regressed the outcome on a minimal model of only PAE (ozAA/day) and sex. In a second model, we added the three cardinal features. In a third model, we added three measures of growth (height, weight and head circumference), and finally, in the full model, we included all dysmorphic features, growth measures, PAE (ozAA/day) and sex. For analyses with a class imbalance, such as ours, accuracy is not a relevant metric, as the model would have high accuracy for simply always predicting the majority class. Thus, for each model, we requested the receive operating characteristic curve and the area under the curve (AUC), and for the full model. AUC is as the probability that the model ranks a random positive example more highly than a random negative example. However, AUC curves are less informative when a class imbalance is present (Carrington et al., 2020). Thus, we also calculated the sensitivity, specificity, positive predictive value, and negative predictive value of the algorithm. To describe AUC performance, we used categories by Hosmer and Lemeshow as follows (0.5= no discrimination; 0.5-0.7= poor discrimination; 0.7-0.8= acceptable discrimination; 0.8-0.9= excellent discrimination; >0.9=outstanding discrimination) (Hosmer Jr et al., 2013) Then, to determine whether the inputs were stronger predictors among children enrolled based on moderate to high PAE, we limited the sample to only children in the PAE group. In sensitivity analyses, we removed PAE (ozAA/day) from the full model to determine whether information on PAE was necessary given all physical features. In the exploratory analysis, we repeated all models with FASD diagnosis (any vs. none) and ARND diagnosis (yes vs. no) as the outcome. ARND models were assessed without the cardinal features to examine the contribution of non-cardinal features. All analyses were performed in SAS 9.4 (Cary, NC).

Results

Seven hundred and seventy-six women enrolled into the study at a mean gestational age of 18.2 weeks. From this cohort, 625 (81%) had an infant with complete dysmorphology exams. From this subsample, 566 had BSID scores at either 6 or 12 months of age, of which 273 also participated in preschool testing. There were an additional 16 children who participated in preschool testing who did not participate in the infant neurodevelopment testing visits. This resulted in an analytic sample of 582 children (75% of initial cohort).

In the analytic sample, 528 (91%) women were married or living with a partner, and 274 (47%) had education past a high school degree. Additionally, 371 (64%) were using prenatal vitamins at enrollment, and 174 (30%) were using tobacco at the time of conception. Tobacco use was more frequent among women enrolled in the alcohol-exposed group. Forty three percent of children were in the PAE group. Alcohol exposure in the PAE group was 0.77 ozAA/day (standard deviation 0.87) at conception and 0.23 ozAA/day (standard deviation 0.50) at enrollment. Slightly over half (52%) of the children were male (Table 1). Some dysmorphic features were common (epicanthal folds, 60%; anteverted nares, 17%; clinodactyly, 16%; thin vermilion border, 16%) while others were rare (ptosis, camptodactyly, pronation/supination of the elbows; all 1%). As expected, the prevalence of dysmorphic features was generally higher among those with PAE than those without. The correlation between physical features was generally <0.5, and varied by PAE status (Supplemental Figures 1–2). Among children with PAE (Supplemental Figure 1), the strongest correlations were observed between thin vermilion border and smooth philtrum (0.6), heart murmur and heart defect (0.6), and the measures of growth (0.3-0.4). Also, there was moderate correlation between the cardinal features and anteverted nares (~0.2) and long philtrum (0.2-0.3). Among children without PAE (Supplemental Figure 2), there was modest correlation between the growth measures (0.2-0.4), heart murmur and heart defect (0.5), and growth measures and short palpebral fissures (0.2-0.3). There was also moderate correlation between clinodactyly and smooth philtrum (0.2), and head circumference and short inner canthal distance (0.2). Generally, correlations between the features were weaker than those observed in the PAE group.

When assessing neurodevelopmental outcomes (Table 1), children with PAE were more likely to fall within the range of neurodevelopmental risk than children without PAE on all measures. Finally, in 307 children evaluated for FASD, 62 (20%) met the study criteria for a diagnosis (6 FAS, 10 partial FAS, and 46 ARND).

Predictive models- neurodevelopment

We first built nested models in the full sample (Table 2). Generally, PAE and sex alone was a poor predictor of each neurodevelopmental outcome (AUC 0.56-0.64). Model performance as measured by the AUC benefitted from the addition of cardinal features (Model 2) and growth measures (Model 3). However, for all outcomes, the highest AUC came from Model 4 when all dysmorphic features and measures of growth were included. The improvement in AUC gained with sequential models improved the prediction of the preschool outcomes compared to prediction of the infant outcomes. Notably, model performance was still poor to acceptable; AUC for the full models ranged from 0.63 to 0.75, and although specificity was

high, sensitivity was low for all outcomes. In the sensitivity analysis, AUC was minimally impacted by removing PAE from the full model (Model 5) (Table 2). AUC dropped by 0.1 to 0.2 across outcomes, and still outperformed Models 1-3.

When the sample was limited to children in the PAE group, model performance as measured by the AUC modestly improved (Table 3). For all outcomes, the full model (Model 4) had the highest AUC. Again, preschool outcomes changed more with the additional inputs than infant outcomes. The AUC ranged from poor to acceptable (0.68-0.76) on infant outcomes, and acceptable to excellent (0.76-0.89) on preschool outcomes. Specificity was high for the full model, although sensitivity, while higher than in the full sample, was still average (31-63%). Positive predictive value (proportion of delayed children among those classified as delayed) ranged from 64-83%, and negative predictive value (proportion of kids not delayed among those classified as having no delay), ranged from 70-90%. In the sensitivity analysis in the PAE group, AUC was again minimally impacted by removing PAE from the full model (Model 5) (Table 3).

Exploratory models- FASD diagnosis

In exploratory analysis with FASD as an outcome in the full sample, the AUC was very similar to the DAS-II and EF factor score models. The full model (Table 2) had an AUC of 0.75, sensitivity of 27%, and positive predictive value of 68%. In models limited to children in the PAE group (Table 3), models had acceptable discrimination, although slightly lower AUC than specific neurodevelopmental outcomes. Sensitivity was 62%, and positive predictive value was 73%. We repeated models with ARND as the outcome (Supplemental Table 2), excluding the cardinal features to evaluate non-cardinal features only. Model performance was essentially unchanged, with the exception of lower sensitivity in the full model (17%).

Discussion

Alcohol use in women of reproductive ages is common, and globally, approximately 10% of pregnancies are exposed to alcohol (Lange et al., 2018; McCaul et al., 2019; Popova et al., 2017). PAE is associated with dysmorphic features, which are identifiable from a young age, often before neurodevelopmental delay becomes apparent. This relationship seeds the desire to predict future neurodevelopmental deficits by dysmorphic features. In this study of 582 children from a birth cohort in Ukraine, we found that measures of growth and most alcohol-related dysmorphic features were higher in children with PAE than those without. When features were used to classify children with or without neurodevelopmental deficits on five developmental assessments, model performance was generally poor for infant outcomes and acceptable to excellent for preschool-aged outcomes. The classifiers performed better when limited to children with any PAE, and for preschool outcomes the AUC was classified as acceptable to excellent. Finally, when all dysmorphic features, growth and sex were included in the models, performance was minimally impacted by removing the PAE variable, suggesting that in the presence of physical features, the features account well for the variance in neurodevelopmental impairment associated with PAE.

When examining our models individually, models had better performance classifying deficits in preschool-aged children as compared to infants, which was not all together surprising. In earlier work from this cohort, we reported reductions in performance on the MDI and PDI in 6-12 month infants as PAE increased, but most scores were not below the one standard deviation threshold for neurodevelopmental risk used here (Bandoli et al., 2019). Given the relative insensitivity of infant neurodevelopmental measures in predicting later cognitive function (Bornstein and Krasnegor, 2013; Matheny Jr, 1989; Olson et al., 2007), poor performance of the classifier for infant outcomes was not unexpected. In contrast, there was better classification by dysmorphic features when assessing preschool aged outcomes, specifically measures of nonverbal functioning, total internalizing and externalizing behaviors, and executive functioning. However, in the full sample, the prevalence of the outcome was low (15-20%), and sensitivity was poor. When models were limited to children with any PAE, dysmorphic features were stronger at correctly classifying children, and sensitivity increased to 55-63% on the DAS-II and executive functioning factor score, and the positive predictive value was 71-83%. Interestingly, the caregiver-rated CBCL had much lower sensitivity (31%). Respective to the FASD diagnosis, the full model classified the diagnosis with the same performance as classifying the DAS-II or executive functioning measure. This is perhaps unsurprising as those two measures are components of the diagnosis. Finally, two additional findings should be highlighted. Particularly on preschool-aged outcomes, model performance was enhanced by including all the dysmorphic features as opposed to just the cardinal features, highlighting the role that non-cardinal features may have in identifying children with alcohol-related neurodevelopmental delay. Also, when the full panel of dysmorphic features, growth and sex were included in models, performance was only minimally improved by the inclusion of the measure of PAE. This is encouraging, as true PAE is often difficult to ascertain, and the ability to identify children affected by PAE in the absence of that information would greatly improve the screening process.

This work follows the efforts of many others to identify inputs that predict the effects of PAE, whether as an FASD diagnosis or neurodevelopmental deficits. The most relevant work by Kalberg et al. studied whether the dysmorphology score (Hoyme et al., 2016) predicted FASD diagnosis. The score is a linear combination of the dysmorphic features with assigned weights (ranging from 1-3). There is substantial overlap in the dysmorphic features in the dysmorphology score and those used in our analysis, the main difference being the lack of weighting in our approach. Kalberg et al. found that the dysmorphology score at 9 months of age predicted FASD diagnosis at age 5 with an AUC of 0.78 (Kalberg et al., 2019), which is very similar to our full model with an AUC of 0.75. The similar model performance is encouraging given the different settings (Ukraine versus South Africa) and study designs. Lussier and colleagues tested a predictive model to classify children with FASD from children without FASD from genome-wide methylation patterns, resulting in an AUC of 0.92, a positive predictive value of 0.90 and a negative predictive value of 0.79 (Lussier et al., 2018). These results were robust to a hold-out sample and to comparisons with children with autism spectrum disorder, confirming the specificity of the findings. The cardiac orienting response has been examined twice in the same cohort as this study; once with infant neurodevelopment and once with FASD as the outcome (Kable et al., 2021). The

cardiac orienting response at 6-months of age was a good predictor (AUC 0.81) of 12-month neurodevelopmental delay (Mesa et al., 2017). In a subsample later evaluated for FASD, the cardiac orienting responses was an acceptable predictor of FASD at age 3.5-4.5 years (AUC 0.77) (Kable et al., 2021). Overall, our classifier relying on the dysmorphic features, growth and sex was comparable to these other efforts. Finally, in a community-based sample of 6-year olds, Coles et al found that maternal and child characteristics, including other substances used during pregnancy, socioeconomic status, were significant inputs into predictive models for FASD-related outcomes (Coles et al., 2020). In future work, we will expand our inputs to include maternal characteristics, other prenatal exposures, and postnatal factors to the dysmorphic features to try to maximize model performance.

Despite acceptable to excellent performance of some of the models, our findings should be viewed with caution. We did not validate our results in an external sample, which is necessary to adequately interrogate a predictive model. In addition, developmental delay was fairly rare, thus AUC can overstate classifier performance. Sensitivity and positive predictive value are key performance metrics, which reflect the proportion of true cases that would be missed in each model and the proportion of children labeled as delayed who, in fact, are not. On those metrics, the classifiers had average performance. Also, specific to our sample, a multivitamin intervention was administered in the study which was found to be beneficial to infant neurodevelopment of the alcohol exposed children (Coles et al., 2015). This intervention may weaken the association between dysmorphology and neurodevelopment, particularly in models restricted to children with PAE, and may not generalize to samples with lower multivitamin use. Further, we examined models with PAE operationalized as average AAoz/day due to our previous studies that found associations between this measure and both dysmorphology (Bandoli et al., 2020) and neurodevelopment (Bandoli et al., 2019). However, models that incorporate binge drinking or peak blood alcohol, which may be stronger predictors of dysmorphology and neurodevelopment, may have better performance. Finally, the level of PAE in this cohort was higher than would be expected in a community-based sample, and thus it is unclear how this classifier, as tested, would perform in a different population.

In summary, although single domains alone will likely never perform as well as a broad spectrum of domains, we have shown that alcohol-related dysmorphic features measured in infancy, coupled with measures of growth and PAE, do a reasonable job at discriminating between preschool-aged children with and without neurodevelopmental delay. Model performance was improved when the classifier was limited to children with any PAE, particularly on non-verbal measures and executive functioning. Validating these findings with an external dataset is an important next step to understanding the relationship between alcohol-related dysmorphic features and neurodevelopment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

All or part of this work was done in conjunction with the Collaborative Initiative on Fetal Alcohol Spectrum Disorders (CIFASD), which is funded by grants from the National Institute on Alcohol Abuse and Alcoholism (NIAAA). Additional information about CIFASD can be found at http://cifasd.org/. Research described in this manuscript was supported by Grant #U01AA014835 to Christina Chambers funded by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the NIH Office of Dietary Supplements (ODS). Gretchen Bandoli is funded by the National Institutes on Alcohol Abuse and Alcoholism (1 K01 AA027811-01). We wish to acknowledge the contribution of: OMNI-Net, Ukraine, and participating families and staff in Rivne and Khmelnytsky, Ukraine.

References

Achenbach TM, Edelbrock C (1991) Child behavior checklist. Burlingt 7:371–392.

Astley SJ, Aylward EH, Olson HC, Kerns K, Brooks A, Coggins TE, Davies J, Dorn S, Gendler B, Jirikowic T, Kraegel P, Maravilla K, Richards T (2009) Magnetic resonance imaging outcomes from a comprehensive magnetic resonance study of children with fetal alcohol spectrum disorders. Alcohol Clin Exp Res 33:1671–1689. [PubMed: 19572986]

Astley SJ, Clarren SK (1996) A case definition and photographic screening tool for the facial phenotype of fetal alcohol syndrome. J Pediatr 129:33–41. [PubMed: 8757560]

Bandoli G, Coles CD, Kable JA, Wertelecki W, Yevtushok L, Zymak-Zakutnya N, Wells A, Granovska IV, Pashtepa AO, Chambers CD (2019) Patterns of Prenatal Alcohol Use That Predict Infant Growth and Development. Pediatrics e20182399. [PubMed: 30610099]

Bandoli G, Jones K, Wertelecki W, Yevtushok L, Zymak-Zakutnya N, Granovska I, Plotka L, Chambers C (2020) Patterns of Prenatal Alcohol Exposure and Alcohol-Related Dysmorphic Features. Alcohol Clin Exp Res 1–8.

Bayley N (1993) Bayley Scales of Infant Development Manual (BSID-II). San Antonio, TX, The Psychological Corporation.

Berch DB, Krikorian R, Huha EM (1998) The Corsi block-tapping task: Methodological and theoretical considerations. Brain Cogn 38:317–338. [PubMed: 9841789]

Bornstein MH, Krasnegor NA (2013) Stability and continuity in mental development: Behavioral and biological perspectives. Psychology Press.

Carrington AM, Fieguth PW, Qazi H, Holzinger A, Chen HH, Mayr F, Manuel DG (2020) A new concordant partial AUC and partial c statistic for imbalanced data in the evaluation of machine learning algorithms. BMC Med Inform Decis Mak 20:4. [PubMed: 31906931]

Chambers CD, Yevtushok L, Zymak-Zakutnya N, Korzhynskyy Y, Ostapchuk L, Akhmedzhanova D, Chan PH, Xu R, Wertelecki W (2014) Prevalence and Predictors of Maternal Alcohol Consumption in Two Regions of Ukraine. Alcohol Clin Exp Res 38:1012–1019. [PubMed: 24834525]

Chasnoff IJ, Wells AM, Telford E, Schmidt C, Messer G (2010) Neurodevelopmental functioning in children with FAS, pFAS, and ARND. J Dev Behav Pediatr 31:192–201. [PubMed: 20375733]

Coles CD, Gailey AR, Mulle JG, Kable JA, Lynch ME, Jones KL (2016) A Comparison Among 5 Methods for the Clinical Diagnosis of Fetal Alcohol Spectrum Disorders. Alcohol Clin Exp Res 40:1000–1009. [PubMed: 27028727]

Coles CD, Kable JA, Granovska IV, Pashtepa AO, Wertelecki W, Chambers CD (2021) Measurement of neurodevelopmental effects of prenatal alcohol exposure in Ukrainian preschool children. Child Neuropsychol a J Norm Abnorm Dev Child Adolesc 27:1088–1103.

Coles CD, Kable JA, Keen CL, Jones KL, Granovska IV, Pashtepa AO, Chambers CD (2015) Dose and Timing of Prenatal Alcohol Exposure and Maternal Supplements: Developmental Effects on 6-Month-Old Infants. Matern Child Health J 19:2605–2614. [PubMed: 26164422]

Coles CD, Kalberg W, Kable JA, Tabachnick B, May PA, Chambers CD (2020) Characterizing Alcohol-Related Neurodevelopmental Disorder: Prenatal Alcohol Exposure and the Spectrum of Outcomes. Alcohol Clin Exp Res 44:1245–1260. [PubMed: 32173870]

del Campo M, Jones KL (2017) A review of the physical features of the fetal alcohol spectrum disorders. Eur J Med Genet 60:55–64. [PubMed: 27729236]

Elliott CD (2007) Differential Ability Scales (2nd ed.). Can J Sch Psychol 22:128–132.

Feldman HS, Jones KL, Lindsay S, Slymen D, Klonoff-Cohen H, Kao K, Rao S, Chambers C (2011) Patterns of prenatal alcohol exposure and associated non-characteristic minor structural malformations: A prospective study. Am J Med Genet Part A 155:2949–2955.

Hosmer DW Jr, Lemeshow S, Sturdivant RX (2013) Applied logistic regression. John Wiley & Sons.

Hoyme HE, Kalberg WO, Elliott AJ, Blankenship J, Buckley D, Marais A-S, Manning MA, Robinson LK, Adam MP, Abdul-Rahman O, Jewett T, Coles CD, Chambers C, Jones KL, Adnams CM, Shah PE, Riley EP, Charness ME, Warren KR, May PA (2016) Updated Clinical Guidelines for Diagnosing Fetal Alcohol Spectrum Disorders. Pediatrics 138:e20154256–e20154256. [PubMed: 27464676]

Hoyme HE, May PA, Kalberg WO, Kodituwakku P, Gossage JP, Trujillo PM, Buckley DG, Miller JH, Aragon AS, Khaole N, Viljoen DL, Jones KL, Robinson LK (2005) A practical clinical approach to diagnosis of fetal alcohol spectrum disorders: Clarification of the 1996 institute of medicine criteria. Pediatrics 115:39–47. [PubMed: 15629980]

Jones KL, Hoyme HE, Robinson LK, del Campo M, Manning MA, Prewitt LM, Chambers CD (2010) Fetal Alcohol Spectrum Disorders: Extending the Range of Structural Defects. Am J Med Genet A 152A:2731–2735. [PubMed: 20949507]

Kable JA, Coles CD, Jones KL, Yevtushok L, Kulikovsky Y, Zymak-Zakutnya N, Dubchak I, Akhmedzhanova D, Wertelecki W, Chambers CD, CIFASD (2021) Infant Cardiac Orienting Responses Predict Later FASD in the Preschool Period. Alcohol Clin Exp Res 45:386–394. [PubMed: 33277942]

Kalberg WO, May PA, Buckley D, Hasken JM, Marais AS, De Vries MM, Bezuidenhout H, Manning MA, Robinson LK, Adam MP, Hoyme DB, Parry CDH, Seedat S, Elliott AJ, Hoyme HE (2019) Early-life predictors of fetal alcohol spectrum disorders. Pediatrics 144.

Korkman M, Kirk U, Kemp S (2007) NEPSY II: Clinical and interpretive manual. Harcourt Assessment, PsychCorp.

Korkman M, Kirk U, Kemp S (1998) NEPSY: A Developmental Neuropsychological Assessment Manual.

Lange S, Probst C, Gmel G, Rehm J, Burd L, Popova S, Lange S, Probst C, Gmel G, Rehm J (2018) Global prevalence of alcohol use and binge drinking during pregnancy, and fetal alcohol spectrum disorder. Biochem Cell Biol 96:237–240. [PubMed: 28834683]

Lussier AA, Morin AM, MacIsaac JL, Salmon J, Weinberg J, Reynolds JN, Pavlidis P, Chudley AE, Kobor MS (2018) DNA methylation as a predictor of fetal alcohol spectrum disorder. Clin Epigenetics 10:1–14. [PubMed: 29312470]

Matheny AP Jr (1989) Assessment of infant mental development: Toward a broader perspective. Clin Perinatol 16:565–576. [PubMed: 2663316]

Mattson SN, Riley EP, Gramling L, Delis DC, Jones KL (1998) Neuropsychological comparison of alcohol-exposed children with or without physical features of fetal alcohol syndrome. Neuropsychology.

May PA, Chambers CD, Kalberg WO, Zellner J, Feldman H, Buckley D, Kopald D, Hasken JM, Xu R, Honerkamp-Smith G, Taras H, Manning MA, Robinson LK, Adam MP, Abdul-Rahman O, Vaux K, Jewett T, Elliott AJ, Kable JA, Akshoomoff N, Falk D, Arroyo JA, Hereld D, Riley EP, Charness ME, Coles CD, Warren KR, Jones KL, Hoyme HE, Daniel F, Arroyo JA, Hereld D, Riley EP, Charness ME, Coles CD, Warren KR, Jones KL, Hoyme HE (2018) Prevalence of fetal alcohol spectrum disorders in 4 US communities. JAMA - J Am Med Assoc 319:474–482.

May PA, Gossage JP, Marais A-S, Adnams CM, Hoyme HE, Jones KL, Robinson LK, Khaole NCO, Snell C, Kalberg WO, Hendricks L, Brooke L, Stellavato C, Viljoen DL (2007) The epidemiology of fetal alcohol syndrome and partial FAS in a South African community. Drug Alcohol Depend 88:259–271. [PubMed: 17127017]

McCaul ME, Roach D, Hasin DS, Weisner C, Chang G, Sinha R (2019) Alcohol and Women: A Brief Overview. Alcohol Clin Exp Res.

McQuire C, Mukherjee R, Hurt L, Higgins A, Greene G, Farewell D, Kemp A, Paranjothy S (2019) Screening prevalence of fetal alcohol spectrum disorders in a region of the United Kingdom: A population-based birth-cohort study. Prev Med (Baltim) 118:344–351.

Mesa DA, Kable JA, Coles CD, Jones KL, Yevtushok L, Kulikovsky Y, Wertelecki W, Coleman TP, Chambers CD, CIFASD (2017) The use of cardiac orienting responses as an early and scalable biomarker of alcohol-related neurodevelopmental impairment. Alcohol Clin Exp Res 41:128–138. [PubMed: 27883195]

Moriarty DG, Zack MM, Holt JB, Chapman DP, Safran M a. (2009) Geographic Patterns of Frequent Mental Distress. U.S. Adults, 1993-2001 and 2003-2006. Am J Prev Med 36:497–505. [PubMed: 19460657]

Olson HC, Jirikowic T, Kartin D, Astley S (2007) Responding to the Challenge of Early Intervention for Fetal Alcohol Spectrum Disorders 20:172–189.

Popova S, Lange S, Poznyak V, Chudley AE, Shield KD, Reynolds JN, Murray M, Rehm J (2019) Population-based prevalence of fetal alcohol spectrum disorder in Canada. BMC Public Health 19:845. [PubMed: 31253131]

Popova S, Lange S, Probst C, Gmel G, Rehm J (2017) Estimation of national, regional, and global prevalence of alcohol use during pregnancy and fetal alcohol syndrome: a systematic review and meta-analysis. Lancet Glob Heal 5:e290–e299.

Sobell L, Sobell M (2000) Alcohol timeline follow-back (TLFB) In: Handbook of Psychiatric Measures. , p 477. Washington, DC, American Psychiatric Association.

Yang Y, Phillips OR, Kan E, Sulik KK, Mattson SN, Riley EP, Jones KL, Adnams CM, May PA, O'Connor MJ (2012) Callosal thickness reductions relate to facial dysmorphology in fetal alcohol spectrum disorders. Alcohol Clin Exp Res 36:798–806. [PubMed: 22150665]

Table 1.

Alcohol-related dysmorphic features and neurodevelopmental outcomes of children enrolled in a cohort study in Western Ukraine (n=582)

	No prenatal alcohol exposure (n=329)		Prenatal alcohol exposure (n=253)	
	n	%	n	%
Sex (male)	181	55.0	124	49.0
Growth measures				
Height < 10th centile	20	6.1	33	13.0
Weight < 10th centile	26	7.9	40	15.8
Head circumference < 10th centile	16	4.9	30	11.9
Alcohol-related dysmorphic features				
Short palpebral fissure	23	7.0	48	19.0
Smooth philtrum	29	8.8	34	13.4
Thin vermilion border	44	13.4	51	20.2
Midface Hypoplasia	6	1.8	4	1.6
Inner canthal distance < 25th centile	25	7.6	44	17.4
Railroad ears	20	6.1	18	7.1
Ptosis	2	0.6	5	2.0
Epicanthal folds	200	60.8	147	58.1
Anteverted nares	47	14.3	50	19.8
Long philtrum	23	7.0	24	9.5
Clinodactyly of the 5th finger	51	15.5	42	16.6
Camptodactyly	4	1.2	3	1.2
Hockey stick crease	25	7.6	20	7.9
Pronation/supination of elbows	0	0.0	4	1.6
Strabismus	6	1.8	8	3.2
Heart murmur	36	10.9	38	15.0
Heart defect	11	3.3	20	7.9
6-12 month infant outcomes[a]				
Psychomotor Development Index (PDI) < 85	75	23.3	89	36.6
Mental Development Index (MDI) < 85	91	28.2	110	45.3
Preschool aged outcomes[b]				
Child behavior checklist total score (CBCL) > 60	30	16.2	29	27.9
Executive functioning factor score (EF) < −1	19	11.0	22	19.5
DAS-II Nonverbal summary < 85	19	11.1	24	22.2
Fetal alcohol spectrum disorder diagnosis[b]	4	2.1	58	50.9

PAE: prenatal alcohol exposure

[a]Sample sizes: MDI/PDI (no PAE: 323 ; PAE: 243)

[b]Sample sizes: CBCL (no PAE: 185, PAE: 104); EF: (no PAE 173; PAE: 113); DAS (no PAE: 171; PAE: 108); FASD diagnosis (no PAE: 193; PAE: 114)

Table 2.

Performance of logistic regression models predicting child neurodevelopmental outcomes (n=582)

Area under the curve for each model	Psychomotor Development Index	Mental Development Index	DAS-II nonverbal summary score	CBCL total score	EF factor score	FASD diagnosis
Model 1: Prenatal alcohol exposure (ozAA/day) and sex	0.58	0.63	0.64	0.56	0.64	0.71
Model 2: Model 1 plus cardinal features	0.60	0.64	0.71	0.58	0.69	0.73
Model 3: Model 2 plus growth	0.60	0.67	0.72	0.56	0.71	0.73
Model 4: Full model (Model 3 plus remaining dysmorphic features)	0.63	0.69	0.75	0.71	0.74	0.75
Sensitivity analysis: Model 5: Model 4 minus prenatal alcohol exposure (ozAA/day)	0.62	0.67	0.75	0.71	0.72	0.75
Sensitivity for Model 4 (full model)	15%	32%	19%	12%	12%	27%
Specificity for Model 4 (full model)	97%	92%	97%	98%	98%	97%
PPV for Model 4 (full model)	68%	68%	57%	58%	50%	68%
NPV for Model 4 (full model)	74%	71%	87%	81%	87%	84%

PPV: positive predictive value; NPV: negative predictive value; DAS-II: Nonverbal and Spatial portions of the Differential Ability Scales, 2nd Edition; CBCL: Child behavior checklist; EF: Executive functioning; FASD: fetal alcohol spectrum disorder

Table 3.

Performance of logistic regression models predicting child neurodevelopmental outcomes among children exposed[a] prenatally to alcohol (n=253)

Area under the curve for each model	Psychomotor Development Index	Mental Development Index	DAS-II nonverbal summary score	CBCL total score	EF factor score	FASD diagnosis
Model 1: Prenatal alcohol exposure (ozAA/day) and sex	0.59	0.67	0.66	0.48	0.72	0.57
Model 2: Model 1 plus cardinal features	0.62	0.70	0.75	0.60	0.77	0.73
Model 3: Model 2 plus growth	0.63	0.72	0.78	0.55	0.83	0.66
Model 4: Full model (Model 3 plus remaining dysmorphic features)	0.66	0.76	0.89	0.76	0.88	0.77
Sensitivity analysis: Model 5: Model 4 minus prenatal alcohol exposure (ozAA/day)	0.65	0.74	0.89	0.76	0.86	0.77
Sensitivity for Model 4 (full model)	35%	56%	63%	31%	55%	62%
Specificity for Model 4 (full model)	91%	83%	96%	93%	95%	77%
PPV for Model 4 (full model)	69%	73%	83%	64%	71%	73%
NPV for Model 4 (full model)	71%	70%	90%	78%	90%	66%

PPV: positive predictive value; NPV: negative predictive value; DAS-II: Nonverbal and Spatial portions of the Differential Ability Scales, 2nd Edition; CBCL: Child behavior checklist; EF: Executive functioning; FASD: fetal alcohol spectrum disorder

[a]Met criteria for enrollment into the PAE group