Supplementary Materials

Artificial intelligence in environmental etiology of autism spectrum disorder: progress, opportunities, and challenges

Ruizhi Zheng[1], Lin Wen[2], Yun Huang[1], Ziying Gan[1], Hongyan Yin[3], Wei Xia[1], Hongxiu Liu[1,4]

[1]Key Laboratory of Environment and Health (HUST), Ministry of Education & Ministry of Environmental Protection, State Key Laboratory of Environmental Health (Incubation), School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei, China.
[2]Department of Nutrition and Food Hygiene, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei, China.
[3]School of Tropical Agriculture and Forestry, Hainan University, Haikou 570228, Hainan, China.
[4]Department of Environmental Health, School of Environmental Science and Engineering, Hainan University, Haikou 570228, Hainan, China.

Correspondence to: Dr. Hongxiu Liu, Department of Environmental Health, School of Environmental Science and Engineering, Hainan University, Haikou 570228, Hainan, China. E-mail: lhx@hainanu.edu.cn

Supplementary Text 1.

Methods

This review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines[1].

1 Literature Search Strategy

The literature search was performed on April 14, 2026, in the PubMed database. Search terms were developed by consulting the Medical Subject Headings (MeSH) thesaurus of the U.S. National Library of Medicine, combined with free-text terms in titles and abstracts to ensure comprehensive coverage.

The complete search string was constructed by combining three thematic keyword groups using Boolean operators (AND/OR), with no temporal restrictions applied:

("Autism Spectrum Disorder"[Mesh] OR autism[Title/Abstract] OR ASD[Title/Abstract] OR "autism spectrum disorder"[Title/Abstract])
AND
("Environmental Exposure"[Mesh] OR exposome[Title/Abstract] OR "environmental pollutants"[Mesh] OR "air pollution"[Mesh] OR "pesticides"[Mesh] OR "metals, heavy"[Mesh] OR "endocrine disruptors"[Mesh] OR "prenatal exposure delayed effects"[Mesh] OR environmental exposure[Title/Abstract] OR air pollution[Title/Abstract] OR pesticides[Title/Abstract] OR heavy metals[Title/Abstract] OR endocrine disruptors[Title/Abstract] OR prenatal exposure[Title/Abstract] OR ambient air[Title/Abstract] OR traffic-related[Title/Abstract] OR phthalates[Title/Abstract] OR PFAS[Title/Abstract] OR PCBs[Title/Abstract] OR noise[Title/Abstract] OR built environment[Title/Abstract])
AND
("Machine Learning"[Mesh] OR "Deep Learning"[Mesh] OR "Artificial Intelligence"[Mesh] OR "Computational Biology"[Mesh] OR "Algorithms"[Mesh]

OR machine learning[Title/Abstract] OR deep learning[Title/Abstract] OR artificial intelligence[Title/Abstract] OR neural network[Title/Abstract] OR random forest[Title/Abstract] OR support vector machine[Title/Abstract] OR gradient boosting[Title/Abstract] OR XGBoost[Title/Abstract] OR predictive model[Title/Abstract] OR algorithm[Title/Abstract] OR computational[Title/Abstract] OR data mining[Title/Abstract] OR big data[Title/Abstract] OR feature selection[Title/Abstract] OR exposure assessment[Title/Abstract] OR risk prediction[Title/Abstract])

2 Inclusion and Exclusion Criteria

Studies were included if they met all of the following criteria: (1) peer-reviewed original research articles; (2) environmental etiology research in ASD; (3) employing AI, ML, DL, or other computational methods.

Studies were excluded if they met any of the following criteria: (1) not related to ASD or environmental exposures; (2) lacking AI or computational methodologies; (3) including but not limited to non-population studies such as animal or cell research; (4) review, conference abstracts, editorials, case reports, commentaries, or study protocols; (5) duplicate publications; (6) not published in English; (7) full text not available.

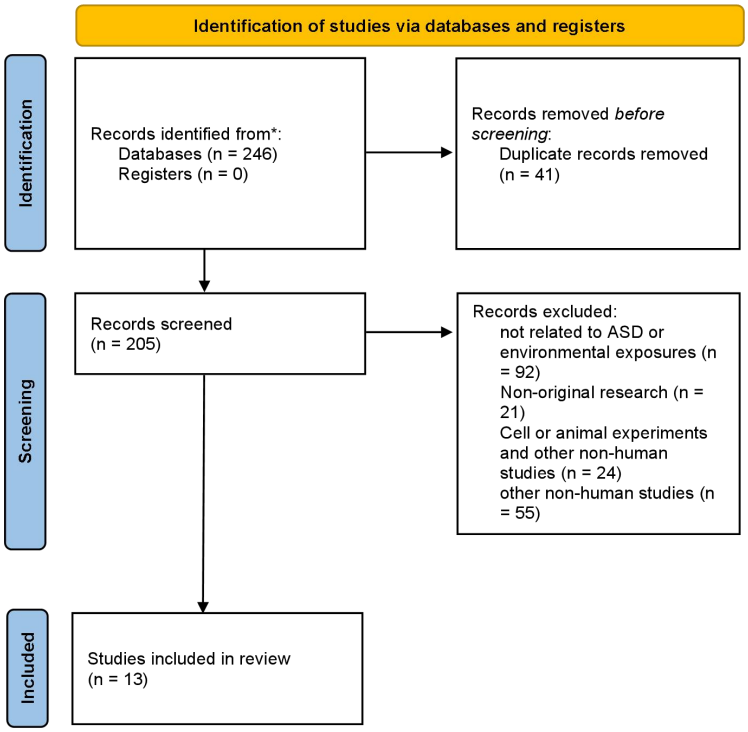
3 Screening Process

The retrieved records were exported to Zotero for literature management and duplicate removal. After removing 41 duplicate records, 205 unique records remained. Two independent reviewers screened the titles and abstracts of these 205 records against the inclusion criteria. At this stage, 192 records were excluded: 92 were not related to ASD or environmental exposures, 21 were non-original research, 24 were non-human studies, and 55 lacked AI or computational methodologies. The remaining 13 full-text articles were assessed for eligibility. Disagreements between reviewers were resolved through discussion or consultation with a third reviewer.

Ultimately, all 13 full-text articles met the inclusion criteria and were included in this review. The screening process is illustrated in the PRISMA flow diagram [Supplementary Figure 1].

4 Data Extraction and Synthesis

For each included study, key information was extracted, including study design, exposure types, AI algorithms employed, model performance metrics, and main findings. Due to heterogeneity in study designs, exposure measurements, and AI methodologies, a narrative synthesis approach was adopted rather than a quantitative meta-analysis.



Supplementary Figure 1. Study identification and selection.

REFERENCE

1. Page, M. J.; McKenzie, J. E.; Bossuyt, P. M.; *et al.* The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *J. Clin. Epidemiol.* **2021**, *134*, 178-89. https://pubmed.ncbi.nlm.nih.gov/33789819/.