



## Article

# A Prediction Model of Autism Spectrum Diagnosis from Well-Baby Electronic Data Using Machine Learning

Ayelet Ben-Sasson <sup>1,\*</sup>, Joshua Guedalia <sup>1</sup> , Liat Nativ <sup>1</sup>, Keren Ilan <sup>1</sup>, Meirav Shaham <sup>1</sup> and Lidia V. Gabis <sup>2,3,4</sup> 

<sup>1</sup> Department of Occupational Therapy, Faculty of Social Welfare and Health Sciences, University of Haifa, Haifa 3498838, Israel; liatnativ@gmail.com (L.N.)

<sup>2</sup> Maccabi Healthcare Services, Tel-Aviv 6812509, Israel; gabis\_l@mac.org.il

<sup>3</sup> Pediatrics, Faculty of Medicine and Health Sciences, Tel-Aviv University, Tel-Aviv 6997801, Israel

<sup>4</sup> Keshet Autism Center Maccabi Wolfson, Holon 5822007, Israel

\* Correspondence: asasson@univ.haifa.ac.il; Tel.: +972-48420364

**Abstract:** Early detection of autism spectrum disorder (ASD) is crucial for timely intervention, yet diagnosis typically occurs after age three. This study aimed to develop a machine learning model to predict ASD diagnosis using infants' electronic health records obtained through a national screening program and evaluate its accuracy. A retrospective cohort study analyzed health records of 780,610 children, including 1163 with ASD diagnoses. Data encompassed birth parameters, growth metrics, developmental milestones, and familial and post-natal variables from routine wellness visits within the first two years. Using a gradient boosting model with 3-fold cross-validation, 100 parameters predicted ASD diagnosis with an average area under the ROC curve of 0.86 (SD < 0.002). Feature importance was quantified using the Shapley Additive explanation tool. The model identified a high-risk group with a 4.3-fold higher ASD incidence (0.006) compared to the cohort (0.001). Key predictors included failing six milestones in language, social, and fine motor domains during the second year, male gender, parental developmental concerns, non-nursing, older maternal age, lower gestational age, and atypical growth percentiles. Machine learning algorithms capitalizing on preventative care electronic health records can facilitate ASD screening considering complex relations between familial and birth factors, post-natal growth, developmental parameters, and parent concern.

**Keywords:** autism spectrum disorders; development; screening; machine learning; electronic health records



**Citation:** Ben-Sasson, A.; Guedalia, J.; Nativ, L.; Ilan, K.; Shaham, M.; Gabis, L.V. A Prediction Model of Autism Spectrum Diagnosis from Well-Baby Electronic Data Using Machine Learning. *Children* **2024**, *11*, 429. <https://doi.org/10.3390/children11040429>

Academic Editor: Zbigniew Starosolski

Received: 12 March 2024

Revised: 28 March 2024

Accepted: 29 March 2024

Published: 3 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Autism spectrum disorders (ASD) manifest through social communication deficits and atypical behaviors, as listed in The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) [1] and it may affect 1:36 individuals [2]. Nonetheless, the prevalence is highly variable between countries, with lower prevalence in lower income communities [3,4], depending primarily on the rate of detection and diagnosis. Recent rates of diagnosis in Israel are 1% among 2–3-year-olds and 2% in 4–5-year-olds [5], which is close to the global median of 1% [4] and higher than previous reports [3]. Early intervention administered before three years of age leads to significantly greater social communication, adaptive functioning, and cognitive competencies [6,7]. The American Academy of Pediatrics and other health organizations recommend [8–10] early screening for ASD at 18 and 24 months during dedicated well-baby visits to promote the initiation of early intervention and improve outcomes [11–13]. In Israel, the well-baby check-ups are performed at a national level and regulated by the Ministry of Health (MoH). The age distribution of developmental milestones examined during the scheduled well-baby visits has been recently studied in this national database [14,15]. The current study focused on a model detecting children with elevated ASD markers in routine well-baby check-up records to provide higher accuracy of ASD detection during developmental screening. Such a model

can, in the long run, be implemented to alert healthcare providers during routine checkups to follow-up with ASD-specific screening tools.

Behavioral markers linked to ASD can be observed during the first and second years of life, even in low-risk populations (e.g., [7,16–19]), yet the average age of diagnosis worldwide remains above three years of age [20,21]. One of the barriers to early detection of ASD is the great heterogeneity in the ‘red flags’, onset, and progression of symptoms, and in the divergence of developmental trajectories in ASD [18,22]. Early detection of ASD relies upon behavioral observations in the social communication and sensory motor domains (e.g., [7,16,23]), and on observing developmental trajectories [7]. Deviations in motor and social communication development significantly differentiated the ASD group from typically developing children and children with other delays during the first [22,24] and second year of life [18,25]. Some early ASD signs can be easily missed by a healthcare provider during a short checkup visit due to their gradual deviation or subtlety, such as social reciprocity, social smile, and eye contact [11,22]. Furthermore, disparities in healthcare screening resources and in provider awareness of individual developmental differences [7,26] challenge early detection. To facilitate and adapt the screening level to different trajectories, the screening procedure needs to match the risk level of the child [27]. Including known risk factors and individual differences may highlight specific variations and increase the level of awareness in particular populations, such as those with prematurity or familial factors [28–31].

Prematurity, like other factors, interacts with additional ASD risk factors, such as cesarean delivery, parental age, and birth complications; therefore, there is a need for an integrative/cumulative approach to risk estimation. These non-linear relationships can be detected using machine learning (ML) analysis of big data, as conducted in the present study. Research shows the value of ML model risk score predictions for accurately stratifying clinical populations [32] and can guide the provider to personalize the checkup according to their risk level. In the ASD literature, ML has been applied for various purposes, including differentiating subgroups [33] and, specifically, parsing the behavioral phenotype variance [34]. Deep learning methods have been applied for ASD detection based on motor abnormalities [35] and the prediction of behavioral intervention efficacy from patient data [36]. Previous electronic health record (EHR) data analysis by ML algorithms differentiated clusters of ASD patients with common co-morbidities (e.g., [37–39]). Furthermore, ML tools can predict ASD likelihood from family medical history in EHRs [40,41]. Such tools address the need for a cost efficient and accurate process which considers neurodevelopmental heterogeneity among children [42].

The well-baby clinic system in Israel provides nation-wide developmental checkups from birth to six years and offers an opportunity for implementing universal ASD screening [43]. The database was digitized to build and test a ML model that can identify significantly deviant early milestones linked to ASD from a child’s EHR at early stages of development and include specific background information, such as pregnancy and familial history. In the current study, we aimed to identify the most significant features linked to a subsequent diagnosis of ASD from a large database of EHRs using ML tools. The ML prediction of ASD was also verified by examining the proportion of children that were flagged as false positives by the model but manifested significant developmental delays. Since the analysis is driven by a large and diverse population, it may provide an applicable model for developmental surveillance and improve ASD detection below the age of two years.

## 2. Materials and Methods

### 2.1. Study Design and Ethics

This was a health records retrospective cohort study. This study was approved by the MoH Helsinki (#15/2021) committee. The need for consent was waived by the ethics committee. EHR data were recorded by the providers of the Israeli national well-baby developmental surveillance program called Tipat Halav (TH, “Drop of Milk”) managed

by the MoH. Data were extracted and anonymized by the TIMNA project personnel (Israel Ministry of Health's Big Data Platform). Data were extracted until 31 December 2021. Researchers accessed de-identified participants' data through this dedicated secured MoH platform.

## 2.2. Data Source

The EHRs analyzed in this study reflected a cohort followed by the national well-baby MoH clinics. The MoH is the main provider of developmental checkups at 503 clinics across the country, covering 75% of the Israeli population. The remaining 25% are followed by two Health Maintenance Organizations according to the same national guidelines. The services include immunizations, growth, and developmental surveillance, as well as medical and environmental risk detection for the mother and child. The services are provided by nurses with specific training in maternal and child health and a developmental pediatrician. There are approximately 11 visits from birth to age 6 years, each with a set of actions based on the child's age and risk status. The nurses enter the data in a digital EHR.

The developmental protocol of TH includes a checkup of progress in 60 age-related milestones across gross motor, fine motor, language, and personal–social developmental domains [43]. The milestones included in the TH protocol are based in part on the Denver Developmental Screening Test (DDST) [44]. At each of the nine age bins, from six weeks to six years, the nurses check between 4–9 milestones. There are 38 milestones tested up to 24 months. Performance results for each milestone are recorded. During developmental checkups, parents are prompted to report their concerns regarding the child's development and/or hearing. An ASD-specific screening tool is not integrated into the checkups.

The EHR inclusion criteria were: (1) children born between 1 January 2014–31 December 2019; (2) the availability of at least two years of computerized data. This led to a database of 780,610 EHRs. In order to develop a model that minimizes data leakage from diagnosis to predictors, we included: (1) as predictors in the model measurements up to two years of age and (2) as the outcome metric of ASD, only children with an ASD diagnosis reported after two years of age. The average age at which diagnosis was recorded for the full ASD group was 37.5 months, SD = 13.3 months (13.22 months–7.72 years).

## 2.3. Case Definitions

The prediction of an ASD diagnosis was based on a diagnosis of ASD reported in the child's EHR, either by an International statistical classification of diseases and related health problems (ICD-9 [45]) code of ASD or reports in provider notes after the age of two years. Following ICD-9 case identification, texts for the rest of the population were searched for ASD-related keywords and verified manually for indication of ASD in the child (as opposed to family members). Children with other conditions with the same acronyms as Pervasive Developmental Disorders (PDD) and ASD were excluded (e.g., Atrial Septal Defect: ASD). Of the total sample, 1275 children were labeled as ASD-diagnosed, 722 children were identified by their ICD-9 diagnostic code, and an additional 553 were identified from the clinical textual notes. In total, 1163 were diagnosed after the age of 2 years.

EHRs of the Typically Developing (TD) group were defined as records of children without an ASD or other pervasive developmental and medical conditions diagnosis. The TD group excluded EHR records of children with another non-ASD pervasive medical or developmental condition. Since there are multiple etiologies linked to ASD, such as Fragile X, Tuberous Sclerosis, metabolic disorders, and more than 200 single genetic disorders, we did not exclude specific etiologies except Down syndrome, as explained below. Global developmental delay and intellectual disability [46] can occur as a comorbidity of ASD or differential diagnosis and are defined as significant delay in more than one developmental domain with performance of at least two standard deviations below age norms. When reported as a comorbidity, ASD symptoms should prevail, according to DSM-5 ASD criteria—E (1). We excluded children who only had these diagnoses reported for them,

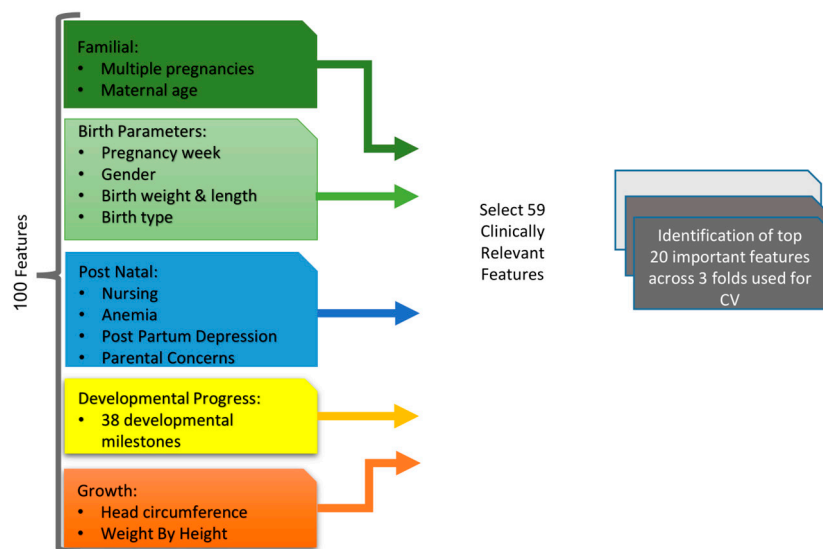
since our goal was to build a system that can assist providers in noticing young children whom they routinely miss.

The only diagnosis that may be causative and linked to ASD but was excluded from our ASD analysis is Down syndrome. Although children with Down syndrome may also have ASD, we decided to exclude this diagnosis since this etiology is overtly obvious from birth and the routine checkups for infants with Down Syndrome are adapted to their different trajectories.

The rationale was to focus the model design on the prediction of ASD and eliminate the prediction of conditions that the provider/family are aware of very early on (e.g., Down syndrome). The list of ICD-9 diagnosis was reviewed by clinical experts, and a list of congenital or other conditions highly associated with developmental challenges was compiled to reach higher level of comparison norms in the control group. Based on this list, 4577 children were excluded from the TD group due to musculoskeletal and nervous system disorders (20.3%), chromosomal abnormalities (4.5%), hearing loss (3%), convulsive disorders (1.7%), Infantile Cerebral Palsy (0.8%), and Spina Bifida (0.6%).

### 2.4. Data Preparation

The EHR structure and, consequently, the database, includes all documented information for a child and mother across all baby wellness visits. As such, the variables are a mix of structured (e.g., concern for development Yes/No, ICD-9 codes) and text data (e.g., provider notes), as well as fixed data points (e.g., birth weight) versus temporal data (e.g., head circumference measurements). For the structured features, there were continuous variables, such as birth weight in kg, and nominal variables, such as type of birth delivery (see Figure 1 for Model process).



**Figure 1.** The iterative model design process from data mining to model testing.

As part of feature engineering, the distribution of variables in the database was investigated. Based on their distribution, several variables were recoded to enable their meaningful integration in the model. Most features included a considerable number of missing values (see Tables 1–3). Median imputation was applied in the case of continuous features, and a missing category was applied for categorical features. To address age differences in timing and number of visits, we calculated first and last measurements, median value, and standard deviation for longitudinal features. Each of the 38 developmental milestones entered the model as two features: only failed or ever passed. This configuration enabled the model to (1) account for cases in which a child had a milestone tested multiple times and (2) to differentiate not performing from a missing entry.

**Table 1.** Comparison of child and family demographic between groups.

Characteristics	TD (N = 774,778)	ASD (N = 1163)	Statistics
Year of child’s birth			$\chi^2(5) = 106.49^{***}$ , $\varphi_c = 0.012$
2014	124,401 (16.1%) <sub>a</sub>	174 (15%) <sub>a</sub>	
2015	126,679 (16.4%) <sub>a</sub>	230 (19.8%) <sub>b</sub>	
2016	129,421 (16.7%) <sub>a</sub>	278 (23.9%) <sub>b</sub>	
2017	131,721 (17%) <sub>a</sub>	212 (18.2%) <sub>a</sub>	
2018	132,522 (17.1%) <sub>a</sub>	182 (15.6%) <sub>a</sub>	
2019	130,034 (16.8%) <sub>a</sub>	87 (7.5%) <sub>b</sub>	
Mother birth country			$\chi^2(5) = 375.48^{***}$ , $\varphi_c = 0.023$
Israel	631,588 (81.5%) <sub>a</sub>	816 (70.2%) <sub>b</sub>	
Russia	43,766 (5.6%) <sub>a</sub>	209 (18%) <sub>b</sub>	
Europe	1750 (1.6%) <sub>a</sub>	6 (0.5%) <sub>b</sub>	
Ethiopia	11,114 (1.4%) <sub>a</sub>	40 (3.4%) <sub>b</sub>	
North America	8799 (1.1%) <sub>a</sub>	3 (0.3%) <sub>b</sub>	
Other	8360 (1.3%) <sub>a</sub>	18 (1.5%) <sub>a</sub>	
Missing	58,401 (7.5%)	71 (6.1%)	
Mother Employment Status			$\chi^2(2) = 11.34$ , $\varphi_c = 0.005$
Employed	334,375 (43.2%) <sub>a</sub>	512 (44%) <sub>a</sub>	
Unemployed	164,258 (21.2%) <sub>a</sub>	271 (23.3%) <sub>a</sub>	
Student	32,129 (4.1%) <sub>a</sub>	27 (2.3%) <sub>b</sub>	
Missing	244,016 (31.5%)	353 (30.4%)	

Note. Continuous variables were compared using independent samples *t*-test, while nominal variables with Chi square or Fisher exact tests. Missing data were not included in the comparisons. ASD: Autism spectrum disorder. TD: Typical development. <sub>a,b</sub> Groups with different subscripts are significantly different based on pairwise adjusted Z-tests.  $^{***} p < 0.001$ .

**Table 2.** Comparison of familial and birth parameters between groups.

Characteristics	TD (N = 774,778)	ASD (N = 1163)	Statistics
Mother’s age <sub>c</sub> Mean (SD, Min–Max)	30.02 (5.71, 15–55)	31.43 (5.65, 19–52)	$t(767,831) = -8.27^{***}$ , $d = 0.24$
Missing	8688 (1.12%)	2 (0.17%)	
Birth weight (Kg) Mean (SD, Min–Max)	3.21 (0.51, 0.4–6)	3.15 (0.58, 0.72–4.77)	$t(747,034) = 3.96^{***}$ , $d = 0.1$
Missing	28,871 (3.72%)	34 (2.92%)	
Pregnancy week <sub>c</sub> Mean (SD, Min–Max)	39.09 (1.8, 23–43)	38.65 (2.07, 25.2–42.3)	$U = 437,787,198^{***}$ , $d = 0.004$
Missing	29,465 (3.8%)	32 (2.75%)	
Males <sub>c</sub> N (%)	398,069 (51.4%)	911 (78.3%)	$^{***}$ , $\varphi = 0.021$
Preterm births (pregnancy week < 37)	54,937 (7.1%)	132 (11.3%)	$^{***}$ , $\varphi = 0.006$
Missing	41,496 (5.4%)	52 (4.5%)	
Birth type			$\chi^2(2) = 86.04^{***}$ , $\varphi_c = 0.011$
Spontaneous	539,004 (69.6%) <sub>a</sub>	693 (59.6%) <sub>b</sub>	
Cesarean	129,685 (16.7%) <sub>a</sub>	309 (26.6%) <sub>b</sub>	
Instrumental	37,232 (4.8%) <sub>a</sub>	71 (6.1%) <sub>b</sub>	
Missing	68,857 (8.9%)	90 (7.7%)	
Multiple Pregnancy			$\chi^2(3) = 18.87^{***}$ , $\varphi_c = 0.005$
Single	741,871 (95.8%) <sub>a</sub>	1084 (93.2%) <sub>b</sub>	
Twin	31,783 (4.1%) <sub>a</sub>	76 (6.5%) <sub>a</sub>	
Triplet	1058 (0.1%) <sub>a</sub>	3 (0.3%) <sub>a</sub>	
Other	66 (0.001%) <sub>a</sub>	0 (0%) <sub>a</sub>	

Note. Continuous variables were compared using independent samples *t*-test, while nominal variables with Chi square or Fisher exact tests. Missing data were not included in the comparisons. ASD: Autism spectrum disorder. TD: Typical development. <sub>a,b</sub> Groups with different subscripts are significantly different based on pairwise adjusted Z-tests. <sub>c</sub> These features were among the top important features in the final model.  $^{***} p < 0.001$ .

**Table 3.** Comparison of post-natal parameters between groups.

Characteristics	TD (N = 774,778)	ASD (N = 1163)	Statistics
Postpartum depression			$\chi^2(3) = 7.19, \varphi_c = 0.004$
None	539,565 (69.6%) <sup>a</sup>	817 (70.2%) <sup>a</sup>	
Mild	11,941 (1.5%) <sup>a</sup>	16 (1.4%) <sup>a</sup>	
Moderate	2306 (0.3%) <sup>a</sup>	8 (0.7%) <sup>b</sup>	
Severe	1062 (0.1%) <sup>a</sup>	3 (0.3%) <sup>a</sup>	
Missing	219,904 (28.3%)	319 (27.4%)	
Ever had Concern for Development <sup>c</sup>	71,516 (9.2%)	820 (70.5%)	***, $\varphi = 0.08$
Missing	10,162 (1.3%)	8 (0.7%)	
Ever had Hearing Suspicion	22,319 (2.9%)	272 (23.4%)	***, $\varphi = 0.047$
Missing	16,200 (2.1%)	12 (1%)	
Ever Nursing <sup>c</sup>	584,345 (75.4%)	708 (60.9%)	***, $\varphi = -0.014$
Missing	11,830 (1.5%)	13 (1.1%)	
Ever had Anemia	41,589 (5.4%)	58 (5%)	
Missing	388 (0.1%)	0 (0%)	

Note. Continuous variables were compared using independent samples *t*-test, while nominal variables with Chi square or Fisher exact tests. Missing data were not included in the comparisons. TD: Typical development. ASD: Autism spectrum disorder<sup>a,b</sup> Groups with different subscripts are significantly different based on pairwise adjusted Z-tests. <sup>c</sup> These features were among the top important features in the final model. \*\*\*  $p < 0.001$ .

## 2.5. Analysis

### 2.5.1. ML Modeling

Gradient boosting (CatBoost) method with a 3-fold cross validation was applied, a method implemented in EHR research [32,41]. Gradient boosting is a ML method that is a model evaluation technique where the dataset is divided into three subsets and the model is trained and tested three times, with two subsets for training and one for testing each time, to assess its performance across different data partitions. This method builds multiple decision trees, where each successive tree is fine-tuned to focus on the errors of the previous trees. The ensemble of trees is then used for the final prediction, working in a stage-wise manner rather than a parallel manner [47]. The high-risk group was defined as a predicted risk score above the Youden Threshold, an objective measure that sets a cutoff point that optimizes the sensitivity and specificity of the model [48]. The model was evaluated using AUC and its corresponding accuracy metrics. The model was designed by the team with the guidance of the third author, an expert in applying advanced machine learning techniques to predict outcomes from medical data.

We used the SHapley Additive exPlanations (SHAP) tool to quantify feature importance. The tool is employed to provide a comprehensive method for interpreting the results of ML models, attributing values to features to quantify their respective impacts on individual predictions. SHAP is a method used to explain the output of machine learning models by attributing the importance of each feature to the model’s predictions [49].

### 2.5.2. Group Comparisons

The ASD and TD groups were compared across child and family characteristics, birth parameters, and familial and post-natal variables excluding missing values. For nominal variables, Chi square or Fisher’s exact tests were conducted with their corresponding measures of association, Cramer’s V, or Phi. Pairwise-adjusted Z-tests were used for variables with more than two categories. Normality tests showed that most continuous variables were normally distributed ( $p > 0.05$ ), except for pregnancy week ( $p < 0.001$ ). For normally distributed continuous variables, independent samples *t*-tests with Cohen’s *d* were conducted. Mann–Whitney U test compared pregnancy week between groups. A *p*-value threshold of 0.001 was applied given the large sample size [50] and multiple number of comparisons.



### 3. Results

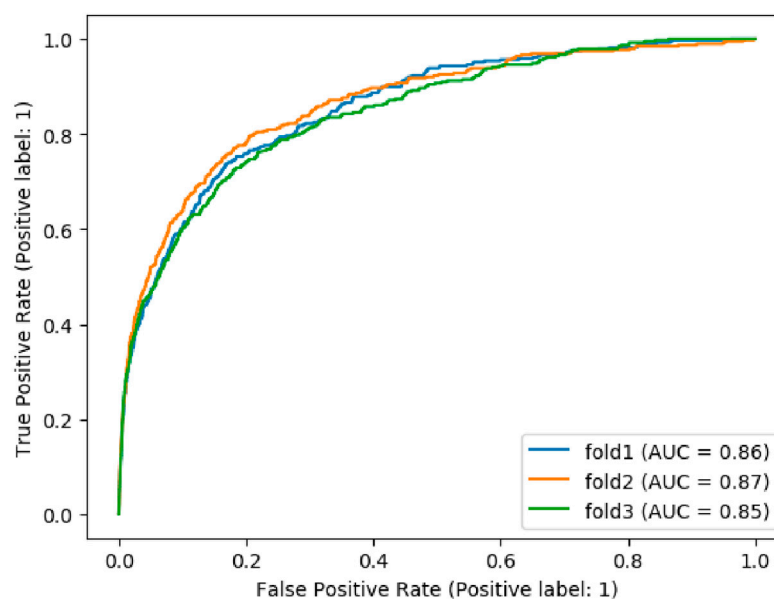
#### 3.1. Descriptive Statistics

The model analyzed EHRs of 774,778 TD children and 1163 ASD children (males 51.4%, 78.3% respectively). Table 1 presents the comparison of child and mother demographic features of TD versus ASD children. The TD group is equally distributed between birth years, while for the ASD group, as expected, there is a smaller proportion in 2019. There is a higher rate of ASD diagnosis for mothers born in Russia and Ethiopia. Table 2 presents familial and birth parameters, and Table 3 shows the post-natal features which were included in the training of our model. Tables show a significantly higher rate of preterm births in the ASD group, with older mothers, lower gestational week, lower birth weight, lower proportion of nursing, and much higher rates of parental concerns.

In the ASD group, there was an average of 13.7 visits overall (SD = 3.18; e.g., vaccination, growth, development) which was significantly ( $t(775,939) = -23.05, p < 0.001, d = 0.7$ ) higher than for the TD group, with an average of 11.44 visits (SD = 3.26). Up to two years of age, the average age at the first visit was 0.85 months (SD = 1.37), and at the last visit, 18.94 months (SD = 2.85 months). The EHR visit ages indicated that, following the age of two years, the average age at the first visit was 28.19 months (SD = 0.26), and at the last visit, 37.37 months (SD = 13.93).

#### 3.2. ML Model Predicting Risk Score for ASD

The data initially included 100 features, from which we selected 59 clinically relevant features to build a 3-fold ML model (see Figure 1 and Table S1). All three folds showed similar results, with an average Area Under the ROC Curve (AUC) = 0.86 (SD < 0.009; see Figure 2) and corresponding sensitivity = 0.75 (SD = 0.03) and specificity = 0.81 (SD = 0.02). The average AUC for the training data was higher than the test (0.88, SD = 0.006), indicating that the model is slightly overfit to the training data, as expected (see Table S2 for AUC by fold). The full tested model classified a total of 81.4% children ( $n = 631,308$ ) as low risk, and 18.6% children ( $n = 144,632$ ) as high-risk. This high-risk group included 874 children with ASD, which is an incidence rate of 0.006, 4.3-fold higher than the ASD incidence in the entire cohort (0.001). We also present results for different distributions between sensitivity and specificity based on changing multiples of the Youden Threshold (see Table S3). Notable is the decrease in rates of false positive cases as threshold increases. In addition, comparing results from different methods showed overall stable AUC with a small variance between methods (see Table 4).



**Figure 2.** ROC curves for the 3 folds predicting ASD diagnosis.

**Table 4.** Comparison of 3-fold cross-validation using different methods.

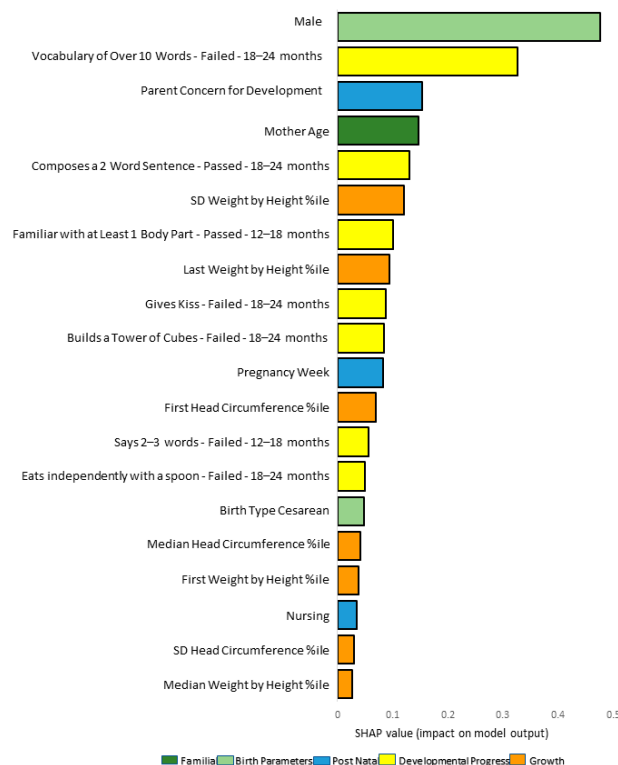
Method	Average AUC	SD	TP	FP	TN	FN	Sensitivity	Specificity
Logistic regression	0.86	0.010	897	159,472	615,305	266	0.77	0.79
Random Forest	0.84	0.009	861	166,793	607,984	302	0.74	0.78
Naive bayes	0.83	0.011	888	185,027	589,750	275	0.76	0.76

We reviewed the top 20 important features based upon their SHAP values within each fold. There was a high level of overlap between the top 20 features, with 15 (65%) in all three folds, 6 (25%) in two folds, and 3 (12.5%) unique to a specific fold, out of the total 24 unique features across the three folds (see Table S1).

3.3. Results from a Representative Fold

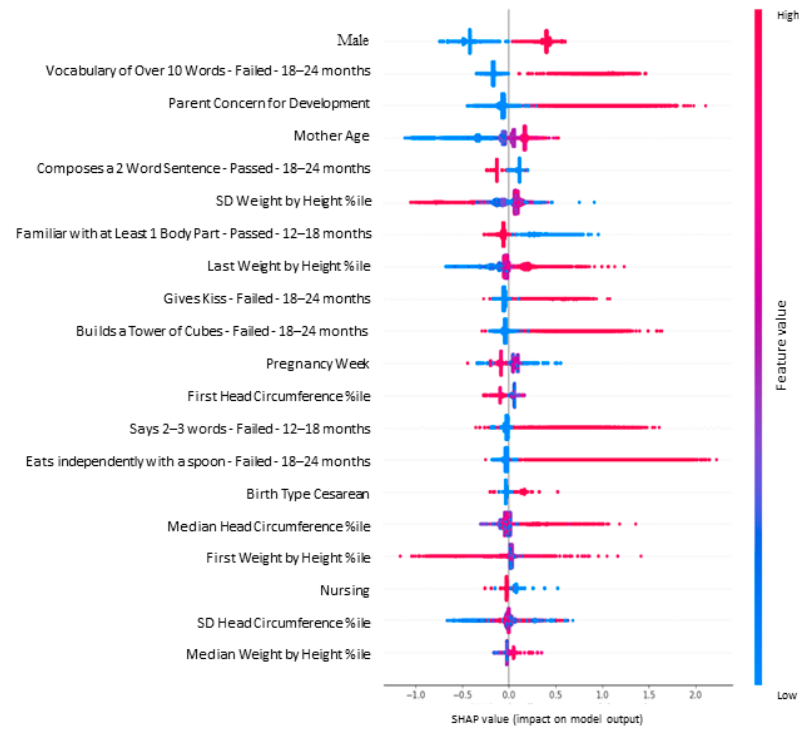
We present a single fold as a representative model (see Figures S1 and S2 for the distribution of features in other folds). Model summary plots show the relative importance of 20 features and their distribution for a representative fold (see Figures 3 and 4). Seven of the important features are related to developmental milestones from different domains. Fisher’s Exact tests showed significant differences in ( $p < 0.001$ ) the rates of attaining seven important model milestones in the ASD group (see Figure 5):

1. has a vocabulary over 10 words (language domain), tested 18–24 months;
2. builds a tower from cubes (fine motor domain), tested 18–24 months;
3. gives kiss (personal–social domain), tested 18–24 months;
4. eats independently with spoon (personal–social domain), tested 18–24 months;
5. knows at least one body part (language domain), tested 12–18 months;
6. composes 2-word sentences (language domain), tested 18–24 months);
7. says 2–3 words (language domain), tested 12–18 months.

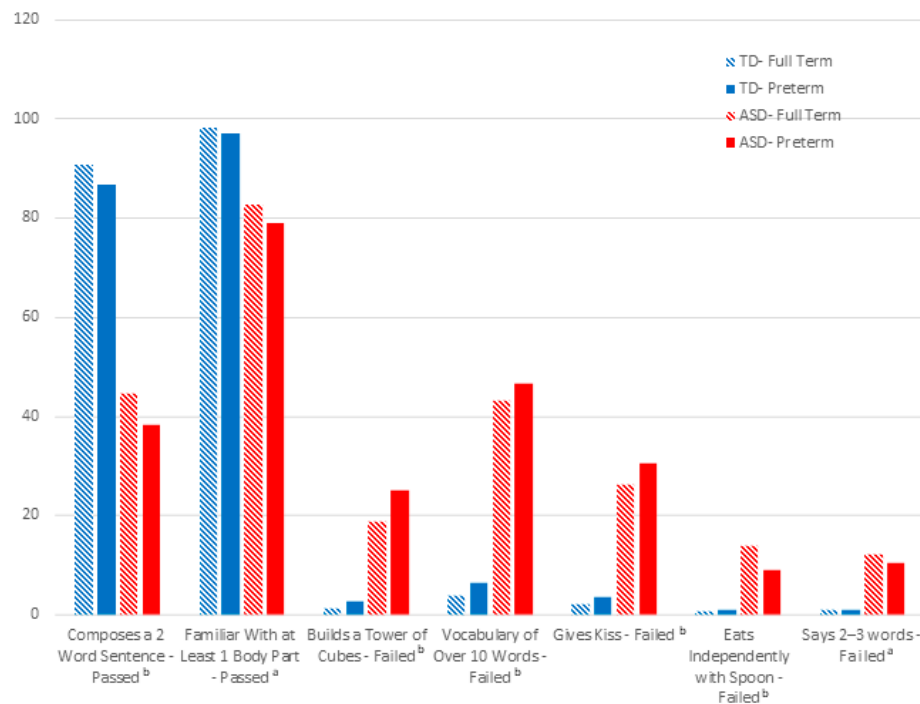


**Figure 3.** Feature importance bar plot for a representative fold from the full model. Note. The feature importance plot presents the importance of the features with their contribution to the model prediction as calculated using SHAP features. %ile: Percentile.





**Figure 4.** SHAP summary plot for a representative fold from the full model. Note. This plot represents the direction of the contribution of the individual values for each feature, to the right of the grey line indicates higher risk for ASD vs. to the left, indicating lower risk. The red reflects larger values, and the blue reflects lower values. %ile: Percentile.



**Figure 5.** Comparison of percentages between groups considering prematurity for the seven important milestones consistent across folds. Note. One milestone entered as passed and the others as failed. All milestones in both ASD and TD showed no statistically significant difference between full term and preterm children. <sup>a</sup> 12–18 months. <sup>b</sup> 18–24 months.

As can be seen in Figure 5, the group differences in the seven important milestones were not accounted for by prematurity; in other words, group comparisons were also significant when restricted to preterm versus full term samples ( $p < 0.001$ ).

The remaining 13 features and their direction of association with ASD in this fold were as follows (8 of which overlap between folds. Features 10–13 appeared in the presented fold and not in all 3 folds):

1. boy
2. more parental developmental concerns
3. never nursing
4. advanced maternal age
5. lower pregnancy week (prematurity) (In some folds, pregnancy week was entered, and prematurity status in others)
6. lower weight by height percentile at first visit
7. higher weight by height percentile at last visit
8. larger weight by height percentile SD (larger variance across visits)
9. higher head circumference percentile at first visit
10. cesarian birth
11. higher head circumference percentile median (across visits)
12. larger head circumference percentile SD
13. higher weight by height percentile median (across visits)

#### 3.4. Subgroup ML Analyses by Sex and Birth Year

Running a 3-fold cross validation model with boys only yielded a lower average AUC of 0.82 (SD = 0.03), with a sensitivity of 0.68 (SD = 0.02) and specificity of 0.82 (SD = 0.04). The average AUC for the training data (0.85, SD = 0.01) was slightly higher than the test (see Table S2 for AUC by fold). The 3 folds of the boys-only model had 12 overlapping features between folds out of the 29 unique features that appear in the 20 most important features in each fold. Of these, 10 features overlapped with the consistent features across the full model's folds (see Table S4 and Figures S3–S5 for details of the boys-only model). Birth weight was a unique consistent feature for the boys-only model.

In addition, to verify stability of the results, we evaluated a model without children born in the year 2019 and saw stable results in the model's AUC (see Table S4 for details of the model without 2019). The 2019-born children had a lower likelihood of ASD, therefore it was important to examine their impact on false positive results.

#### 3.5. Validity of the Model's High-Risk Group

To evaluate the high-risk status (i.e., false positive and true positive) identified by the above 3-fold model, we examined the rate of developmental provider referrals in the high risk versus low-risk groups (i.e., false negative and true negative). Provider's developmental referral in the EHR was defined as a reported provider's referral of the child for an extensive developmental assessment due to developmental delay. This referral variable was not included as a feature in the model and was therefore an external validity indicator of the model's classification. Those who the model identified as high-risk had a 3.29 odds ratio (95% CI 3.18–3.39) of being referred for developmental delay than the low-risk group.

## 4. Discussion

This study demonstrates how ML models can utilize existing routine preventative care data to estimate a child's ASD risk prior to the age of 2 years. The model predicting early ASD signs capitalized on the ongoing recording of real-life data by baby wellness providers, leading to a rich array of contributing real-world clinical measurements, including biometric predictors, such as familial, birth, post-natal, growth, and clinical predictors of developmental progress. At the healthcare policy level, this evidence calls for integrating technology to facilitate ASD detection in routine care. This is crucial for earlier detection of

ASD globally, overcoming screening barriers related to access to high quality healthcare. The model detected 75% of ASD children and remained relatively stable in performance when only boys were included. Such a system can flag risk factors in a child's early profile and facilitate a formal ASD screening during routine checkups [29]. While the model tested was slightly overfit to training data, it displayed similar results. Furthermore, this study presents the impact of different thresholds upon the balance of sensitivity and specificity. As threshold increases, the number of false positive cases decreases at the cost of fewer true positives. Healthcare providers can select different thresholds to support their context-specific policy and decision making. Future research with ample data would enable employment of methods such as CatBoost overfit detection during the training process to prevent overfitting. The current model's contributing factors represented well-known features—being a boy, lower pregnancy week, older mother's age, parental developmental concern, delayed language and personal–social development [51,52]—and less known features—never nursing [53], fine motor delay [54], and higher growth percentiles [55]. While the latter have some supporting evidence, they are not well-established factors. The model highlighted the most significant atypical developmental features among those that are well-known. Such discoveries are made possible by ML algorithms learning the interactions between features and improving the power of features by specific combination of familial data and examination of the child. The early ASD diagnosis process must start with identifying a child's biometric and clinical data. Integrating ML tools as decision support systems in EHRs offers a timely efficient way to alert providers regarding a young child's ASD risk.

Of the 38 milestones checked up to 24 months, six were important features consistent across the model folds. These milestones reflect a mixture of the language (three milestones), personal–social (two milestones) and fine motor (one) domains. Most were from the 18–24 months checkup, but one was earlier, from 12–18 months (i.e., knows at least one body part). While language expression and comprehension delay are not a diagnostic feature of ASD, they appear to reflect the child's social communication challenges early on and are easier for providers and parents to observe. This is consistent with evidence of robust divergence in ASD from 14–24 months, specifically in language and social development [18,22,24]. These important milestones were identified by the model via their interactions among themselves and with other features. Nonetheless, differences between ASD and TD groups on each of the six milestones were not explained by prematurity or by sex differences, indicating that they are a unique combination associated with the ASD status in this population. The findings also demonstrate that about 50% of ASD children do not fail these milestones during the second year of life. Thus, ASD-specific screening is needed at later ages to detect children without early developmental concerns.

Some milestones from the second year of life which appear in early screeners for ASD did not enter the model including eye contact, response to name, and expressing needs. It is possible that these markers had a weaker signal since they require closer clinical observation than a brief surveillance checkup and/or warrant a quality measuring scale rather than a pass or fail mark. Further improvement of the developmental protocol to target ASD-specific markers is warranted.

Implementing the model with subgroups of the population yielded relatively stable results. For the boys-only model, 10 of the consistently important features overlapped with the consistent ones in the full-model. Birthweight was only a consistent feature across the boys-only model, supporting evidence for differences in birth parameters of boys versus girls with ASD [28]. Since most of the ASD sample comprised males, there is need to replicate the model with a large sample of girls to detect different early "red flags".

There were strikingly higher rates of ASD among mothers born in Russia and, to some degree, among those from Ethiopian origins relative to their proportion in the population (see Table 1). Ethnicity and socioeconomic status are inconsistently linked with lower rates of ASD [4]. In contrast with our findings, lower rates of ASD were reported in lower socioeconomic status (SES) groups [3] and, specifically, in non-native Israeli families, including

Ethiopian ones [56]. Previous research differs dramatically in terms of ascertainment and publication years. Higher rates of early ASD diagnosis in immigrant communities may relate to a combination of biological risk factors, increased discrimination, isolation and stress levels, and poor prenatal care, as well as provider biases in immigrant groups. Future research into the healthcare and child contributing factors is critical to ensure equality in early detection.

While our study does not explicitly focus on the disparities between developed and developing countries in the detection of ASD, it is crucial to consider how such a model would perform in diverse healthcare contexts. Though our research does not delve into the specifics of this issue, we have included data that explore the role of immigration in mothers on their children's ASD detection rates. Previous findings using the Modified Checklist for Autism in Toddlers (MCHAT) in a pediatric healthcare setting also showed higher rates of positive screenings among children of color and those from lower-income households compared with white, higher-income, privately insured, and suburban children [57]. As opposed to a specific tool, designing a screening system that relies on routinely collected data enables ASD detection that is based upon multi-cultural representative data. Carrying over such technology to developing countries requires direct cross-cultural validation at multiple levels, including the expression of signs, parental recognition, interpretation, and reporting levels [58].

## 5. Limitations

Integrating developmental milestones in the model required accounting for the variability in timing of checkups as well as multiple entries per milestone per child. To avoid attributing a-priori developmental assumptions, for instance to: 'never tested' and 'late tested' data, milestones were transformed to binary features of 'ever passed' and 'only failed'. While this enabled us to minimize the attribution of late performance to later development, we were not able to account for children who passed after several attempts. Future ML work that integrates a child's pattern of performance over time may improve the model's capacity to capture delay.

The ML model is limited to the list of milestones and risk factors examined and recorded in national surveillance check-ups. Additional early signs linked to subsequent ASD diagnosis were not examined, and as such, were not identified in our study. Milestones such as head lag at 3 months, raising hands to be picked up at 6 months, and nodding at 12 months are not included in the formal examination and may potentially be important "red flags" that strengthen the model [59]. In addition, father age and family history of ASD are two known risk factors [51] that are not routinely collected and would be important to add in future research for promoting better ASD detection. Advanced maternal age was found to be significant, and it may be a proxy for advanced paternal age.

An additional limitation is that our identification of ASD is based on reports during visits. It is plausible that there is a much larger ASD group that has been diagnosed later or not reported during well-baby visits. Based on the national incidence of ASD, the true positive ASD group is at least 10 times higher in the examined population [5]. Lower rates of ASD also relate to the young age of this community sample (born 2014–2019). However, as the model relied on accurate and early ASD diagnoses, it enhances its reliability in identifying the group that has not been previously identified in the EHR through earlier detection in future assessments. We attempted to identify some of these children through the estimation of global developmental delay among the high-risk group. This adds a lot of noise to the outcome label, and thus, requires replication with a later stable outcome label. In addition, there was variability in the way providers recorded diagnosis, with 43.37% of the ASD diagnoses ascertained via provider notes rather than ICD-9 code. Further research into inter-rater assessment differences in clinics is warranted.

## 6. Conclusions

This study highlights the value of ML systems for detecting elevated ASD signs in preventative care. The most important predicting features in our models reflected a mixture of biometric and clinical data. The representative nature of this rich database enables us to account for the medical and behavioral heterogeneity of young children. Although clinical data are noisy, the model showed high accuracy levels, identifying most children with ASD prior to 2 years of age. While the model had a cost of high false positive rates, some of these children showed other developmental issues that required verification. The model pointed to significant milestones and features that are easy to recognize and can be incorporated in detecting children that mandate a more thorough and systematic ASD-focused evaluation. Future work is needed to integrate such data-driven ML systems to support providers in personalizing early detection of ASD during routine developmental checkups.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/children11040429/s1>, Table S1: List of 100 features; Table S2: Model performance for training data; Table S3: Comparison of model performance with different thresholds (various multiples of the Youden Threshold); Table S4: Accuracy measures of the Full Model, Boys Model, and No-2019 Model; Figure S1: Feature importance bar plot and SHAP summary plot for a fold 1 from the Full Model; Figure S2: Feature importance bar plot and SHAP summary plot for a fold 3 from the Full Model; Figure S3: Feature importance bar plot and SHAP summary plot for a fold 1 from the Boys Model; Figure S4: Feature importance bar plot and SHAP summary plot for a fold from the Boys Model; Figure S5: Feature importance bar plot and SHAP summary plot for fold 3 from the Boys Model.

**Author Contributions:** Conceptualization and methodology, A.B.-S., J.G. and L.V.G.; formal analysis, A.B.-S., J.G., L.N., M.S. and K.I.; data curation, A.B.-S., L.N., K.I. and M.S.; writing—original draft preparation, A.B.-S., J.G. and L.V.G.; writing—review and editing, A.B.-S., L.V.G., J.G., K.I. and M.S.; visualization, A.B.-S., J.G., L.N. and M.S.; supervision, A.B.-S. and L.V.G.; project administration, A.B.-S.; funding acquisition, A.B.-S. and L.V.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Gertner Institute of Health Policy and Epidemiology, grant number 2020.351.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of the Israeli Ministry of Health (#15/2021 and 13 June 2021).

**Informed Consent Statement:** Patient consent was waived by the ethics committee due to the anonymized data analysis nature of research.

**Data Availability Statement:** The data presented in this study are available on request from the TIMNA-Israel Ministry of Health's Big Data Platform, Ministry of Health and are available at <https://govextra.gov.il/ministry-of-health/big-data-research/home/> (accessed on 31 December 2021), Jerusalem, Israel. The data are not publicly available due to privacy and ethical restrictions.

**Acknowledgments:** Thank you to Roe Cohen and Galit Shefer from TIMNA-Israel Ministry of Health's Big Data Platform, Ministry of Health, Jerusalem, Israel for their assistance with administrative and database management issues. We appreciate Dina Zimmerman, MoH, for her input regarding the interpretation of baby-wellness records.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed.; American Psychiatric Association: Washington, DC, USA, 2013.
2. Maenner, M.J.; Warren, Z.; Williams, A.R.; Amoakohene, E.; Bakian, A.V.; Bilder, D.A.; Durkin, M.S.; Fitzgerald, R.T.; Furnier, S.M.; Hughes, M.M.; et al. Prevalence and characteristics of autism spectrum disorder among children aged 8 years—Autism and developmental disabilities monitoring network, 11 sites, United States, 2020. *MMWR Surveill Summ* **2023**, *72*, 1–14. [[CrossRef](#)] [[PubMed](#)]



3. Davidovitch, M.; Hemo, B.; Manning-Courtney, P.; Fombonne, E. Prevalence and incidence of autism spectrum disorder in an Israeli population. *J. Autism Dev. Disord.* **2012**, *43*, 785–793. [[CrossRef](#)] [[PubMed](#)]
4. Zeidan, J.; Fombonne, E.; Scora, J.; Ibrahim, A.; Durkin, M.S.; Saxena, S.; Yusuf, A.; Shih, A.; Elsabbagh, M. Global prevalence of autism: A systematic review update. *Autism Res.* **2022**, *15*, 778–790. [[CrossRef](#)] [[PubMed](#)]
5. Dinstein, I.; Solomon, S.; Zats, M.; Shusel, R.; Lottner, R.; Gershon, B.B.; Meiri, G.; Menashe, I.; Shmueli, D. Two-fold increase in the prevalence of autism in Israel between 2017 and 2021. *medRxiv* **2023**. [[CrossRef](#)]
6. Fuller, E.A.; Kaiser, A.P. The effects of early intervention on social communication outcomes for children with autism spectrum disorder: A meta-analysis. *J. Autism Dev. Disord.* **2020**, *50*, 1683–1700. [[CrossRef](#)] [[PubMed](#)]
7. Zwaigenbaum, L.; Bauman, M.L.; Choueiri, R.; Kasari, C.; Carter, A.; Granpeesheh, D.; Mailloux, Z.; Roley, S.S.; Wagner, S.; Fein, D.; et al. Early intervention for children with autism spectrum disorder under 3 years of age: Recommendations for practice and research. *Pediatrics* **2015**, *136* (Suppl. S1), S60–S81. [[CrossRef](#)] [[PubMed](#)]
8. American Academy of Pediatrics; Council on Children with Disabilities; Section on Developmental Behavioral Pediatrics; Bright Futures Steering Committee; Medical Home Initiatives for Children with Special Needs Project Advisory Committee. Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics* **2006**, *118*, 405–420. [[CrossRef](#)] [[PubMed](#)]
9. Siu, A.L.; Bibbins-Domingo, K.; Grossman, D.C.; Baumann, L.C.; Davidson, K.W.; Ebell, M.; García, F.A.R.; Gillman, M.; Herzstein, J.; Kemper, A.R.; et al. Screening for autism spectrum disorder in young children: US preventive services task force recommendation statement. *JAMA* **2016**, *315*, 691–696. [[CrossRef](#)] [[PubMed](#)]
10. Workgroup, Bright Futures Periodicity Schedule; Committee on Practice and Ambulatory Medicine; Simon, G. R.; Baker, C.N.; Barden, G.A.; Brown, O.S.W., III; Hackell, J.M.; Hardin, A.P.; Meade, K.E.; Moore, S.B.; et al. 2016 recommendations for preventive pediatric health care. *Pediatrics* **2016**, *137*, 1.
11. Barbaro, J.; Dissanayake, C. Autism spectrum disorders in infancy and toddlerhood: A review of the evidence on early signs, early identification tools, and early diagnosis. *J. Dev. Behav. Pediatr.* **2009**, *30*, 447–459. [[CrossRef](#)]
12. Dawson, G. Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. *Dev. Psychopathol.* **2008**, *20*, 775–803. [[CrossRef](#)] [[PubMed](#)]
13. Webb, S.J.; Jones, E.J.H.; Kelly, J.; Dawson, G. The motivation for very early intervention for infants at high risk for autism spectrum disorders. *Int. J. Speech Lang. Pathol.* **2014**, *16*, 36–42. [[CrossRef](#)] [[PubMed](#)]
14. Sudry, T.; Zimmerman, D.R.; Yardeni, H.; Joseph, A.; Baruch, R.; Grotto, I.; Greenberg, D.; Eilenberg, R.; Amit, G.; Akiva, P.; et al. Standardization of a developmental milestone scale using data from children in Israel. *JAMA Netw. Open* **2022**, *5*, e222184. [[CrossRef](#)] [[PubMed](#)]
15. Sadaka, Y.; Sudry, T.; Zimmerman, D.R.; Tsadok, M.A.; Baruch, R.; Yardeni, H.; Ben Moshe, D.; Akiva, P.; Amit, G. Assessing the attainment rates of updated CDC milestones using a new Israeli developmental scale. *Pediatrics* **2022**, *150*, e2022057499. [[CrossRef](#)] [[PubMed](#)]
16. Ben-Sasson, A.; Carter, A.S. The contribution of sensory–regulatory markers to the accuracy of ASD screening at 12 months. *Res. Autism Spectr. Disord.* **2013**, *7*, 879–888. [[CrossRef](#)]
17. Ben-Sasson, A.; Habib, S.; Tirosh, E. Feasibility and validity of early screening for identifying infants with poor social-communication development in a well-baby clinic system. *J. Pediatr. Nurs.* **2014**, *29*, 238–247. [[CrossRef](#)] [[PubMed](#)]
18. Landa, R.J.; Gross, A.L.; Stuart, E.A.; Faherty, A. Developmental trajectories in children with and without autism spectrum disorders: The first 3 years. *Child. Dev.* **2013**, *84*, 429–442. [[CrossRef](#)]
19. Pierce, K.; Carter, C.; Weinfeld, M.; Desmond, J.; Hazin, R.; Bjork, R.; Gallagher, N. Detecting, studying, and treating autism early: The one-year well-baby check-up approach. *J. Pediatr.* **2011**, *159*, 458–465. [[CrossRef](#)] [[PubMed](#)]
20. Brett, D.; Warnell, F.; McConachie, H.; Parr, J.R. Factors affecting age at ASD diagnosis in UK: No evidence that diagnosis age has decreased between 2004 and 2014. *J. Autism Dev. Disord.* **2016**, *46*, 1974–1984. [[CrossRef](#)]
21. Daniels, A.M.; Mandell, D.S. Explaining differences in age at autism spectrum disorder diagnosis: A critical review. *Autism* **2013**, *18*, 583–597. [[CrossRef](#)]
22. Ozonoff, S.; Heung, K.; Byrd, R.; Hansen, R.; Hertz-Picciotto, I. The onset of autism: Patterns of symptom emergence in the first years of life. *Autism Res.* **2008**, *1*, 320–328. [[CrossRef](#)] [[PubMed](#)]
23. Rogers, S.J. What are infant siblings teaching us about autism in infancy? *Autism Res.* **2009**, *2*, 125–137. [[CrossRef](#)]
24. Davidovitch, M.; Stein, N.; Koren, G.; Friedman, B.C. Deviations from typical developmental trajectories detectable at 9 months of age in low risk children later diagnosed with autism spectrum disorder. *J. Autism Dev. Disord.* **2018**, *48*, 2854–2869. [[CrossRef](#)]
25. Macari, S.L.; Campbell, D.; Gengoux, G.W.; Saulnier, C.A.; Klin, A.J.; Chawarska, K. Predicting developmental status from 12 to 24 months in infants at risk for autism spectrum disorder: A preliminary report. *J. Autism Dev. Disord.* **2012**, *42*, 2636–2647. [[CrossRef](#)]
26. Lipkin, P.H.; Macias, M.M.; Chen, B.B.; Coury, D.; Gottschlich, E.A.; Hyman, S.L.; Sisk, B.; Wolfe, A.; Levy, S.E. Trends in pediatricians’ developmental screening: 2002–2016. *Pediatrics* **2020**, *145*, e20190851. [[CrossRef](#)]
27. Zwaigenbaum, L.; Bryson, S.; Lord, C.; Rogers, S.; Carter, A.; Carver, L.; Chawarska, K.; Constantino, J.; Dawson, G.; Dobkins, K.; et al. Clinical assessment and management of toddlers with suspected autism spectrum disorder: Insights from studies of high-risk infants. *Pediatrics* **2009**, *123*, 1383–1391. [[CrossRef](#)] [[PubMed](#)]



28. Allen, L.; Leon-Attia, O.; Shaham, M.; Shefer, S.; Gabis, L.V. Autism risk linked to prematurity is more accentuated in girls. *PLoS ONE* **2020**, *15*, e0236994. [[CrossRef](#)]
29. Johnson, C.P.; Myers, S.M. Identification and evaluation of children with autism spectrum disorders. *Pediatrics* **2007**, *120*, 1183–1215. [[CrossRef](#)] [[PubMed](#)]
30. Wang, C.; Geng, H.; Liu, W.; Zhang, G. Prenatal, perinatal, and postnatal factors associated with autism. *Medicine* **2017**, *96*, e6696. [[CrossRef](#)]
31. Ozonoff, S.; Young, G.S.; Carter, A.; Messinger, D.; Yirmiya, N.; Zwaigenbaum, L.; Bryson, S.; Carver, L.J.; Constantino, J.N.; Dobkins, K.; et al. Recurrence risk for autism spectrum disorders: A baby siblings research consortium study. *Pediatrics* **2011**, *128*, e488–e495. [[CrossRef](#)]
32. Guedalia, J.; Lipschuetz, M.; Novoselsky-Persky, M.; Cohen, S.M.; Rottenstreich, A.; Levin, G.; Yagel, S.; Unger, R.; Sompolinsky, Y. Real-time data analysis using a machine learning model significantly improves prediction of successful vaginal deliveries. *Am. J. Obstet. Gynecol.* **2020**, *223*, 437.e1–437.e15. [[CrossRef](#)] [[PubMed](#)]
33. Parlett-Pelleriti, C.M.; Stevens, E.; Dixon, D.; Linstead, E.J. Applications of unsupervised machine learning in autism spectrum disorder research: A review. *Rev. J. Autism Dev. Disord.* **2023**, *10*, 406–421. [[CrossRef](#)]
34. Stevens, E.; Dixon, D.R.; Novack, M.N.; Granpeesheh, D.; Smith, T.; Linstead, E. Identification and analysis of behavioral phenotypes in autism spectrum disorder via unsupervised machine learning. *Int. J. Med. Inform.* **2019**, *129*, 29–36. [[CrossRef](#)] [[PubMed](#)]
35. Milano, N.; Simeoli, R.; Rega, A.; Marocco, D. A deep learning latent variable model to identify children with autism through motor abnormalities. *Front. Psychol.* **2023**, *14*, 1194760. [[CrossRef](#)] [[PubMed](#)]
36. Maharjan, J.; Garikipati, A.; Dinunno, F.A.; Ciobanu, M.; Barnes, G.; Browning, E.; DeCurzio, J.; Mao, Q.; Das, R. Machine learning determination of applied behavioral analysis treatment plan type. *Brain Inform.* **2023**, *10*, 7. [[CrossRef](#)] [[PubMed](#)]
37. Lingren, T.; Chen, P.; Bochenek, J.; Doshi-Velez, F.; Manning-Courtney, P.; Bickel, J.; Wildenger Welchons, L.; Reinhold, J.; Bing, N.; Ni, Y.; et al. Electronic health record based algorithm to identify patients with autism spectrum disorder. *PLoS ONE* **2016**, *11*, e0159621.
38. Leroy, G.; Gu, Y.; Pettygrove, S.; Galindo, M.K.; Arora, A.; Kurzius-Spencer, M. Automated extraction of diagnostic criteria from electronic health records for autism spectrum disorders: Development, evaluation, and application. *J. Med. Internet Res.* **2018**, *20*, e10497. [[CrossRef](#)] [[PubMed](#)]
39. Maenner, M.J.; Yeargin-Allsopp, M.; Van Naarden Braun, K.; Christensen, D.L.; Schieve, L.A. Development of a machine learning algorithm for the surveillance of autism spectrum disorder. *PLoS ONE* **2016**, *11*, e0168224. [[CrossRef](#)] [[PubMed](#)]
40. Rahman, R.; Kodesh, A.; Levine, S.Z.; Sandin, S.; Reichenberg, A.; Schlessinger, A. Identification of newborns at risk for autism using electronic medical records and machine learning. *Eur. Psychiatry* **2020**, *63*, e22. [[CrossRef](#)]
41. Ejlskov, L.; Wulff, J.N.; Kalkbrenner, A.; Ladd-Acosta, C.; Fallin, M.D.; Agerbo, E.; Mortensen, P.B.; Lee, B.K.; Schendel, D. Prediction of autism risk from family medical history data using machine learning: A national cohort study from denmark. *Biol. Psychiatry Glob. Open Sci.* **2021**, *1*, 156–164. [[CrossRef](#)]
42. Thabtah, F. An accessible and efficient autism screening method for behavioural data and predictive analyses. *Health Inform. J.* **2018**, *25*, 1739–1755. [[CrossRef](#)] [[PubMed](#)]
43. Israel Ministry of Health. Developmental Assessments Guidelines for up to Age Six. 2016. Available online: [https://www.health.gov.il/hozer/bz12\\_2016.pdf](https://www.health.gov.il/hozer/bz12_2016.pdf) (accessed on 3 January 2021).
44. Frankenburg, W.K.; Dodds, J.B. The Denver developmental screening test. *J. Pediatr.* **1967**, *71*, 181–191. [[CrossRef](#)] [[PubMed](#)]
45. World Health Organization. *International Statistical Classification of Diseases and Related Health Problems*, 9th ed.; World Health Organization: Geneva, Switzerland, 1979.
46. Shevell, M.I.; Ashwal, S.; Donley, D.; Flint, J.; Gingold, M.; Hirtz, D.; Majnemer, A.; Noetzel, N.; Sheth, R.D. Practice parameter: Evaluation of the child with global developmental delay: Report of the quality standards subcommittee of the american academy of neurology and the practice committee of the child neurology society. *Neurology* **2003**, *60*, 367–380. [[CrossRef](#)] [[PubMed](#)]
47. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased boosting with categorical features. *Adv. Neural Inf. Process Syst.* **2018**, *31*, 6638–6648. [[CrossRef](#)]
48. Youden, W.J. Index for rating diagnostic tests. *Cancer* **1950**, *3*, 32–35. [[CrossRef](#)]
49. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process Syst.* **2017**, *30*, 4765–4774. [[CrossRef](#)]
50. Vidgen, B.; Yasseri, T. P-values: Misunderstood and misused. *Front Psychol.* **2016**, *4*, 6. [[CrossRef](#)]
51. Hyman, S.L.; Levy, S.E.; Myers, S.N.; Kuo, D.Z.; Apkon, S.; Davidson, L.F.; Ellerbeck, K.A.; Foster, J.E.A.; Noritz, G.H.; Leppert, M.O.; et al. Identification, evaluation, and management of children with autism spectrum disorder. *Pediatrics* **2020**, *145*, e20193447. [[CrossRef](#)] [[PubMed](#)]
52. Zwaigenbaum, L.; Bryson, S.; Garon, N. Early identification of autism spectrum disorders. *Behav. Brain Res.* **2013**, *251*, 133–146. [[CrossRef](#)]
53. Tseng, P.-T.; Yen, C.-F.; Chen, Y.-W.; Stubbs, B.; Carvalho, A.F.; Whiteley, P.; Chu, C.-S.; Li, D.-J.; Chen, T.-Y.; Tang, C.-H.; et al. Maternal breastfeeding and attention-deficit/hyperactivity disorder in children: A meta-analysis. *Eur. Child. Adolesc. Psychiatry* **2018**, *28*, 19–30. [[CrossRef](#)]

54. Landa, R.; Garrett-Mayer, E. Development in infants with autism spectrum disorders: A prospective study. *J. Child. Psychol. Psychiatry* **2006**, *47*, 629–638. [[CrossRef](#)] [[PubMed](#)]
55. Surén, P.; Stoltenberg, C.; Bresnahan, M.; Hirtz, D.; Lie, K.K.; Lipkin, W.I.; Magnus, P.; Reichborn-Kjennerud, T.; Schjølberg, S.; Susser, E.; et al. Early growth patterns in children with autism. *Epidemiology* **2013**, *24*, 660–670. [[CrossRef](#)] [[PubMed](#)]
56. Kamer, A.; Zohar, A.H.; Youngmann, R.; Diamonds, G.W.; Inbar, D.; Senecky, Y.A. Prevalence estimate of pervasive developmental disorder among immigrants to Israel and Israeli natives: A file review study. *Soc. Psychiatry Psychiatr. Epidemiol.* **2004**, *39*, 141–145. [[CrossRef](#)] [[PubMed](#)]
57. Guthrie, W.; Wallis, K.; Bennett, A.; Brooks, E.; Dudley, J.; Gerdes, M.; Pandey, J.; Levy, S.E.; Schultz, R.T.; Miller, J.S. Accuracy of autism screening in a large pediatric network. *Pediatrics* **2021**, *144*, e20183963. [[CrossRef](#)] [[PubMed](#)]
58. de Leeuw, A.; Happé, F.; Hoekstra, R.A. A conceptual framework for understanding the cultural and contextual factors on autism across the globe. *Autism Res.* **2020**, *13*, 1029–1050. [[CrossRef](#)]
59. Gabis, L.V.; Shaham, M.; Attia, O.L.; Shefer, S.; Rosenan, R.; Gabis, T.; Daloya, M. The weak link: Hypotonia in infancy and autism early identification. *Front. Neurol.* **2021**, *12*, 612674. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.