DOI: 10.58490/ctjump.2026i11TA.4353

APPLICATION OF ARTIFICIAL INTELLIGENCE IN SCREENING FOR AUTISM SPECTRUM DISORDER IN CHILDREN AT CAN THO CHILDREN'S HOSPITAL

Tran Thi Ngoc Phuong*, Nguyen Van Trinh, Ayyakannu Elavarasan Aishwarya, Srinivasan Jyothi Dyana, Vo Van Thi, Lu Tri Dien

Can Tho University of Medicine and Pharmacy

**Corresponding author: 2253010674@student.ctump.edu.vn*

Received: 30/01/2026

Reviewed: 25/02/2026

Accepted: 25/6/2026

ABSTRACT

Background**: Autism Spectrum Disorder is a complex neurodevelopmental disorder characterized by social-communication deficits and restricted, repetitive behaviors, with rising global prevalence, including in Vietnam, where early detection remains limited. Artificial intelligence (AI) provides a data-driven approach to improve diagnostic accuracy.* ***Objectives: *To develop an AI-based screening model for Autism Spectrum Disorder in children aged 18-36 months at Can Tho Children's Hospital and to identify perinatal factors associated with Autism Spectrum Disorder.* ***Materials and methods:*** *A descriptive cross-sectional study was conducted on 980 children aged 18-36 months (November 2024-November 2025). Screening used M-CHAT-R/F, followed by DSM-5 confirmation. Logistic regression identified independent risk factors, and multiple AI models were evaluated using standard performance metrics.* ***Results:*** *Of 980 children, 43 (4.4%) were diagnosed with Autism Spectrum Disorder. Independent predictors included interventional delivery (aOR = 2.10), abnormal labor duration (aOR = 4.21), preterm birth (aOR = 2.31), and birth asphyxia (aOR = 7.24). M-CHAT-R/F Questions 1, 6, and 20 were most predictive. Decision Tree and Random Forest showed the best performance, with AUC values of approximately 0.80.* ***Conclusions:*** *Autism Spectrum Disorder prevalence was 4.4%, with interventional delivery, prolonged labor, preterm birth, and birth asphyxia as major perinatal risk factors. AI-based screening models demonstrated good discriminatory performance and may serve as useful tools to support early ASD screening in resource-limited healthcare settings.*

***Keyword**: Autism spectrum disorder, Autism Spectrum Disorder diagnosis, Machine learning.*

I. INTRODUCTION

Autism Spectrum Disorder (ASD) is an increasing global public health concern due to its rising prevalence and substantial social, emotional, and economic impact. ASD is a neurodevelopmental disorder characterized by deficits in social communication and restricted, repetitive behaviors, with wide variability in symptom severity. Its etiology involves genetic and environmental factors, including maternal stress, pesticide exposure, and prenatal or perinatal hypoxia [1]. The CDC (2023) reports that approximately 1 in 36 children are affected worldwide [2]. In Vietnam, Le Thi Vui (2021) reported an ASD prevalence of 0.758% among children aged 18-30 months, particularly in boys and urban areas [3]. These findings confirm the increasing ASD burden and emphasize the importance of early screening. Undiagnosed ASD may result in long-term mental health issues, social difficulties, and limited educational or occupational outcomes [4]. In Vietnam, diagnosis remains challenging due to limited expertise, low awareness, and unequal healthcare

access. Traditional diagnosis requires multidisciplinary evaluation and is often time-consuming and costly.

Recently, artificial intelligence (AI) has shown promise in improving ASD detection. Machine learning models can analyze behavioral and clinical data to support diagnostic decisions with improved accuracy and reduced assessment time [5]. Recent systematic reviews have demonstrated promising performance of AI-based ASD screening tools, particularly those employing multimodal approaches that integrate diverse data types [6]. However, most studies have been conducted in high-income countries using datasets that may not represent the demographic and clinical characteristics of children in Vietnam. Moreover, despite the growing need for widespread ASD screening and increasing interest in AI applications in Vietnam, the use of AI for ASD diagnosis remains limited. Therefore, this study was conducted to develop and evaluate an AI-assisted screening model for ASD in Vietnamese children and to identify perinatal factors independently associated with ASD.

II. MATERIALS AND METHODS

2.1. Participants

The study included children aged 18-36 months receiving care at Can Tho Children's Hospital from November 2024 to November 2025.

- Inclusion criteria: Were informed parental consent.

- Exclusion criteria: Severe comorbidities (e.g., cerebral palsy, epilepsy), incomplete data, or withdrawal before completion.

2.2. Research methods

The sample size was estimated based on the requirement for developing a clinical prediction model. The calculation followed the Van Smeden-Riley framework, which considers the outcome event rate, the number of candidate predictor variables, and the target mean absolute prediction error to reduce the risk of model overfitting and ensure predictive stability. In this study, the outcome event was defined as children at risk of autism spectrum disorder, with an estimated event rate of approximately 8.4% [7]. Ten candidate predictor variables were included in the model, and a target mean absolute prediction error of 0.03 was selected. After adjusting for 10% potential missing or incomplete data, the minimum required sample size was approximately 715 children. A total of 980 eligible children were ultimately enrolled and included in the final analysis, exceeding the minimum sample size required for model development.

- Sampling method: convenience sampling

- Data Collection and Research Process: Demographic and perinatal data were collected, including age, sex, birth order, parental age, family history, and maternal or perinatal conditions. The M-CHAT-R/F was used for initial screening. Children scoring ≥3 underwent follow-up screening and DSM-5-based diagnostic confirmation by child psychiatrists. Research assistants were trained in standardized data collection and screening procedures.

- Data Analysis and AI Model Evaluation

All data were verified, coded, and analyzed using Excel and SPSS 20.0. Descriptive statistics included frequencies, percentages, and mean ± standard deviation. Group comparisons used Chi-square or Fisher's exact tests, with statistical significance set at $p < 0.05$.

Machine-learning analyses were performed using R software and RStudio. Because

the proportion of ASD cases was relatively low, the dataset was imbalanced, which could reduce the model's ability to identify ASD cases. Therefore, SMOTE was applied only to the training set to synthetically increase the minority ASD class, while the testing set was kept unchanged to ensure an unbiased evaluation of model performance. All machine learning models were implemented using the Python scikit-learn library. For all experimental conditions, the dataset was partitioned into a 75% training set and a 25% test set.

The AI model was trained on demographic, clinical, and behavioral data from ASD and control groups. Its performance was evaluated against physicians' diagnoses using sensitivity, specificity, accuracy, F1-score, and AUC.

- Ethical Approval: The research was conducted in accordance with an approved research protocol, with consent from the Can Tho University of Medicine and Pharmacy Council and the Ethics Committee of Can Tho University of Medicine and Pharmacy, under approval number 24.159.SV/PCT-HĐĐĐ. Written informed consent was obtained from all parents or legal guardians prior to participation. All procedures adhered to ethical principles, ensuring participants were not exposed to any physical or psychological harm. The study was conducted solely for scientific purposes, and personal information was kept strictly confidential. Participants continued to receive standard care and counseling during and after the study.

III. RESULTS

From November 2024 to November 2025, a total of 980 children aged 18-36 months were enrolled at Can Tho Children's Hospital.

3.1. General Characteristics

Table 1. Age and gender characteristics of the study subjects (n = 980)

Child characteristics		Frequency (n)	Percentage (%)
Age group	18-<24 months	79	8.1
	24-36 months	901	91.9
Gender	Male	483	49.3
	Female	497	50.7
Total		980	100.0

Most participants were 24-36 months old (91.9%), with an almost equal distribution of male (49.3%) and female (50.7%). Over half were first-born (55.9%), 38.5% were second-born, and only 5.7% were third-born or later.

Table 2. Characteristics of factors associated with the risk of Autism Spectrum Disorder

Predictive factors	Total (n = 980) n (%)	DSM-5 (+) (n = 43) n (%)	DSM-5 (-) (n = 937) n (%)
Maternal age > 35 years	158 (16.1)	8 (18.6)	150 (16.0)
Rural residence	327 (33.4)	17 (39.5)	310 (33.1)
Miscarriage/abortion	39 (4.0)	1 (2.3)	38 (4.1)
Interventional delivery	266 (27.1)	22 (51.2)	244 (26.0)
Abnormal labor duration	35 (3.6)	8 (18.6)	27 (2.9)
Preterm birth	49 (5.0)	8 (18.6)	41 (4.4)
Low birth weight	14 (1.4)	3 (7.0)	11 (1.2)
Birth asphyxia	7 (0.7)	4 (9.3)	3 (0.3)

Children with ASD had substantially higher rates of interventional delivery (51.2%

vs. 26.0%), abnormal labor (18.6% vs. 2.9%), prematurity (18.6% vs. 4.4%), and birth asphyxia (9.3% vs. 0.3%). In contrast, maternal age >35 years, rural residence, and miscarriage history showed little difference.

3.2. Screening and Diagnostic Outcomes

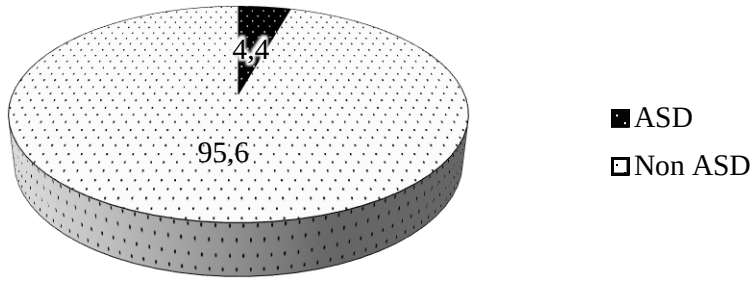

Figure 1. Proportion of children diagnosed with Autism Spectrum Disorder

Based on DSM-5 diagnostic confirmation, 43 of 980 children (4.4%) were diagnosed with ASD, while 937 children (95.6%) were classified as non-ASD. This prevalence is higher than that reported in community-based studies but reflects the hospital-based nature of the study population and the systematic screening process employed.

3.3. Factors Associated with ASD

Table 3. Relationship between perinatal and postnatal factors and autism

Variables	Crude OR (95% CI)	p	Adjusted OR (95% CI)	p
Gender	1.050 (1.035-1.066)	0.042	1.205 (0.890-1.066)	0.088
Interventional delivery	2.808 (1.570-5.021)	0.001	2.104 (1.140-7.123)	0.032
Abnormal labor duration	6.171 (3.095-12.304)	<0.001	4.214 (1.130-8.285)	<0.001
Preterm birth	4.343 (2.130-8.854)	<0.001	2.312 (1.313-9.321)	<0.001
Low birth weight	5.175 (1.815-14.758)	0.020	3.264 (0.725-18.944)	0.185
Birth asphyxia	14.256 (6.999-29.039)	<0.001	7.242 (1.442-17.228)	< 0.001
Question 1	1.236 (0.842-1.814)	0.179	2.126 (1.212-3.628)	0.009
Question 6	3.587 (1.980-6.501)	<0.001	4.388 (1.402-13.730)	0.011
Question 7	1.586 (1.028-2.445)	0.036	0.576 (0.327-1.017)	0.057
Question 10	2.312 (1.428-3.744)	0.001	2.325 (0.978-5.528)	0.056
Question 20	1.537 (1.019-2.323)	0.042	2.366 (1.302-4.299)	0.005

After adjustment, interventional delivery (aOR = 2.10), abnormal labor duration (aOR = 4.21), prematurity (aOR = 2.31), and birth asphyxia (aOR = 7.24) remained independently associated with ASD, while low birth weight was no longer significant (aOR = 3.26, p = 0.185). Significant M-CHAT-R/F predictors included Question 1 (aOR = 2.13), Question 6 (aOR = 4.39), and Question 20 (aOR = 2.37).

3.4. M-CHAT-R/F Items and Model Performance

Table 4. Performance metrics of the Refined Model for Autism Prediction

Rank	Model	AUC	95% CI	Acc	Se	Sp	Pre	F1
1	Random Forest + SMOTE	0.8	0.65-0.88	0.79	0.70	0.73	0.61	0.70
2	Decision Tree + SMOTE	0.8	0.55-0.84	0.80	0.70	0.75	0.75	0.80
3	AdaBoost +SMOTE	0.799	0.59-0.85	0.79	0.58	0.75	0.80	0.74
4	Logistic Regression + SMOTE	0.79	0.66-0.86	0.79	0.58	0.75	0.80	0.74
5	SVM Radial + SMOTE	0.786	0.59-0.89	0.76	0.25	0.73	0.42	0.38
6	KNN + SMOTE	0.784	0.66-0.86	0.76	0.58	0.71	0.37	0.49
7	Naïve Bayes + SMOTE	0.76	0.54-0.83	0.03	0.70	0.19	0.15	0.10

After SMOTE, all models showed fair performance (AUC ≈ 0.78-0.80). Decision Tree and Random Forest performed best (AUC = 0.80; F1 = 0.80 and 0.70) with balanced sensitivity and specificity (~70-75%), while AdaBoost and Logistic Regression showed similar results (AUC ≈ 0.79, F1 ≈ 0.74). SVM and KNN had moderate performance (AUC ≈ 0.785), whereas Naïve Bayes performed poorly (AUC = 0.76, F1 = 0.10).

IV. DISCUSSION

4.1. General characteristics of the study population

Among 980 children aged 18-36 months, most were 24-36 months (91.9%), with 8.1% aged 18-24 months. Sex distribution was nearly balanced (female 50.7%, male 49.3%), consistent with the study Vo Van Thi (2023) [7]. This suggests that population-based ASD screening tends to show gender parity compared with clinical samples. Regarding birth order, first-born children accounted for 55.9%, followed by second-born (38.5%) and third-born or later (5.7%), similar to findings by Vo Van Thi (2023) [7]. The predominance of first-born children may reflect smaller family sizes. Although some studies suggest a slightly higher ASD risk among first-borns, evidence remains inconsistent, indicating that birth order alone is not a strong ASD determinant in Vietnamese toddlers.

4.2. Screening performance and ASD prevalence

Abnormal M-CHAT-R/F responses most commonly involved sensory sensitivity, impaired nonverbal communication, and limited imaginative play, consistent with Vietnamese studies by Vo Van Thi (2023) and Nguyen Minh Phuong (2021) [7]. After two-stage screening and clinical confirmation, ASD prevalence was 4.4%, higher than some community studies (0.758%) [3] but consistent with the rising global trend reported by the CDC (2023) [2]. This difference may reflect hospital-based recruitment, improved detection, and use of DSM-5 criteria. These findings highlight the increasing ASD burden in Vietnam and the need for routine developmental screening.

4.3. Factors associated with ASD risk

Perinatal factors including interventional delivery, prolonged labor, preterm birth, and birth asphyxia were significantly associated with ASD and remained independent predictors in multivariate analysis. These results align with international studies. Crump *et al.* (2021) reported increased ASD risk in preterm births, particularly extremely preterm infants [8]. Arun (2023) found birth asphyxia increased ASD odds by 12.6 times [9], while Silva *et al*. (2022) reported higher ASD risk following emergency cesarean delivery [10]. These factors may impair oxygenation and neural development during critical periods.

Birth asphyxia, prolonged or abnormal labor might all effect oxygen concentration in new born , impact on neural development pathways that are already vulnerable in individuals who later present with ASD [11],[12]. In contrast, no significant associations were found with sex, maternal age ≥35 years, or birth order. Overall, the findings emphasize the role of perinatal rather than demographic factors in ASD risk and support improved obstetric and neonatal care for early prevention.

4.4. AI prediction model

In this study, all machine-learning models demonstrated fair diagnostic ability (AUC ≈ 0.78-0.80), with Decision Tree and Random Forest achieving the best overall performance. This result was supported by Zambre's study, which reported that Random Forest algorithms' predictive power made them suitable in aiding ASD detection [13]. In addition, Fatemeh Mohammadi [5] concluded that integrating voice analysis and facial analysis was proven to be effective in assisting ASD diagnosis.

In the Vietnamese healthcare context, AI-based models may provide valuable support for early ASD screening. Due to the relatively limited development of psychiatric and developmental health services in Vietnam, many parents may be unfamiliar with the early manifestations of autism and may not seek evaluation until symptoms become more apparent. In addition, teachers, particularly those in preschool and primary school settings, often lack formal training in recognizing ASD. As a result, AI-assisted screening tools could help identify children at risk at an earlier stage, thereby facilitating timely referral for diagnostic evaluation and appropriate intervention.

This study has several limitations. First, due to the small number of Autistic cases, this study is at risk of overfitting. Therefore, to further validate the effectiveness of the AI Model, further study studies should be carried out to test the performance of the AI model on other children in and outside Can Tho Children's Hospital. Second, the study population included only participants who were brought to the hospital because their symptoms affected their daily functioning. Therefore, certain characteristics and symptoms may have been underrepresented. Third, although Can Tho Children's Hospital is one of the largest pediatric hospitals in the Mekong Delta, the findings may not be representative of all children with autism in Vietnam. Future studies should include larger, multicenter populations to further validate the findings and improve model performance.

V. CONCLUSIONS

Among 980 children aged 18-36 months at Can Tho Children's Hospital, 4.4% were diagnosed with autism spectrum disorder based on DSM-5 criteria. The most frequently abnormal M-CHAT-R/F items were noise sensitivity, limited pretend play, atypical finger movements, and reduced attention-seeking behaviors. ASD was independently associated with interventional delivery, prolonged labor, preterm birth, and birth asphyxia, whereas no significant associations were found with age, sex, or maternal age. The AI-based diagnostic models demonstrated promising performance as a supportive tool for early ASD screening. Future multicenter studies involving larger and more diverse populations, including children recruited from hospital, kindergarten, and community settings, are warranted to further validate and improve model performance.

ACKNOWLEDGEMENTS

The research team sincerely thanks Can Tho University of Medicine and Pharmacy

for financially supporting the implementation of this study under Decision No. 4618/QD-DHYDCT dated December 17, 2024 of Can Tho University of Medicine and Pharmacy.

REFERENCES

1. Bozkurt A, İpek H, Bayındır K, *et al.* Evaluation of prenatal and perinatal risk factors in autism spectrum disorder according to disease severity. *Namik Kemal Med J.* 2025. 13(1), 69-75, doi: 10.4274/nkmj.galenos.2025.80270.
2. Maenner MJ, Warren Z, Williams AR, *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(2), 1-14, doi: 10.15585/mmwr.ss7202a1.
3. Thi Vui L, Duc DM, Thuy Quynh N, *et al.* Early screening and diagnosis of autism spectrum disorders in Vietnam: A population-based cross-sectional survey. *J Public Health Res.* 2022.11(2), 2460, doi: 10.4081/jphr.2021.2460.
4. Eigsti IM. The autism constellation and neurodiversity: Long-term and adult outcomes in autism spectrum disorder. *Pediatr Clin North Am.* 2024. 71(2), 327-341, doi: 10.1016/j.pcl.2024.01.003.
5. Mohammadi F, Shahrokhi H, Asadzadeh A, Pirmoradi S, Moghtader A, *et al.* Artificial intelligence in autism spectrum disorder diagnosis: A scoping review of face, voice, and text analysis methods. *Health Sci Rep.* 2025. 8(11), e71476, doi: 10.1002/hsr2.71476.
6. Shahini A, Kamath AP, Sharma E, Salvi M, Tan RS, Siuly S, *et al.* A systematic review for artificial intelligence-driven assistive technologies to support children with neurodevelopmental disorders. *Inf Fusion.* 2025. 124, 103441, doi: 10.1016/j.inffus.2025.103441.
7. Vo VT, Tran DT, Nguyen MP. Study on the prevalence of children from 18 to 36 months at risk of autism spectrum disorder by using M-CHAT-R/F scaleat Ca Mau province, Vietnam in 2022. *Vietnam Medical Journal.* 2023. 520(1B), doi: 10.51298/vmj.v520i1B.3905.
8. Crump C, Sundquist J, Sundquist K. Preterm or early term birth and risk of autism. *Pediatrics.* 2021.148(3),e2020032300, doi: 10.1542/peds.2020-032300.
9. Arun P, Azad C, Kaur G, Sharma P. A community-based study of antenatal and neonatal risk factors in autism spectrum disorder. *Prim Care Companion CNS Disord.* 2023. 25(2), 22m03339, doi: 10.4088/PCC.22m03339.
10. Silva VBD, Maia FA, Oliveira AJS, *et al*. Association between autism spectrum disorder and peripartum events: A case-control study. *Rev Paul Pediatr*. 2023. 41, e2021220, doi: 10.1590/1984-0462/2023/41/2021220.
11. Yuan JJ, Zhao YN, Lan XY, Zhang Y, Zhang R. Prenatal, perinatal and parental risk factors for autism spectrum disorder in China: a case-control study. *BMC Psychiatry*. 2024. 24(1), 219, doi: 10.1186/s12888-024-05643-0.
12. Preciado C, Baida M, Li Y, Li Y, Demopoulos C. Prenatal exposure to hypoxic risk conditions in autistic and neurotypical youth: Associated ventricular differences, sleep disturbance, and sensory processing. *Autism Res*. 2024. 17(12), 2547-2557, doi: 10.1002/aur.3250.
13. Zambre G, Albhar H, Masli T, Patil S. Detection and analysis of autism spectrum disorder using random forest classifier. In: Proceedings of the 2024 2nd International Conference on Recent Trends in Microelectronics, Automation, Computing and Communications Systems (ICMACC). IEEE. 2024, doi: 10.1109/ICMACC62921.2024.10894067.