
Machine Learning-Based Screening for Fetal Alcohol Spectrum Disorder

Chen Zhang, Laurent Itti*
University of Southern California
Los Angeles, CA 90089

Po-He Tseng
Duke University
Durham, NC 27708

Angelina Paolozza, James N. Reynolds, Douglas P. Munoz
Queen's University
Kingston, ON K7L 3N6

Abstract

Fetal alcohol spectrum disorders (FASD) is one of the most prevalent neurodevelopmental disorders, placing high psychological pressures and large economic burdens on families and society. Yet, the current diagnostic process and screening tools are challenging and time- and money- consuming, with underreported profiles of neurocognitive and neurobehavioral impairments, and an underestimated proportion of at-risk population because of limited clinical resources. In this work, we assessed children/youth with FASD from a multi-modal perspective and developed a high-throughput, low-cost screening protocol using a machine learning framework. This annual screening procedure for children/youth at risk of FASD can be easily and widely deployed with high expected monetary benefit, potentially leading to earlier intervention and treatment crucial for neurodevelopmental disorders.

1 Introduction

Fetal alcohol spectrum disorder (FASD), a consequence of prenatal exposure to alcohol [1], is one of the most common causes of developmental disabilities and neurobehavioral deficits. The estimated prevalence of FASD among school-age children may be as high as 2-5% in the U.S. [2,3]. Costs associated with FASD, in areas such as medical care, special education and facilities, and social services, can run into billions of dollars annually, which places a large burden on both families and society [4,5]. Early screening is of significant importance because it can lead to early intervention, which can mitigate the disruptive effects of the disorder and also reduce the probability of developing secondary disabilities. However, the current clinical diagnostic procedures may take up to two full days, requiring a multidisciplinary team comprised of a physician, psychologist, facial dysmorphologist and occupational therapist [6]. The diagnostic criteria vary across clinics and countries, leading to inconsistent diagnoses and treatments. The high rate of co-morbidity with other developmental disorders such as attention deficit hyperactivity disorder [7] may also contribute to misdiagnosis. Most screening or discrimination studies have so far relied on demographic, behavioral, physical, and psychometric data, face morphometric analysis, and history of maternal alcohol consumption [8-12], which are not time- and cost-efficient. Therefore, an easy, objective, and effective procedure with globally unified standards to assess the deficits and differentiate the neurological groups is desperately needed as a screening tool for children/youth at risk for FASD, to help target the at-risk population, accelerate the clinical process, and improve diagnostic accuracy and efficiency.

Compared to age-matched typically developing (TD) controls, previous studies have revealed abnormalities in multiple brain regions, including the corpus callosum of children/youth with FASD.

Other types of deficits, including dysfunctions in oculomotor control, visuospatial ability, attention and working memory, and alternations in executive functions and academic functions, are frequently reported for the FASD group. In this paper, we start from a multi-modal perspective of those structural, functional and cognitive impairments, including measurements from saccadic eye movement and free viewing behavior, psychometric performance, and neuroimaging, to construct classifiers that demonstrate high performances in identifying the clinical population. Considering time, cost, age restriction, required administration, and accessibility of different measurements, we also propose a high-throughput and low-cost screening procedure with high expected monetary benefit, based on a subset of eye movement and psychometric tests, which could be widely deployed and lead to the earlier intervention and treatment that is crucial for neurodevelopmental disorders. To quantify the potential benefits of early screening on a large scale, we used tools from the theory of value of information [13]. We propose a cost-benefit model based on this theory to evaluate the expected benefits of our screening procedure.

2 Materials and Methods

2.1 Multi-modal tests and data collection

This study included children/youth aged 5-18 years who were either TD healthy controls (n=116) or had received a diagnosis of an FASD (n=91). Participants completed up to six tests, including saccadic eye movement tasks (prosaccade, antisaccade and memory-guided/M-G saccade), free viewing of videos, psychometric tests, and neuroimaging of the corpus callosum. See Table 1 for the number of participants involved in each and in all of the tests.

Table 1: Number of participants for different tests

	Prosaccade	Antisaccade	M-G saccade	Free viewing	Psych	DTI	Six tests
n(FASD)	71	67	61	47	58	41	22
n(control)	115	106	93	53	71	35	24

Saccadic eye movement tasks. Participants were seated in front of a monitor with a built-in infrared illuminator and infrared camera. The coordinates of the left pupil were sampled in the vertical and horizontal axes using the Eyelink 1000 system (SR Research, Kanata, ON). Each trial started with the illumination of a central fixation point (FP). In the **prosaccade** and **antisaccade** tasks [14,15], a peripheral target appeared randomly to the left or right after the FP disappeared. Participants needed to complete a saccade to the correct location and were instructed to look **towards** the target (prosaccade) or **away** from the target (antisaccade). In the **memory-guided** task, two peripheral targets appeared sequentially. After the disappearance of the FP, participants were required to make two saccades, as accurately as possible, to the remembered locations of the peripheral targets in the same sequence in which they were presented [16]. Measurements including performance accuracy, variability, main sequence, inhibition (special for anti- and memory-guided saccade), working memory (special for memory-guided saccade), and other sensorimotor functions were then derived from the recordings and treated as features for further classification analysis.

Free viewing. Participants watched a series of five 1-minute video clips, with each clip consisting of a sequence of uncorrelated short video snippets of 2-4 seconds duration, chosen from a set of 70 snippets. Eye movements were recorded as described above for the right eye, and the participants were instructed to simply “watch and enjoy the clips” [17].

Psychometric tests. Participants completed 20 subtests in domains of attention and executive functioning, memory and learning, visuospatial processing, working memory, and math and reading ability [18]. The scores of those subtests were pre-processed for age correction and standardization [19] and then used as features for classification.

Diffusion tensor imaging (DTI). The corpus callosum was divided into six regions of interest (ROI). Three eigenvalues, average length and angle were obtained for each ROI averaged across all voxels in a given white matter tract [20,21], and used as input to the classifier.

2.2 Data classification

Single test classification. The data from saccadic eye movements, psychometric tests and DTI were analyzed using support vector machine-recursive feature elimination (SVM-RFE) [22]. SVM-RFE allows group classification and feature selection at the same time. Data from participants who finished all six tests were left out for testing and the remaining data was used for training with leave-one-out (LOO) as the cross validation method.

The free viewing data were first pre-processed into attentional eye traces. The *Itti-Koch* saliency model [23,24] was applied to obtain saliency maps for the visual properties of the scene elements looked at by each participant. Another top-down map was derived from the Gaussian-smoothed spatiotemporal gaze map of a group of healthy young adults (n=19) from USC [17]. The standardized values of those maps were extracted at each recording point of the eye traces, resulting in what we call “attentional eye traces” for each of the 70 video snippets, and then passed through two stacked tiled convolutional neural networks [25]. 70 weak classifiers were trained from the learned representations, after a dimension reduction step via L1-regularized logistic regression (LR). A second layer of LR classifier was built on top of them to obtain the final prediction.

Multi-test classification. To take advantage of data from different tests, we then performed a classification analysis on the dataset of the participants who finished all six tests (N = 46, see Table 1). An iterative train-test procedure was applied. During each iteration: one participant’s data set was excluded as test data; the remaining data were used for training with cross validation. Probabilities of a participant being identified as having FASD as predicted from every test were concatenated as input for training a LR classifier. The iteration terminated when every participant’s data was used as the test sample once and the accuracy was calculated based on the prediction labels acquired during testing.

3 Results

Classification accuracy of each test is summarized in Table 2. A classifier operating at chance level would achieve 52.2% correct (naïve Bayes). The combination of subset measurements from prosaccade, antisaccade, psychometrics, and free viewing resulted in the best multi-test classification accuracy of **84.8%**. The sensitivity and specificity of classification were 81.8% and 87.5% respectively.

Table 2: Number of participants for different tests

	Pro-saccade	Anti-saccade	M-G saccade	Free viewing	Psych	DTI
Test accuracy	69.6%	76.1%	65.2%	71.7%	78.3%	67.4%
Time	7min	7min	10min	5min	45min	60min
Required administration	Anyone trained to use Eyelink 1000				psychometrist or clinical psychologist	MRI technician and imaging center
Age restriction	≥ 3 yrs	≥ 6 yrs	≥ 6 yrs	≥ 2 mos	7~15 yrs	≥ 7 yrs
Total cost	\$50/h				\$200/h	\$530~540/h
Accessibility	****				**	*

During the classification process, we also discovered that features related to oculomotor control functions (such as accuracy, variability, main sequence), sensorimotor functions, inhibition, working memory, attention, math ability, and the posterior parts of the corpus callosum were distinguishing in the detection of the FASD group, suggesting a comprehensive profile of the impairments.

The six tests could be investigated from various aspects besides classification accuracy (See Table 2 for details). Free viewing and saccadic eye movements are the lowest-cost tests with the easiest accessibility and the least age restriction, whereas brain imaging is the most expensive paradigm with the most difficult accessibility.

4 Screening procedure

A high-throughput and low-cost screening procedure could be proposed at this point (see Figure 1). A participant is first assessed by the prosaccade and free viewing tests, with sensitivity 77.3%, specificity 79.2%, and accuracy 78.3%, for estimating the risk of having FASD. If the participant is classified as high risk of having FASD (FASD score returned by the classifier higher than 0.55 on a 0 to 1 scale), then a full diagnostic evaluation could be suggested. This is the most cost-effective approach, with the least age restriction and test load, and can be further evaluated from a “value of information” perspective of view [13]. For older children/youth, the addition of the antisaccade test and/or of a short battery of psychometric tests adds to the classification accuracy, leading to greater confidence that at-risk children/youth are not missed.

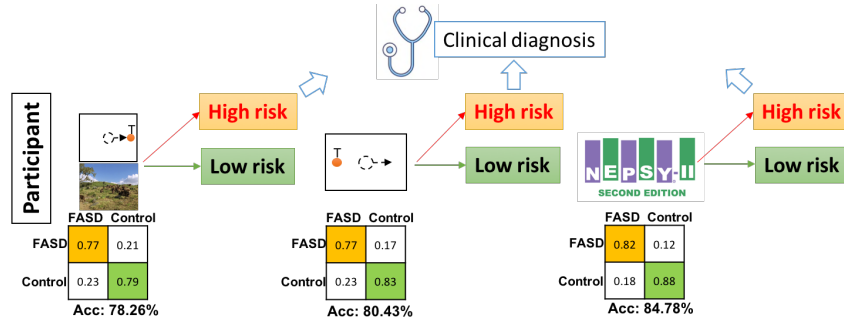


Figure 1: Screening procedure.

For an individual with FASD, the expected annual savings with prosaccade and free viewing tests only is about \$7,670 (see **Appendix** for details about the cost-benefit analysis). The probability of missing a patient in a consecutive 4-year period is less than 0.3% (0.23^4), which means that, with annual screening using a low-cost, high-capacity procedure, a child/youth with FASD can be discovered with a probability higher than 99.7% within 4 years. Thus the expected savings in 4 years is at least $\$7670(1 + 0.23 + 0.23^2 + 0.23^3) = \$9,933$. For our group of 46 children/youth with FASD in the testing group (the fraction of participants with FASD in the screened population p_F is 0.48), the expected value of individual annual savings multiplied by 46 is \$148,065 for a screening procedure composed of prosaccade and free viewing, \$151,892 with the addition of antisaccade, and \$187,220 with another addition of three psychometric subtests. Extrapolating this computation to a screened population of 1,000 individuals results in savings of more than 3 to 4 million dollars. If only children with known or suspected prenatal alcohol exposure or any suspected neurodevelopmental deficits were to be screened using our method, a reasonable estimation of p_F is likely to be higher than 0.48, leading to even more savings based on our model.

5 Conclusion

The proposed screening procedure consisting of three brief eye movement tests and three subtests from the psychometric battery achieved a mean sensitivity of 82%, a mean specificity of 88%, and an overall accuracy of 85%. It can be administered in under 1.5 hours. The method provides a quantitative and objective signature of the disorder, for each individual, along multiple dimensions that encompass a range of cognitive and oculomotor functions. It alleviates the need for screening or discrimination tools that strongly rely on demographic, behavioral, physical, psychometric, face morphometric analysis, and history of maternal alcohol consumption, though some of these could be included in the referral process to ensure a relatively large p_F . The screening procedure is also much easier to implement with early or pre- school age children compared to the standard diagnostic process. The data can be analyzed through the same machine learning pipelines and thus all the screening estimations are achieved through the same standards. Such a screening tool could be widely used at clinics, schools, or health units where young children are seen routinely, across different regions and areas, promoting communications within an interdisciplinary context. Early screening that streamlines a more in-depth clinical diagnostic process would have a substantial impact for children with FASD because it could lead to earlier intervention and treatment, which is crucial for neurodevelopmental disorders.

References

- [1] Jones K & Smith D. Recognition of the fetal alcohol syndrome in early infancy. *The Lancet* 1973; 302(7836): 999–1001.
- [2] May PA, Gossage JP, Kalberg WO, Robinson LK, Buckley D, Manning M, et al. Prevalence and epidemiologic characteristics of FASD from various research methods with an emphasis on recent in-school studies. *Developmental disabilities research reviews* 2009; 15(3): 176–192.
- [3] May PA, Baete A, Russo J, Elliott AJ, Blankenship J, Kalberg WO, et al. Prevalence and characteristics of fetal alcohol spectrum disorders. *Pediatrics* 2014; 134(5): 855–866.
- [4] Popova S, Lange S, Burd L, Rehm J. The burden and economic impact of fetal alcohol spectrum disorder in Canada. Centre for Addiction and Mental Health Toronto, Canada; 2015.
- [5] Popova S, Lange S, Burd L, Rehm J. Burden and social cost of fetal alcohol spectrum disorders 2016.
- [6] Chudley AE, Conry J, Cook JL, Looock C, Rosales T, LeBlanc N. Fetal alcohol spectrum disorder: Canadian guidelines for diagnosis. *Canadian Medical Association Journal* 2005; 172(5 suppl): S1–S21.
- [7] Peadon E & Elliott EJ. Distinguishing between attention-deficit hyperactivity and fetal alcohol spectrum disorders in children: clinical guidelines. *Neuropsychiatric Disease and Treatment* 2010; 6: 509-515.
- [8] Goh YI, Chudley AE, Clarren SK, Koren G, Orrbine E, Rosales T, et al. Development of Canadian screening tools for fetal alcohol spectrum disorder. *The Canadian journal of clinical pharmacology = Journal canadien de pharmacologie clinique* 2007; 15(2): e344–66.
- [9] Goh PK, Doyle LR, Glass L, Jones KL, Riley EP, Coles CD, et al. A Decision Tree to Identify Children Affected by Prenatal Alcohol Exposure. *The Journal of pediatrics* 2016; 177: 121–127.
- [10] Burd L, Klug MG, Li Q, Kerbeshian J, Martsolf JT. Diagnosis of fetal alcohol spectrum disorders: a validity study of the fetal alcohol syndrome checklist. *Alcohol* 2010; 44(7): 605–614.
- [11] Mutsvangwa TE, Meintjes EM, Viljoen DL, Douglas TS. Morphometric analysis and classification of the facial phenotype associated with fetal alcohol syndrome in 5-and 12-year-old children. *American Journal of Medical Genetics Part A* 2010; 152(1): 32–41.
- [12] Kalberg WO, May PA, Blankenship J, Buckley D, Gossage JP, Adnams CM. A practical testing battery to measure neurobehavioral ability among children with FASD. *The international journal of alcohol and drug research* 2013; 2(3): 51.
- [13] Feltham GA. The value of information. *The Accounting Review* 1968; 43(4): 684–696.
- [14] Munoz DP & Everling S. Look away: the anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience* 2004; 5(3): 218–228.
- [15] Coe BC & Munoz DP. Mechanisms of saccade suppression revealed in the anti-saccade task. *Phil. Trans. R. Soc. B.* 2017 Apr 19; 372(1718): 20160192.
- [16] Paolozza A, Rasmussen C, Pei J, Hanlon-Dearman A, Nikkel SM, Andrew G, et al. Working memory and visuospatial deficits correlate with oculomotor control in children with fetal alcohol spectrum disorder. *Behavioural brain research* 2014; 263: 70–79.
- [17] Tseng PH, Paolozza A, Munoz DP, Reynolds JN, Itti L. Deep learning on natural viewing behaviors to differentiate children with fetal alcohol spectrum disorder. In: *Intelligent Data Engineering and Automated Learning–IDEAL 2013* Springer; 2013.p. 178–185.
- [18] Korkman M, Kirk U, Kemp S. NEPSY-II: A developmental neuropsychological assessment. San Antonio, TX, Harcourt Assessment; 2007.
- [19] Paolozza A, Rasmussen C, Pei J, Hanlon-Dearman A, Nikkel SM, Andrew G, et al. Deficits in response inhibition correlate with oculomotor control in children with fetal alcohol spectrum disorder and prenatal alcohol exposure. *Behavioural brain research* 2014; 259: 97–105.
- [20] Paolozza A, Treit S, Beaulieu C, Reynolds JN. Response inhibition deficits in children with Fetal Alcohol Spectrum Disorder: Relationship between diffusion tensor imaging of the corpus callosum and eye movement control. *NeuroImage: Clinical* 2014; 5: 53–61.
- [21] Witelson SF. Hand and sex differences in the isthmus and genu of the human corpus callosum: a postmortem morphological study. *Brain* 1989; 112(3): 799–835.
- [22] Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. *Machine learning* 2002; 46(1-3): 389–422.

[23] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on Pattern Analysis & Machine Intelligence 1998; (11): 1254–1259.

[24] Itti L & Koch C. A saliency-based search mechanism for overt and covert shifts of visual attention. Vision research 2000; 40(10): 1489–1506.

[25] Ngiam J, Chen Z, Chia D, Koh PW, Le QV, Ng AY. Tiled convolutional neural networks. In: Advances in Neural Information Processing Systems; 2010. p. 1279–1287.

Acknowledgements: This work was supported by the National Science Foundation (CCF-1317433 and CNS-1545089), and Intel Corporation. The authors affirm that the views expressed herein are solely their own, and do not represent the views of the United States government or any agency thereof.

Appendix

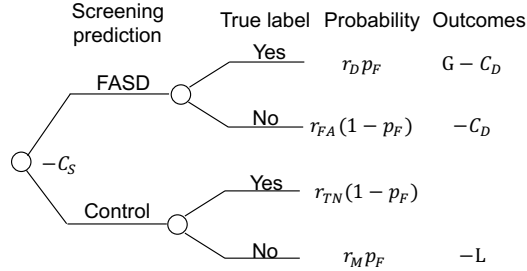


Figure 2: Model structure for cost evaluation of the screening procedure.

Figure 2 shows the model structure used for the annual cost-benefit analysis of the screening procedure for each individual. The probabilities in the model are derived from *a priori* probability of having FASD (p_F , which could be the fraction of participants with FASD in the screened population), and the confusion matrix (detection rate r_D , false alarm rate r_{FA} , true negative rate r_{TN} , and miss rate r_M , see Figure 1 for those values for different screening procedures). A general cost of the screening is denoted as C_S , of value smaller than \$50 using initially only prosaccade and free viewing (Table 2). Adding the antisaccade task does not increase this value since it can be done together with prosaccade and free viewing. The cost of screening will increase by less than \$200 when a short battery of psychometric tests is also included. A participant predicted to be at high-risk of having FASD will have a further clinical diagnostic evaluation, which costs less than \$4,000 (excluding medication, hospitalization and other non-diagnostic costs), denoted as C_D . If our prediction of FASD is correct, then we will have a gain G for early detection. Conversely, a loss L occurs when we miss a patient. The values for G and L are difficult to estimate, and no quantitative studies about them were found. However, a significant difference in the individual annual cost was reported for different severities of FASD [1]. Early detection could mean that the progress of the disorder is mitigated by providing a supportive, enriched environment for the child, and access to services for the family, and the benefit gained from it could be at least \$20,000 per year (difference of average annual health care costs between severe and mild FASD patients). Thus $G \geq \$20,000$. If the screening procedure can be applied annually, then the cost of missing a patient, L , should be less than G , under the assumption that not all missed mild children/youth with FASD will develop severe symptoms within the year before the next screening. The expected value of the model for an individual is computed as follows:

$$EV = r_D p_F (G - C_D - C_S) - r_{FA} (1 - p_F) (C_D + C_S) - r_{TN} (1 - p_F) C_S - r_M p_F (L + C_S)$$

For an individual with FASD, the expected annual savings with the aforementioned assumptions for screening with eye movement tasks only is computed as:

$$EV = 0.77(G - C_D) - 0.23L - C_S \geq 0.77 \times (20,000 - 4,000) - 0.23 \times 20,000 - 50 = 7,670$$

[1] Stade B, Ali A, Bennett D, Campbell D, Johnston M, Lens C, et al. The burden of prenatal exposure to alcohol: revised measurement of cost. Can J Clin Pharmacol 2009; 16(1): e91–e102.