TU Delft

Delft University of Technology

Document Version
Final published version

Licence
CC BY

Citation (APA)
Equal-Life scientific team, Peuters, C., & Cheval, B. (2026). Prenatal environmental exposures associated with ADHD diagnosis in children: a multicentre exposome analysis. *Environment International*, *214*, Article 110415. https://doi.org/10.1016/j.envint.2026.110415

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

Environment International 214 (2026) 110415

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Full length article

Prenatal environmental exposures associated with ADHD diagnosis in children: a multicentre exposome analysis

Libor Šulc [a,*], Fan Wu [a], Sanne Meijering [b], Hendriek Boshuizen [b], Angel M. Dzhambov [c,d], Peter Lercher [c], Carmen Peuters [e,f,g], Boris Cheval [h], Nuria Botella [e], Payam Dadvand [e,f,g], Jordi Júlvez [i], Jolanda M.A. Boer [b], Christoph Giehl [j], Jan Spilski [j], Yingxin Chen [k,l], Anna L. Hansell [k,l,m], Rik Bogers [b], Tanja Vrijkotte [n,o,p], Gabriele Bolte [q], Achilleas Psyllidis [r], Karl Lundin Remnélius [s], Sven Bölte [s,t,u], Elise van Kempen [b], Irene van Kamp [b], Jenny Selander [a], Equal-Life scientific team

[a] Unit of Occupational Medicine at the Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
[b] National Institute for Public Health and the Environment, Bilthoven, the Netherlands
[c] Institute of Highway Engineering and Transport Planning, Graz University of Technology, Graz, Austria
[d] Environmental Health Division, Research Institute at Medical University of Plovdiv, Medical University of Plovdiv, Plovdiv, Bulgaria
[e] ISGlobal, Barcelona, Spain
[f] Universitat Pompeu Fabra (UPF), Barcelona, Spain
[g] Spanish Consortium for Research on Epidemiology and Public Health (CIBERESP), Spain
[h] Univ Rennes, École normale supérieure de Rennes, VIPS^2, F-35000 Rennes, France
[i] Clinical and Epidemiological Neuroscience (NeuroÈpia), IISPV, Reus, Spain
[j] Center for Cognitive Science, RPTU University Kaiserslautern-Landau, Kaiserslautern, Germany
[k] Centre for Environmental Health and Sustainability, Division of Epidemiology and Public Health, School of Medical Sciences, University of Leicester, Leicester, UK
[l] The National Institute of Health Research (NIHR) Health Protection Research Unit (HPRU) in Chemical Threats and Hazards at the University of Leicester, Leicester, UK
[m] National Institute for Health Research (NIHR) Leicester Biomedical Research Centre (BRC), Leicester General Hospital, Leicester, UK
[n] Department of Public and Occupational Health, Amsterdam UMC, location University of Amsterdam, Amsterdam, the Netherlands
[o] Amsterdam Reproduction & Development Research Institute, Amsterdam, the Netherlands
[p] Amsterdam Public Health Research Institute, Amsterdam, the Netherlands
[q] University of Bremen, Institute of Public Health and Nursing Research, Department of Social Epidemiology, Bremen, Germany
[r] Department of Sustainable Design Engineering, Faculty of Industrial Design Engineering, Delft University of Technology, Delft, the Netherlands
[s] Center of Neurodevelopmental Disorders (KIND), Department of Women's and Children's Health, Centre for Psychiatry Research, Karolinska Institutet & Region Stockholm, Stockholm, Sweden
[t] Child and Adolescent Psychiatry, Stockholm Health Care Services, Region Stockholm, Stockholm, Sweden
[u] Curtin Autism Research Group, Curtin School of Allied Health, Curtin University, Perth, Australia

ARTICLE INFO

Keywords:
ADHD
Exposome
Prenatal Exposure
Children
Socioeconomic Factors

ABSTRACT

Attention deficit hyperactivity disorder (ADHD) is one of the most diagnosed neurodevelopmental conditions. While its aetiology is primarily attributed to genetic factors, the complex interplay with environmental exposures remains poorly understood. We investigated the relationship between prenatal environmental exposures (excluding biomarkers) and ADHD, utilizing data from seven cohorts (n = 156,563 children) across Europe.

A supervised machine learning model (Random Forest) was run at each cohort site to identify the environmental exposures most informative for classifying ADHD. The results were pooled by averaging variable importance ranks across cohorts and highest-ranking variables were then used in logistic regression to obtain interpretable associations. Finally, a random-effects *meta*-analysis with restricted maximum likelihood estimation was used to infer the associations across cohorts.

Based on Random Forest, built and natural urban characteristics, air pollution, and maternal socioeconomic factors were selected as variables with the greatest predictive contribution to ADHD classification. Associations with built and natural urban characteristics and air pollution were not robust across cohorts and largely attenuated in fully adjusted models. Meta-analysis across cohorts showed increased odds of ADHD for maternal

* Corresponding author at: C6 Institutet för miljömedicin, C6 Arbetsmedicin Selander, 171 77 Stockholm, Sweden.
E-mail address: libor.sulc@ki.se (L. Šulc).

https://doi.org/10.1016/j.envint.2026.110415
Received 10 March 2026; Received in revised form 8 July 2026; Accepted 10 July 2026
Available online 11 July 2026
0160-4120/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

smoking during pregnancy (OR = 1.67, 95% CI: 1.31, 2.14, $I^2 = 65.5\%$) and low maternal education (OR = 1.62, 95% CI: 1.27, 2.07, $I^2 = 51.3\%$), with substantial heterogeneity.

Our findings indicated that maternal smoking during pregnancy and low maternal education were associated with increased odds of ADHD diagnoses in children, suggesting the potential role of socioeconomic disparities. However, gene-environment correlation cannot be ruled out.

1. Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is one of the most frequently diagnosed neurodevelopmental conditions, with a worldwide prevalence of around 5.3% among school-aged children (Aghaei et al., 2019). ADHD is often diagnosed between the ages of 6 and 18 when related challenges emerge in school and other social settings – although, for various reasons, in girls, ADHD diagnoses are often delayed by several years, compared to boys (Rocco et al., 2021; Williams et al., 2025). While ADHD was long considered a childhood condition, it is increasingly recognised as a condition that affects individuals throughout their whole lifespan (Sharma and Couture, 2014).

ADHD is defined by patterns of age-inappropriate inattention, impulsivity, and hyperactivity, associated with problems in functioning across major life arenas, such as education, family, school performance, peer relationships, and workplace. ADHD increases risk for mental health issues, accidents, criminality, premature mortality, and suicide (Faraone et al., 2021; Sonuga-Barke et al., 2023). International guidelines recommend multimodal intervention for ADHD, including a combination of both pharmacological and behavioural treatment (Coghill et al., 2023).

ADHD is an etiologically complex condition with both genetic and environmental factors and their interaction contributing. Genetic factors are considered a major contributor to ADHD development (Demontis et al., 2023), with a heritability of up to 80% (Faraone et al., 2024). Nevertheless, the contribution of environmental factors, on the other hand, is estimated to be 10–40% (Oudin et al., 2019). Environmental factors may not only play a role in the causes of ADHD, but also in whether individuals are referred to services and the likelihood of getting diagnosed. Prenatal and early-life exposures, including maternal factors, exposure to environmental chemicals, and nutritional deficiencies, may impact the development of ADHD (Carlsson et al., 2021). Gaining a better understanding of the role of prenatal exposures could be essential for reducing ADHD-related risks, thereby promoting preventative strategies of early adversities to improve both individual outcomes and societal costs (Aghaei et al., 2019; Mikeš et al., 2023; Sciberras et al., 2017).

To address this complexity and better characterize the multifactorial environmental contributions to neurodevelopmental disorders such as ADHD, the exposome framework was proposed by Wild in 2005 (Wild, 2005). The exposome aims to capture the totality of environmental exposures that an individual encounters throughout life, beginning in the prenatal period, thereby complementing genetic approaches to disease aetiology. It encompasses a broad range of external exposures such as air pollution, built environment, socioeconomic conditions, and lifestyle factors as well as internal markers measured in biological samples. By integrating information of multiple exposures across different life domains, the exposome framework provides a holistic perspective on how complex exposure patterns shape health and development (Gutierrez-Ortiz et al., 2025). With the growing adoption of the exposome framework, research has expanded beyond traditional physical health outcomes, increasingly including neurodevelopmental and psychiatric conditions into exposome-based investigations (Erzin and Guloksuz, 2021). The present study focuses on the general and specific external exposome, specifically environmental and socioeconomic exposures during pregnancy, and does not include internal exposome.

This work focuses on analysing data from more than 150,000 children collected in seven cohorts, all part of the Equal-Life project, which is a large European effort to understand how children's physical and social environments influence their mental health and cognitive development from conception to early adulthood (van Kamp et al., 2022). The main aims were to i) identify prenatal exposures associated with ADHD, ii) assess the relationship between the identified exposures and ADHD diagnosis for each study population, and iii) assess the consistency of these relationships across all analysed studies. By utilising diverse datasets and employing robust statistical methods, this study aimed to investigate the prenatal factors influencing neurodevelopment towards ADHD.

2. Methods

2.1. Study populations

Altogether, seven European population-based cohorts and cross-sectional studies involving children were included: ABCD, the Netherlands (van Eijsden et al., 2011), ALPINE, Austria (Dzhambov et al., 2019), ALSPAC, United Kingdom (Fraser et al., 2013), BREATHE, Spain (Forns et al., 2016), FAIR, Sweden (Skröder et al., 2021), PIAMA, the Netherlands (Wijga et al., 2014), and WALNUTs, Spain (Julvez et al., 2021). The total number of participants was 156,563, ranging from 645 (WALNUTs) to 138,971 (FAIR). More details on each study were included in the **SI** (Supplementary Information) Table S1.

2.2. ADHD diagnosis assessment

Three main sources of data on ADHD diagnoses were used: clinical register data, parental reports, and teacher assessments. Parental reports were used in ABCD, PIAMA, ALSPAC, and WALNUTs. In ABCD, mothers reported whether their child had long-term medical, behavioural, or school problems, with ADHD listed among the options (children aged 11–12, 15–16). In PIAMA, parents were asked whether their child had ever been diagnosed with ADHD (children aged 14), in ALSPAC, mothers reported whether their child had ADHD symptoms using DSM-IV (Willcutt, 2012) criteria for children aged 10 years old, and WALNUTS asked whether a doctor or psychologist had diagnosed the child, and to specify the condition (children aged 11–16). Teacher assessments were used in BREATHE and ALPINE. In BREATHE, teachers rated ADHD symptoms using DSM-IV criteria across 18 items (0–3 scale, total 0–54, children aged 7–11). In ALPINE, teachers of children aged 8–11 years completed the Needleman questionnaire (Ullmann et al., 1984) with yes/no responses indicating ADHD status. FAIR captured physician and psychologist diagnoses using the National Medical Birth Register and Patient Register, which covered nearly all deliveries in Sweden (children aged 0–12). These registers include child diagnosis in the form of International Classification of Diseases codes (10th version). Statements on ethical approvals for each study were provided in the Supplementary Information.

2.3. Prenatal exposome characterisation

Two types of data sources on the exposome were used: 1) cohort-specific data curated in the framework of each study (e.g., air pollution, traffic noise), and 2) enriched data generated for most of the studies for the needs of the Equal-Life project (e.g., land use).

Cohort-specific data covered indicators such as parental education,

household income, number of adults living in the household, maternal smoking during pregnancy, and parental occupation. These indicators were extracted from cohort-specific questionnaires or national registers. Furthermore, indicators like NO_2 or road/plane/rail noise were estimated based on national monitoring services.

Enriched data included multiple land use or neighbourhood-level indicators. These included the population density, building density, elevation, and slope, land use mix (e.g., urban green land, agricultural land, water land), roads (e.g., lengths of the streets, distance to the closest road), connectivity (e.g., closeness and number of intersections, betweenness), green and blue areas (e.g., access to major blue areas, distance to closest green areas), tree cover, and neighbourhood-level socioeconomic indicators (e.g., age structure, unemployment rate, net migration rate). Mentioned indicators were often generated in multiple iterations based on buffer area around the geocode (residence during pregnancy, 100/300/500 m) and year (based on data availability). Data sources of the indicators were included in **SI** Table S2, along with files containing detailed information on data preparation, harmonisation, and enrichment.

2.4. Data analysis

A significant number of variables, both from physical and social exposome, were collected across the seven cohorts, including the data enrichment efforts within the Equal-Life project. Variables available only in one cohort were removed from the selection, as well as variables with a high proportion of missing data (>25%). Iterations of the same variables, with a buffer radius of 100 m and 500 m around a geocode (home address), were removed in favour of the 300 m buffer (Pratt et al., 2014; Who, 2017). Likewise, many enriched variables were year or period-specific; thus, only the variable relevant to the birth year or closest to the birth year of each child in each cohort was used. Finally, variables not relevant to the prenatal period were removed (e.g., breastfeeding, child dietary habits, peer relationships). Per study, 46–49 variables were available for analysis (ALPINE and ALPSAC had only 24 and 29 variables, respectively), with varying levels of overlap between studies. Variable availability, overlap, and description of each variable were included in **SI Table S3**.

A machine learning algorithm – Random Forest (RF) was employed to perform two tasks: 1) to further reduce the pool of variables, and 2) to identify the variables most informative for classifying ADHD diagnosis. The separate results of RF produced at different locations were centrally integrated to procure a list of variables with the greatest predictive contribution to ADHD diagnosis across all the cohorts. As a next step, logistic regression was run to estimate the risk of ADHD associated with each variable. In the final step, estimates of the logistic regression were centrally *meta*-analysed to glean the overall effect estimates. All analyses were done in the R programming language.

2.4.1. Random forest

Two variations of RF were used: Classification and Regression Trees (RF-CART) (Lampa et al., 2014) and Conditional Inference Trees (RF-CI) (Strobl et al., 2008). RF-CART prioritised continuous variables in its ranking, while RF-CI incorporated a statistical hypothesis testing framework for variable selection, resulting in a bias toward categorical variables. This dual approach complemented the selection of both continuous and categorical variables. Additionally, RF-CART is robust to outlying values (Breiman, 2001), while RF-CI may be to a degree influenced by them, though this affects only the choice of splitting variable not the location of the split (Hothorn et al., 2006). Therefore, the use of both methods serves as a form of sensitivity analysis with respect to outliers. The analysis was performed separately using RF-CART and RF-CI. A shared R code using the *missForest* (Stekhoven and Bühlmann, 2012); *ranger* (Wright and Ziegler, 2017), and *partykit* (Hothorn et al., 2006) packages, was written for the RF analysis (Meijering, 2026). Variables with up to 25% missing values were imputed, while variables exceeding this threshold were excluded from the analysis (Madley-Dowd et al., 2019). Each data owner performed the analysis separately to avoid issues associated with data privacy and analysis inconsistency. RF-CI analysis on FAIR cohort exceeded available computational power. To address this, the FAIR cohort was randomly divided into five subsets of approximately 28,000 participants each. The proportion of ADHD cases was preserved in each subset to maintain representativeness of the outcome. The RF-CI analysis was then run separately on each subset, producing five sets of variable importance rankings. These were subsequently pooled into a single FAIR-level result using the same rank-averaging approach that was applied across cohorts (see below). The pooled result was included in the overall cross-cohort analysis alongside the results from the remaining six cohorts. In each RF run across all studies, permutation variable importance (Altmann et al., 2010) was given to all analysed variables (the higher the importance, the greater the variable's contribution to classification of ADHD diagnosis). Based on permutation variable importance, *more reliable* and *less reliable* variables for the prediction of ADHD diagnosis were identified. Variables with a permutation variable importance below the lowest absolute permutation variable importance were deemed *less reliable* (Strobl et al., 2009). Analysed variables were then ranked in order of importance (highest importance = rank 1). Any *less reliable* results were attributed with the lowest rank possible in the specific RF run, signifying lower predictive contribution to ADHD classification. Thus, all analysed variables were included, but the increased uncertainties were accounted for.

To pool the results from all studies, the ranks of the same variables were averaged to produce an overall ranking of variables by their predictive contribution to ADHD classification (Rengasamy et al., 2021). Variables appearing in fewer than four cohorts (half), regardless of rank, were excluded to further compare variables available in as many studies as possible, while still being relevant for ADHD diagnosis. This step was performed separately for RF-CART and RF-CI results, as each prioritized different sets of variables and thus yielded different results.

2.4.2. Logistic regression

Pooled RF results focusing on the 15 variables with the greatest predictive contribution to ADHD classification (RF-CART and RF-CI separately) served as a guide for variable selection in the logistic regression. The number of covariates included in each logistic regression model was limited to a maximum of 15. This specification was well within contemporary methodological recommendations (Vittinghoff and McCulloch, 2007), which emphasized model complexity, outcome prevalence, and predictor correlations rather than rigid "events-per-variable" thresholds as the key determinants of model stability. Given the large sample size and the substantial number of ADHD cases in our dataset, including up to 15 independent variables did not meaningfully increase the risk of overfitting. Moreover, several covariates were categorical with limited levels, further reducing the effective number of parameters. Sensitivity analyses confirmed that model estimates were stable across different variable sets.

Previous knowledge of exposure linked with ADHD in children was used as a complement to the pooled results to produce a more robust selection of exposures. The following rules were then applied for the final selection: 1) include demographics as adjusting variables, 2) prioritize variables appearing in both RF-CART and RF-CI, 3) include variables from diverse exposure subcategories (e.g., built environment, socioeconomics), 4) prioritize high-ranking variables over lower-ranking variables, 5) avoid overly complex models, and 6) consider existing literature. Logistic regression with selected variables (exposures) and ADHD as the dependent variable was run. Beta coefficients were expressed as odds ratios (OR) with a 95% confidence interval (CI). Increment of change for each variable (see **SI Table S4**) used to calculate OR + 95% CI was based on the units each variable was measured in. Data multicollinearity was assessed based on the generalized variance inflation factor (**SI** Table S12) (Fox and Monette, 1992). We applied

multiple imputation using the *mice* package (Buuren and Groothuis-Oudshoorn, 2011), generating five imputed datasets with five iterations per imputation using classification and regression trees. Imputation was restricted to variables with ≤ 25% missing data (Madley-Dowd et al., 2019), while variables exceeding this threshold were excluded from the analysis. Pooling of the regression results was done according to the Rubin's rules (Little and Rubin, 2019).

Demographic indicators were considered a distinct group of variables that were treated as covariates. These included the child's birth year and sex, and the mother's age at birth and country of birth. In total, three models were run: basic model (each exposure analysed separately, including covariates), fully adjusted model (all exposures and covariates), and harmonised model (including only exposures and covariates available in all involved cohorts). To evaluate the robustness of our exposure selection strategy, we ran sensitivity analyses in which specific exposures were replaced by alternative variables suggested by the RF-CART and RF-CI importance rankings.

A standardized script for logistic regression and imputation was first distributed to all data owners, who applied it locally to their respective datasets (Šulc, 2026). The resulting study-specific estimates were pooled centrally using a random-effects *meta*-analysis. Meta-analysis used the restricted maximum likelihood estimator to account for both within-study and between-study variability. This approach assumes that the true effect size may vary across studies due to differences in study populations or methodologies and provides a weighted average of the effect estimates while incorporating this heterogeneity. To assess heterogeneity across studies, we calculated Cochran's Q statistic, the I^2 statistic, and the between-study variance (τ^2). The *meta*-analyses were performed using the *metafor* package in R (Viechtbauer, 2010). To investigate methodological heterogeneity, a subgroup *meta*-analysis was conducted, stratifying cohorts into parent reported, teacher assessment, and non-parent assessment of ADHD diagnosis. Meta-regression further testing the effect of study size, percent of ADHD cases, median year of the study, and ADHD diagnosis assessment method (parent, teacher, registry) was done.

To assess the potential impact of unmeasured confounding, E-values were calculated for all found associations (OR and lower 95% CI bound), including both individual cohort estimates and *meta*-analysis results. The E-value represents the minimum strength of association an unmeasured confounder would need to have with both the exposure and the outcome to explain the observed association (VanderWeele and Ding, 2017).

3. Results

3.1. Descriptive statistics of study populations

See Table 1 for an overview of the study populations. The percentage of ADHD diagnoses/ADHD status ranged from 1.6% (ALSPAC) to 13.8% (ALPINE). In all studies, boys had a noticeably higher number of positive diagnoses as compared to girls. Median maternal age at birth spanned from 28 years in ALPINE to 33 years in ABCD and WALNUTs. Smoking during the first trimester of pregnancy ranged from 4.9% (FAIR) to 17.2% (ALSPAC), highlighting varying maternal health behaviours across populations and social groups. Education levels also varied across cohorts, with ABCD having the highest proportion of highly educated people (67.7%) and ALSPAC the lowest (15.5%). At the same time, ALSPAC also had the highest proportion of the least educated people (53.3%). Mothers' country of birth was the study country in > 79% of participants across cohorts. Children from ALSPAC, ALPINE, and PIAMA were born between 1991 and 1997, children from WALNUTs, ABCD, and BREATHE between 2000 and 2004, and children from FAIR between 2007 and 2012.

3.2. Integrated results from random forest

Information on the RF analyses (permutation variable importance, rank) for each included cohort was provided in **SI** **Table S5-S7**. As expected, RF-CART prioritised continuous variables while RF-CI prioritised the categorical ones. The full integrated results of both RF types were provided in **SI** Figs. S1 and S2.

Overall, both methods identified a mix of environmental, built environment, and sociodemographic factors as important, although their relative importance differed by approach. In the RF-CART model, the highest-ranked variables were tree cover, followed by distance to major roads (Major road), general road distance, and residential NO_2 concentrations. Population and topographic measures, including population count and elevation, and low-density residential areas (LDR) were also prominent. Additional contributions came from indicators of surrounding land cover and greenness, such as the normalized difference vegetation index (NDVI), natural land, distance to green space, and distance to blue space (Blue distance).

RF-CI model prioritised sociodemographic and familial characteristics. The most important variable was father's and mother's education, mother's occupation code, smoking during pregnancy (Smoking), and the number of adults in the household. Environmental variables were also represented, with population density, area of green space, LDR, tree cover, and elevation ranked as important. Average rank and more details on each selected variable are presented in **SI** **Table S8**.

Based on the pooled results and the rules stated above, the tree cover,

Table 1
Descriptive statistics of the study populations.

Statistic	Study						
	ABCD	ALPINE	ALSPAC	BREATHE	FAIR	PIAMA	WALNUTs
Study sample size	2,853	1,251	7,683	2,852	138,971	2,308	645
ADHD – positive diagnosis (%)	3.8	13.8	1.6	3.8	7.7	7.1	9.2
of which (% boys)	75.9	65.9	50.2	72.2	69.5	73.2	76.3
Mothers' age at birth – median (5th percentile, 95th percentile)	33 (25, 40)	28 (20, 37)	29 (21, 37)	–	32 (23, 40)	33 (25, 37)	33 (26, 41)
Mothers' age at birth – NA (%)	10.8	1.5	6.8	–	0	1.2	9.0
Mother smoking during pregnancy (%)	8.8	11.9	17.2	9.1	4.9	16.5	10.5
Mother smoking during pregnancy – NA (%)	0	2.0	5.8	7.3	0	0.8	3.1
Mothers' education – low (%)	6.8	22.3	53.3	12.1	9.3	17.5	6.7
Mothers' education – mid (%)	16.9	54.6	24.2	27.3	31.3	41.6	32.7
Mothers' education – high (%)	67.7	23.1	15.5	55.7	58.1	40.6	60.5
Mothers' education – NA (%)	8.7	0	7.1	4.9	1.4	0.3	0.2
Mothers' country of birth (%)*	79.9	–	91.1	84.1	86.6	94.7	83.7
Mothers' country of birth – NA (%)*	0.0	–	7.4	0.4	0	1.1	1.4
Birthyear (children)	2003–2004	1992–1997	1991–1994	2001–2004	2007–2012	1996–1997	2000–2003

*Country of birth is the same as the country where the study is being conducted.

LDR, elevation, NO_2, Blue distance, Major roads, population count, the number of adults in the household, Smoking, and the mother's education were selected as the main variables of interest. The secondary variables of interest, used in the sensitivity analyses, were road distance, population density, built-up area, NDVI, and household income. More details on each variable and the process of its selection were provided in the Supplementary Information.

Table 2
Results of the logistic regression using the variables indicated by the RF-CART and RF-CI and ADHD.

Basic models														
Study	**ABCD**		**ALPINE**		**ALSPAC**		**BREATHE**		**FAIR**		**PIAMA**		**WALNUTs**	
Variable	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
Blue distance	1.02	[0.88, 1.19]	1.00	[0.99, 1.01]	n/a	n/a	0.99	[0.94, 1.04]	**1.04**	**[1.03, 1.05]**	0.99	[0.97, 1.02]	1.06	[0.97, 1.15]
Mother's Education	**2.06**	**[1.33, 3.19]**	**1.47**	**[1.04, 2.08]**	1.37	[0.87, 2.17]	**2.58**	**[1.76, 3.78]**	**1.73**	**[1.66, 1.81]**	1.22	[0.89, 1.67]	**3.44**	**[1.81, 6.54]**
Elevation	0.48	[0.21, 1.07]	1.00	[0.99, 1.01]	n/a	n/a	1.00	[0.96, 1.04]	**1.09**	**[1.07, 1.11]**	0.77	[0.59, 1.01]	1.03	[0.98, 1.10]
Built-up area	0.81	[0.59, 1.11]	0.98	[0.79, 1.21]	n/a	n/a	0.95	[0.70, 1.29]	**0.89**	**[0.86, 0.92]**	1.02	[0.83, 1.26]	0.73	[0.47, 1.12]
Population count	1.01	[0.55, 1.88]	0.87	[0.44, 1.72]	n/a	n/a	0.98	[0.82, 1.17]	**0.90**	**[0.89, 0.92]**	0.95	[0.64, 1.40]	0.87	[0.65, 1.16]
Major roads	1.31	[0.76, 2.25]	0.99	[0.90, 1.09]	n/a	n/a	1.02	[0.65, 1.59]	**1.02**	**[1.01, 1.03]**	0.95	[0.75, 1.20]	0.89	[0.31, 2.55]
Number of adults	n/a	n/a	n/a	n/a	n/a	n/a	0.95	[0.64, 1.40]	**0.62**	**[0.59, 0.66]**	1.31	[0.92, 1.87]	0.55	[0.29, 1.05]
NDVI	**1.34**	**[1.02, 1.75]**	1.13	[0.86, 1.48]	0.85	[0.60, 1.19]	1.10	[0.71, 1.70]	**1.17**	**[1.14, 1.20]**	0.96	[0.75, 1.23]	1.54	[0.71, 3.35]
NO_2	0.97	[0.72, 1.31]	0.96	[0.80, 1.16]	1.22	[0.73, 2.04]	1.03	[0.86, 1.24]	**0.70**	**[0.67, 0.73]**	1.15	[0.93, 1.42]	0.83	[0.66, 1.04]
Population density	1.00	[0.97, 1.03]	n/a	n/a	1.04	[0.97, 1.12]	n/a	n/a	**0.94**	**[0.93, 0.95]**	0.94	[0.82, 1.07]	n/a	n/a
Smoking	1.41	[0.77, 2.58]	**1.60**	**[1.02, 2.51]**	**1.86**	**[1.24, 2.77]**	**2.66**	**[1.55, 4.56]**	**2.31**	**[2.16, 2.48]**	1.07	[0.79, 1.43]	**3.63**	**[1.78, 7.42]**
Tree cover	0.65	[0.24, 1.76]	0.85	[0.66, 1.10]	n/a	n/a	1.34	[0.63, 2.83]	**1.09**	**[1.05, 1.12]**	1.28	[0.89, 1.85]	1.62	[0.42, 6.35]
LDR	**3.73**	**[1.60, 8.70]**	n/a	n/a	1.00	[0.99, 1.01]	1.05	[0.61, 1.82]	**1.16**	**[1.13, 1.20]**	0.96	[0.59, 1.56]	1.69	[0.65, 4.42]
Fully adjusted models														
Blue distance	1.04	[0.88, 1.23]	0.99	[0.97, 1.01]	n/a	n/a	0.97	[0.87, 1.07]	1.00	[0.99, 1.01]	1.00	[0.97, 1.03]	1.09	[0.93, 1.28]
Mother's Education	**1.82**	**[1.15, 2.88]**	**1.50**	**[1.05, 2.14]**	1.06	[0.63, 1.78]	**2.38**	**[1.60, 3.52]**	**1.46**	**[1.39, 1.53]**	1.22	[0.88, 1.69]	**2.95**	**[1.47, 5.90]**
Elevation	0.47	[0.14, 1.57]	1.00	[0.99, 1.02]	n/a	n/a	1.01	[0.94, 1.09]	**1.06**	**[1.04, 1.08]**	0.76	[0.57, 1.02]	0.99	[0.82, 1.18]
Population count	1.11	[0.50, 2.50]	0.66	[0.28, 1.58]	n/a	n/a	0.96	[0.70, 1.33]	**0.95**	**[0.93, 0.98]**	0.81	[0.50, 1.33]	1.10	[0.58, 2.07]
Major roads	1.14	[0.62, 2.10]	0.96	[0.84, 1.08]	n/a	n/a	1.20	[0.62, 2.32]	1.00	[0.99, 1.01]	0.99	[0.77, 1.26]	0.49	[0.10, 2.41]
Number of adults	n/a	n/a	n/a	n/a	n/a	n/a	0.89	[0.61, 1.29]	**0.71**	**[0.68, 0.76]**	1.35	[0.94, 1.94]	0.58	[0.29, 1.14]
NO_2	1.32	[0.90, 1.93]	0.88	[0.67, 1.15]	1.21	[0.73, 2.03]	1.12	[0.93, 1.36]	**0.84**	**[0.78, 0.90]**	1.20	[0.92, 1.55]	0.78	[0.53, 1.15]
Smoking	1.29	[0.70, 2.39]	1.52	[0.96, 2.41]	**1.84**	**[1.09, 3.08]**	**2.12**	**[1.20, 3.76]**	**1.75**	**[1.62, 1.88]**	1.01	[0.74, 1.36]	**3.04**	**[1.32, 6.96]**
Tree cover	1.12	[0.35, 3.58]	0.74	[0.55, 1.01]	n/a	n/a	1.31	[0.34, 5.06]	**0.93**	**[0.89, 0.97]**	1.45	[0.95, 2.21]	0.68	[0.05, 8.66]
LDR	**3.41**	**[1.32, 8.81]**	n/a	n/a	1.00	[0.99, 1.01]	1.12	[0.48, 2.62]	0.99	[0.95, 1.03]	0.99	[0.55, 1.76]	1.36	[0.20, 9.43]
Harmonised models														
Mother's Education	1.53	[0.99, 2.35]	**1.44**	**[1.01, 2.05]**	1.19	[0.72, 1.96]	**2.39**	**[1.65, 3.46]**	**1.61**	**[1.54, 1.68]**	1.30	[0.95, 1.78]	**2.70**	**[1.41, 5.15]**
NDVI	1.18	[0.87, 1.60]	1.97	[0.95, 4.09]	0.87	[0.54, 1.40]	1.16	[0.57, 2.34]	1.01	[0.96, 1.05]	1.13	[0.75, 1.70]	1.47	[0.27, 8.08]
NO_2	1.19	[0.84, 1.67]	0.97	[0.78, 1.21]	1.22	[0.73, 2.03]	1.11	[0.93, 1.33]	**0.86**	**[0.80, 0.92]**	1.24	[0.93, 1.64]	0.83	[0.58, 1.18]
Smoking	1.35	[0.73, 2.47]	1.56	[0.99, 2.48]	**1.91**	**[1.14, 3.20]**	**2.10**	**[1.20, 3.67]**	**1.92**	**[1.79, 2.06]**	1.04	[0.77, 1.40]	**2.98**	**[1.40, 6.35]**
Grey space – Built-up area	0.87	[0.58, 1.30]	1.64	[0.91, 2.94]	n/a	n/a	n/a	n/a	n/a	n/a	1.00	[0.75, 1.34]	n/a	n/a
Grey space – Population count	n/a	n/a	n/a	n/a	n/a	n/a	0.94	[0.72, 1.23]	**0.97**	**[0.94, 0.99]**	n/a	n/a	1.12	[0.61, 2.05]
Grey space – Population density	n/a	n/a	n/a	n/a	1.03	[0.93, 1.13]	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

OR: odds ratio; 95% CI: 95% confidence interval; LDR: low-density residential areas; NDVI: normalized difference vegetation index; Basic model: included each tested variable separately and the child's birth year and sex, and the mother's age at birth and country of birth; Fully adjusted model: included all the variables together and the child's birth year and sex, and the mother's age at birth and country of birth; Harmonized model: included only the variables available in all the involved cohorts and the child's birth year and sex.

3.3. Logistic regression and meta-analysis

Descriptive statistics of the exposome variables were included in **SI Table S9** and correlation matrices in **SI** Table S10. Detailed results of the logistic regression are shown in Table 2. In the basic models, all the exposures were tested separately, adjusting only for demographic indicators (the child's birth year and sex, and the mother's age at childbirth and country of birth). When inspected separately, many of the basic models provided associations suggesting increased odds of ADHD in association with various exposures. Apparent was the effect of maternal smoking during pregnancy and (low) maternal education, indicating an association with higher odds of ADHD, significant across most studies. LDR was also indicated as a risk factor for ADHD. This was, however, seen only in ABCD and FAIR cohorts. In the same cohorts, the results of NDVI counterintuitively suggested an adverse association of this green space indicator with ADHD. Similar effects were observed for blue space and tree cover.

In fully adjusted models, the most consistent associations across cohorts were those with exposure to maternal smoking during pregnancy and the (low) educational status of the mother. In ABCD, LDR was still associated with increased ADHD risk, but this was no longer the case in FAIR. No other associations were significant in any cohort, except for the FAIR cohort. In FAIR, NO_2, tree cover, population count, and the number of adults were associated with decreased odds of ADHD diagnosis. On the other hand, elevation was associated with an increased risk of ADHD. Across cohorts, the results for environmental exposures were inconsistent often with relatively wide CI, negative or positive point estimates depending on the cohort.

In the harmonised model, only the exposures available in all the involved cohorts were included. This limited the selected variables to elevation, maternal smoking during pregnancy, maternal education, NO_2, NDVI, and an indicator of grey space (built-up areas, population count, or population density; these were treated equally). The results from this model were in line with the pattern of association provided thus far. Maternal smoking during pregnancy and lower maternal education were still associated with an increased odds of ADHD diagnosis, while no associations were found for other exposures. The only exceptions were NO_2 and population count, both of which showed inverse association in FAIR.

In the sensitivity analyses, we replaced some of the original variables with household income, road distance, built-up area, population density, and NDVI (**SI Table S9**). Overall, these substitutions produced results similar to the fully adjusted models. The most notable change appeared in the ALPINE cohort, where the OR for maternal smoking during pregnancy indicated higher odds of ADHD compared to the fully adjusted model. Slightly higher ORs were also seen in FAIR, BREATHE, and WALNUTs compared with the fully adjusted models for the same variable. Household income showed a pattern similar to maternal education: lower income was associated with higher odds of ADHD diagnosis, although the association was weaker than the association with education.

Finally, the results of the harmonised models were *meta*-analysed. Since most of the results did not indicate an association, only the results for mothers' education and smoking during pregnancy were presented. The remaining results from the *meta*-analyses were included in the **SI Figs. S3-S10**.

Meta-analysis of the results for mothers' education, including all the cohorts, showed an OR = 1.61, 95% CI [1.54, 1.68]. However, the FAIR cohort contributed to almost 93% of this weight. As a result, the pooled estimate was largely driven by a single cohort and therefore provided limited information about consistency across studies. When FAIR was excluded, the contribution of the remaining cohorts was noticeably more balanced, while a similar OR was observed as before, OR = 1.62, only with a wider 95% CI [1.27, 2.07] (Fig. 1). A moderate heterogeneity was observed ($Q = 10.44$, $p = 0.06$, $I^2 = 51.3\%$, $\tau^2 = 0.05$). For this reason, the FAIR-excluded model was treated as the primary result, as it more meaningfully reflects the degree of agreement across the contributing populations.

For maternal smoking during pregnancy, the weight of all the cohorts was more balanced, with FAIR contributing approximately 26% to the pooled OR (Fig. 2). When pooled, smoking during pregnancy increased the odds of ADHD diagnosis; OR = 1.67, 95% CI [1.31, 2.14]. Exclusion of the FAIR cohort had only a minor impact on the pooled result (OR = 1.60, 95% CI [1.19, 2.16]). Observed heterogeneity was higher compared to the previous *meta*-analysis ($Q = 18.78$, $p = 0.005$, $I^2 = 65.5\%$, $\tau^2 = 0.06$).

Although not statistically significant, the *meta*-analysis point estimate for NO_2 indicated an elevated risk of ADHD (**SI Fig. S8 and S9**). This was also the case for NDVI, but not for grey space. Notably, FAIR cohort results often add over 75% of the weight.

The strongest associations (maternal education, smoking) showed consistently higher E-values both in individual cohorts and in the *meta*-analysis estimates, where E-values for the point estimates were 2.60 and 2.73 respectively (**SI** Table S13). These values indicate that a moderately strong unmeasured confounder could potentially explain several of the observed associations, particularly those with E-values close to 1. Meanwhile the associations for maternal education and smoking would require a confounder with stronger association to be fully attributed to bias.

A subgroup *meta*-analysis stratifying cohorts into parent reported, teacher assessment, and non-parent assessment of ADHD diagnosis was provided in supplement (**SI** Fig. S11 to S14). For maternal education, both groups indicated an increased risk of ADHD. Heterogeneity was notably lower in the parent reported group ($Q = 4.64$, $p = 0.199$, $I^2 = 7.6\%$, $\tau^2 = 0.004$) compared to the teacher assessment group ($Q = 3.80$, $p = 0.051$, $I^2 = 73.7\%$, $\tau^2 = 0.094$). Regarding maternal smoking, both

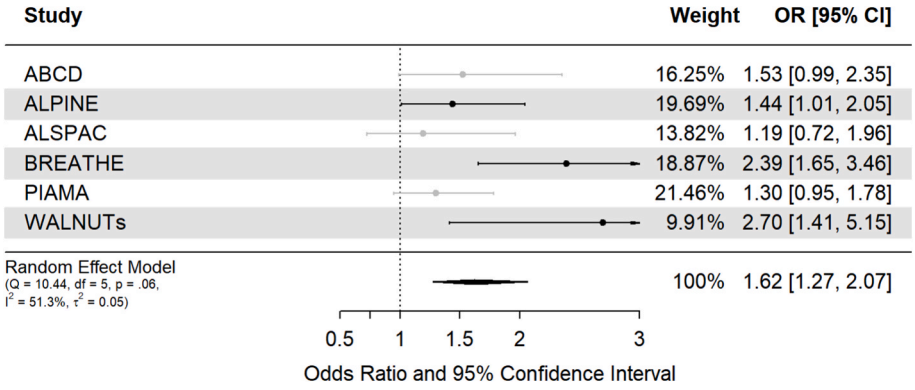

Fig. 1. Meta-analysis of the effect of (low) maternal education on the odds of ADHD diagnosis. Based on the results of the harmonised model, excluding FAIR cohort (black point estimate with confidence interval whiskers represent found associations).

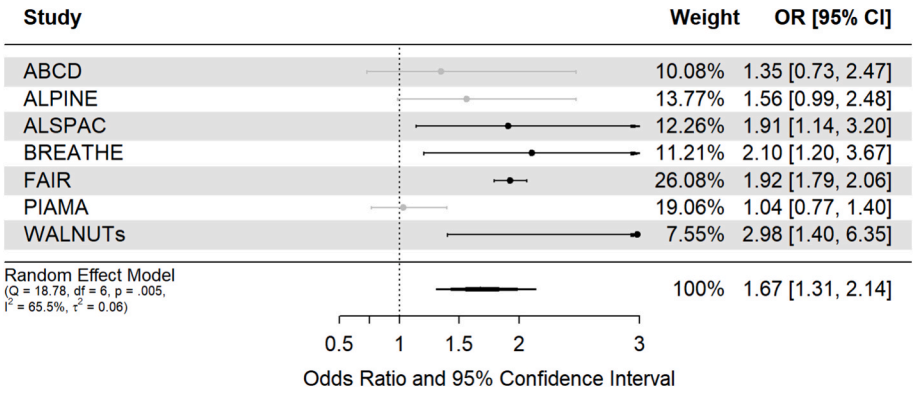

Fig. 2. Meta-analysis of the effect of maternal smoking during pregnancy on the odds of ADHD diagnosis. Based on the results of the harmonised model (black point estimate with confidence interval whiskers represent found associations).

groups indicated increased risk as well. Interestingly, for smoking, the parent reported group showed substantial heterogeneity ($Q = 8.96$, $p = 0.030$, $I^2 = 65.0\%$, $\tau^2 = 0.12$), whereas no heterogeneity was observed in teacher assessment group ($Q = 0.65$, $p = 0.42$, $I^2 = 0\%$, $\tau^2 = 0$). Lastly, the results of *meta*-regression (study size, percent of ADHD cases, median year of the study, and ADHD diagnosis assessment method) was done (**SI** Table S14). None of the tested moderators had meaningful impact on *meta*-analysis heterogeneity.

4. Discussion

Across seven European cohorts totalling over 150,000 children, maternal smoking during pregnancy and low maternal education were associated with increased odds of ADHD diagnosis. Associations between environmental exposures such as air pollution, green space, and road proximity and ADHD were often null or not robust across cohorts.

Previous studies identified a range of factors linked to the likelihood of ADHD. A recent *meta*-analysis, among others, reported that smoking during pregnancy, low parental education, and a single-parent family were associated with ADHD, albeit with weaker supporting evidence (Kim et al., 2020). Adverse family conditions (e.g., maltreatment, incarceration, divorce, and harsh discipline) have additionally been linked to higher ADHD risk (Claussen et al., 2024). Regarding environmental exposures, a literature review highlighted particulate matter, NO_2, and polycyclic aromatic hydrocarbons as potential risk factors, though findings across studies remain mixed (Aghaei et al., 2019). Other chemicals, such as mercury, cadmium, lead, phthalates, bisphenol A, brominated flame retardants, certain pesticides, and per- and polyfluoroalkyl substances, have also been implicated to varying degrees (Moore et al., 2022). Furthermore, road traffic noise has been associated with ADHD symptoms, whereas green space has been identified as a potential protective factor (Hjortebjerg et al., 2016; Yuchi et al., 2022).

Our findings strengthen the evidence base for maternal smoking during pregnancy as a modifiable factor contributing to ADHD. Both basic and fully adjusted models as well as *meta*-analysis across cohorts demonstrated increased odds of ADHD among children of mothers who smoked during pregnancy, and this association persisted in sensitivity analyses. These results align with previous studies and *meta*-analyses highlighting prenatal smoking as one of the most consistent, though not uncontested, factors for ADHD diagnosis. A previous *meta*-analysis reported an OR = 1.60, 95% CI [1.45, 1.76], $I^2 = 79\%$ for the association between prenatal smoking and ADHD (Kim et al., 2020). This is consistent with our pooled estimate albeit with wider confidence intervals reflecting the smaller number of cohorts included in our *meta*-analysis (seven compared to twenty). The biological plausibility is supported by experimental and epidemiological evidence. Nicotine can interfere with foetal brain development (Jansone et al., 2023) and induce DNA methylation (Nonkovic et al., 2024) possibly leading to changes in how genes related to brain development are expressed, increasing the likelihood of ADHD (Huang et al., 2018).

Nonetheless, several studies suggest that the association with smoking may reflect other underlying causal mechanisms rather than smoking itself. ADHD is heritable and within-family studies have reported substantial attenuation of the association with ADHD compared to standard cohort estimates. This suggests that a meaningful portion of the observed effect may reflect familial confounding rather than a direct intrauterine effect of nicotine exposure (Gustavson et al., 2017; Langley et al., 2012). Young adults with ADHD are more likely to smoke and tend to experience more severe withdrawal symptoms than their non-ADHD counterparts (Mitchell et al., 2019), which may increase the likelihood of smoking during pregnancy among affected women. In addition, developmental factors further complicate the interpretation of smoking-related associations. Low birth weight has been linked to an increased risk of ADHD (Crump et al., 2023), while maternal smoking during pregnancy has been linked to low birth weight (Di et al., 2022). These interconnected pathways imply that smoking may influence factors associated with ADHD risk without necessarily being the primary causal driver of ADHD. Together, these findings suggest that both direct biological effects and familial confounding may contribute to the observed association.

The other important result of this study was the association between low maternal education and increased odds of ADHD. Parental socioeconomic disadvantages were associated with impaired children's neurodevelopment (Chin-Lun Hung et al., 2015). Lower parental education was linked to increased odds of ADHD in children (Zhao et al., 2024), even after controlling for genetic predispositions (Torvik et al., 2020). A *meta*-analysis based on six studies reported OR = 1.91, 95% CI [1.20, 3.03], $I^2 = 91\%$ for the association between low maternal education and ADHD (Kim et al., 2020). Our *meta*-analysis, also based on six cohorts, yielded a broadly consistent result (OR = 1.62, 95% CI [1.27, 2.07]), with notably lower heterogeneity ($I^2 = 51.3\%$) and narrower CI. This suggests that the greater methodological consistency in our harmonised analytic framework, including covariate adjustment, and variable definitions, may have helped reduce between-study variability, even though ADHD diagnosis assessment varied between studies.

Education serves as a proxy for socioeconomic position, which covers a wide range of material, behavioural, and psychosocial factors. However, this association may partly reflect passive gene–environment correlation, where parental genetic liability for ADHD contributes to both lower educational attainment and the familial environment shared with offspring (Voronin, et al., 2024; Pingault et al., 2023). Individuals with ADHD often experience educational difficulties, including poorer school performance, grade repetition, and early school dropout (Faraone et al., 2021). As a result, women with an ADHD diagnosis may be less likely to attain higher levels of education due to challenges directly related to their condition. Their children, in turn, have an increased

likelihood of ADHD due to shared genetic predisposition. Therefore, the observed association between low maternal education and ADHD in children may be partly attributable to shared familial genetic liability existing alongside socioeconomic pathways.

E-values indicated that a moderately strong unmeasured confounder would be required to fully nullify the observed associations for smoking during pregnancy and maternal education, suggesting these findings are more robust than those for other examined exposures. Nevertheless, the familial and genetic factors discussed above represent candidates possibly capable of meeting that threshold.

Subgroup *meta*-analyses showed broadly consistent associations for smoking and maternal education across ADHD assessment types (parent/teacher assessment), although heterogeneity varied, indicating that measurement method may have contributed to differences in effect estimates. Meta-regression analyses did not identify study-level characteristics as meaningful sources of heterogeneity. This, however, should be interpreted with caution, given the low number of studies, which arguably limits the ability to detect such patterns (Hempel et al., 2013).

Differences in heterogeneity across ascertainment may reflect that parents and teachers observe the child in different settings and therefore capture partly different behavioural contexts rather than a single homogeneous outcome. Parent reports may better reflect behaviours observed in the home environment, whereas teacher reports may be more influenced by classroom demands, school practices, and referral thresholds (Garcia-Rosales et al., 2021). As a result, variation in effect estimates across cohorts may arise from differences in outcome ascertainment and reporting context rather than from differences in the underlying association itself.

The wide range in ADHD prevalence across cohorts, from 1.6% in ALSPAC to 13.8% in ALPINE, further illustrates this point. This variation is more likely due to differences in how ADHD was operationalized than to cohort eligibility or sampling. The DSM-IV-based parent report used in ALSPAC may have captured more clearly defined cases, whereas the yes/no teacher checklist used in ALPINE may have identified a broader spectrum of symptoms, possibly inflating prevalence (Musullulu, 2025). Taken together, these differences suggest that the cohorts may not have captured a fully homogeneous outcome, which likely contributed to the observed between-study heterogeneity. At the same time, the directional consistency of the two main findings across cohorts and assessment types lends support to their robustness despite this methodological heterogeneity.

Initially, several variables were identified as potentially important by the RF algorithms, including tree cover, NDVI, NO_2, population density and count, and distance to roads. These results may have been influenced by RF-CART's tendency to favour continuous variables, although RF-CI also identified variables such as tree cover and green area size. Nevertheless, associations with ADHD were inconsistent across cohorts and models, and many of the associations were driven by the disproportionately large FAIR cohort. Notably, NO_2 showed protective associations with ADHD in some models, contrary to much of the literature suggesting adverse impacts on neurodevelopment (Zhao et al., 2024). One possible explanation is socioeconomic confounding: families living in more polluted urban centres may have, on average, higher socioeconomic status, which itself was associated with lower odds of ADHD (Faraone et al., 2021). This pattern could be particularly relevant in Sweden (FAIR), where higher-income and more highly educated populations tend to live in areas with higher air pollution levels (Strandell et al., 2024). Nonetheless, these findings did not replicate across all cohorts, and in harmonised models, most associations with air pollution were attenuated or not present. Even so, these results should be interpreted with caution, as they may reflect limitations in exposure assessment rather than true absence of effect. Whether these exposures are genuinely unrelated to ADHD or simply not detectable with employed methods remains uncertain.

Although several variables were selected as informative for ADHD classification by the RF models, many had remarkably small effect sizes or wide CIs in the subsequent logistic regression analyses. This reflects the difference in what each method estimates. In the current study, RF was used as a tool for variable selection. Permutation variable importance captures how much a variable improves classification accuracy in the data, which depends on marginal correlations, nonlinear thresholds, and interactions with other predictors (Breiman, 2001). It does not estimate the association of a variable with the outcome adjusted for the remaining covariates, as logistic regression does. A variable can therefore rank high in RF while showing attenuated or null results in logistic regression when its contribution is highly non-linear, interactive or when it is correlated with other variables. Logistic regression, on the other hand, provides interpretable effect estimates and is more constrained in its assumptions (e.g., linearity, additivity). The discrepancy between RF and logistic regression models does not necessarily imply contradiction but reflects that RF identified exposures carrying predictive information in the data, including through nonlinear and correlated pathways, while logistic regression estimated their independent adjusted associations under more restrictive assumptions.

These findings have implications for research, policy, and public health. Found associations between maternal smoking during pregnancy, low maternal education and ADHD underscore the relevance of established social and behavioural determinants of neurodevelopment. These results highlight the need for interventions that support smoking cessation before and during pregnancy and for broader policies aimed at reducing socioeconomic disadvantage even if part of the observed associations may reflect underlying genetic predisposition. The largely inconsistent associations with environmental exposures suggest that the used variables may be insufficient to capture the complexity of the exposome on neurodevelopment. This underscores the need for future research to integrate improved exposure assessment methods.

This study had several important strengths. It relied on multiple well-established European cohort studies, spanning diverse regions, time periods, and populations. This enhanced the generalizability of findings and allowed for exploration of cohort-specific differences. The use of a harmonised analytic framework across studies ensured consistency in variable definitions, data enrichment, covariate adjustment, and outcome classification. Despite differences in the original cohort designs, we implemented an aligned statistical approach that integrated both machine learning and regression modelling.

Limitations of this study should be acknowledged. ADHD diagnoses were identified using different methods, ranging from register-based clinical diagnoses to parent or teacher reports, which may have introduced misclassification and likely contributed to heterogeneity. Differences in cohort time periods could also have biased diagnosis assessment, as methods for identifying ADHD have advanced over the years (Knyazhansky and Shrot, 2025). The FAIR cohort contributed a disproportionately large sample, which may have driven many of the findings. The use of residential exposures as proxies for individual-level environmental exposures introduces additional potential for misclassification. Alternative methods such as conditional permutation importance or SHAP values may provide more robust handling of correlated predictors in future work. Multiple imputation was performed using five imputed datasets with five iterations per imputation, due to computational constraints related to the large sample size of the FAIR cohort. Finally, unmeasured familial confounding, including parental ADHD, parenting behaviours, and health-seeking behaviours, may have influenced found associations.

5. Conclusion

In this large multi-cohort European study, we evaluated a wide range of prenatal environmental and sociodemographic variables in relation to childhood ADHD. We found that maternal smoking during pregnancy and low maternal education were associated with increased odds of ADHD in children across involved cohorts. Associations with

environmental exposures such as air pollution, green space, or population density were generally small, inconsistent, and often cohort specific. While our results reinforce the evidence of established social and behavioural determinants of ADHD, the interpretation of these associations is limited by the inability to account for genetic and familial factors, which may partly explain some of the observed relationships. In future studies, we recommend a continued focus on the multi-exposure approach, perhaps with greater emphasis on socioeconomic disparities, as these may play an important role in neurodevelopmental vulnerability. From a public health perspective, maternal smoking during pregnancy and socioeconomic disadvantage remain plausible targets for preventive action regardless of the underlying causal mechanisms, given their broader benefits for maternal and child health. Even where part of the association may reflect gene-environment correlation, both factors are amenable to population-level intervention: smoking cessation support during pregnancy and policies that reduce educational and socioeconomic inequality.

6. Data Statement

The datasets generated during and/or analysed during the current study are not publicly available due to the confidential nature of cohort participant data and the restrictions imposed by the data governance agreements under which the data were accessed. Access to the data may be requested through the relevant institutional data access committee, subject to applicable ethical and legal requirements.

CRediT authorship contribution statement

Libor Šulc: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis. **Fan Wu:** Writing – review & editing, Writing – original draft. **Sanne Meijering:** Writing – review & editing, Software, Methodology. **Hendriek Boshuizen:** Writing – review & editing, Supervision, Software, Project administration, Methodology, Conceptualization. **Angel M. Dzhambov:** Writing – review & editing, Software, Methodology, Investigation, Formal analysis. **Peter Lercher:** Writing – review & editing. **Carmen Peuters:** Writing – review & editing, Investigation, Formal analysis. **Boris Cheval:** Writing – review & editing, Methodology. **Nuria Botella:** Writing – review & editing, Investigation, Formal analysis. **Payam Dadvand:** Writing – review & editing, Resources, Methodology, Data curation. **Jordi Júlvez:** Writing – review & editing. **Jolanda M.A. Boer:** Writing – review & editing, Investigation, Formal analysis. **Christoph Giehl:** Writing – review & editing, Software, Methodology, Investigation, Formal analysis. **Jan Spilski:** Writing – review & editing, Conceptualization. **Yingxin Chen:** Writing – review & editing, Investigation, Formal analysis. **Anna L. Hansell:** Writing – review & editing, Conceptualization. **Rik Bogers:** Writing – review & editing, Conceptualization. **Tanja Vrijkotte:** Writing – review & editing, Conceptualization. **Gabriele Bolte:** Writing – review & editing, Resources, Data curation. **Achilleas Psyllidis:** Writing – review & editing, Resources, Data curation. **Karl Lundin Remnélius:** Writing – review & editing, Investigation. **Sven Bölte:** Writing – review & editing, Supervision, Methodology. **Elise van Kempen:** Writing – review & editing, Supervision, Project administration. **Irene van Kamp:** Writing – review & editing, Project administration, Funding acquisition. **Jenny Selander:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Funding acquisition, Conceptualization.

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 874724. Equal-Life is a member of the European Human Exposome Network. ALH acknowledges support from the UK National Institute for Health and Care Research (NIHR) Health Protection Research Unit (HPRU) in Chemical Threats and Hazards, a partnership between the UK Health Security Agency (UKHSA), the Health and Safety Executive (HSE) and the University of Leicester [grant number NIHR207293]; and from the NIHR Leicester Biomedical Research Centre (BRC) [grant number NIHR203327]. The views expressed are those of the author(s) and not necessarily those of the NIHR, UKHSA, HSE, or University of Leicester. The PIAMA study was supported by The Netherlands Organization for Health Research and Development; The Netherlands Organization for Scientific Research; The Long Foundation Netherlands (project number 4.1.14.001); The Netherlands Ministry of Spatial Planning, Housing, and the Environment; The Netherlands Ministry of Health, Welfare, and Sport; and the National Institute for Public Health and the Environment. The WALNUTs project was supported by the Instituto de Salud Carlos III through the projects CP14/00108 and PI16/00261 (co-funded by European Union Regional Development Fund 'A way to make Europe'. The BREATHE project was funded by the European Research Council under the ERC Grant Agreement number 268,479 – the BREATHE project. ISGlobal is a member of the CERCA Programme, Generalitat de Catalunya. Austrian Ministry of Science and Transportation funded the framework of the Environmental Health Impact Assessment (EHIA), the government of the Tyrol region provided GIS data, and informational support was received from the BEG (Brenner Eisenbahn Gesellschaft). The BBT survey got support from the BBT company within a legally required EHIA through EU-support.

Declaration of competing interest

The authors declare that they have no conflict of financial or personal interests that could influence the work reported in this paper.

Acknowledgement

The authors gratefully acknowledge the contribution of all Equal-Life consortium members and the members of the advisory committee (www.equal-life.eu). We are grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The authors thank the participating mothers and their children of the ABCD-study, and all other persons and institutions who contributed to the ABCD study: obstetric care providers, primary schools, students, and youth healthcare centers in Amsterdam (the Netherlands).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2026.110415.

Data availability

The data that has been used is confidential.

References

Aghaei, M., Janjani, H., Yousefian, F., Jamal, A., Yunesian, M., 2019. Association between ambient gaseous and particulate air pollutants and attention deficit hyperactivity disorder (ADHD) in children; a systematic review. Environ. Res. 173, 135–156. https://doi.org/10.1016/j.envres.2019.03.030.

Altmann, A., Toloşi, L., Sander, O., Lengauer, T., 2010. Permutation importance: a corrected feature importance measure. Bioinformatics 26 (10), 1340–1347. https://doi.org/10.1093/bioinformatics/btq134.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32. https://doi.org/10.1023/A:1010933404324.

van Buuren, S., Groothuis-Oudshoorn, K., 2011. **mice** : multivariate imputation by chained equations in *R*. J. Stat. Softw. 45 (3). https://doi.org/10.18637/jss.v045.i03.

Carlsson, T., Molander, F., Taylor, M.J., Jonsson, U., Bölte, S., 2021. Early environmental risk factors for neurodevelopmental disorders – a systematic review of twin and

sibling studies. Dev. Psychopathol. 33 (4), 1448–1495. https://doi.org/10.1017/S0954579420000620.
Chin-Lun Hung, G., Hahn, J., Alamiri, B., Buka, S.L., Goldstein, J.M., Laird, N., et al., 2015. Socioeconomic disadvantage and neural development from infancy through early childhood. Int. J. Epidemiol. 44 (6), 1889–1899. https://doi.org/10.1093/ije/dyv303.
Claussen, A.H., Holbrook, J.R., Hutchins, H.J., Robinson, L.R., Bloomfield, J., Meng, L., et al., 2024. All in the family? a systematic review and meta-analysis of parenting and family environment as risk factors for attention-deficit/hyperactivity disorder (ADHD) in children. Prev. Sci. 25 (S2), 249–271. https://doi.org/10.1007/s11121-022-01358-4.
Coghill, D., Banaschewski, T., Cortese, S., Asherson, P., Brandeis, D., Buitelaar, J., et al., 2023. The management of ADHD in children and adolescents: bringing evidence to the clinic: perspective from the European ADHD Guidelines Group (EAGG). Eur. Child Adolesc. Psychiatry 32 (8), 1337–1361. https://doi.org/10.1007/s00787-021-01871-x.
Crump, C., Sundquist, J., Sundquist, K., 2023. Preterm or early term birth and risk of attention-deficit/hyperactivity disorder: a national cohort and co-sibling study. Ann. Epidemiol. 86. https://doi.org/10.1016/j.annepidem.2023.08.007.
Demontis, D., Walters, G.B., Athanasiadis, G., Walters, R., Therrien, K., Nielsen, T.T., et al., 2023. Genome-wide analyses of ADHD identify 27 risk loci, refine the genetic architecture and implicate several cognitive domains. Nat. Genet. 55 (2), 198–208. https://doi.org/10.1038/s41588-022-01285-8.
Di, H.K., Gan, Y., Lu, K., Wang, C., Zhu, Y., Meng, X., et al., 2022. Maternal smoking status during pregnancy and low birth weight in offspring: systematic review and meta-analysis of 55 cohort studies published from 1986 to 2020. World J. Pediatr. 18 (3), 176–185. https://doi.org/10.1007/s12519-021-00501-5.
Dzhambov, A.M., Markevych, I., Lercher, P., 2019. Associations of residential greenness, traffic noise, and air pollution with birth outcomes across Alpine areas. Sci. Total Environ. 678, 399–408. https://doi.org/10.1016/j.scitotenv.2019.05.019.
Erzin, G., Guloksuz, S., 2021. The exposome paradigm to understand the environmental origins of mental disorders. Alpha Psychiatry. 22 (4), 171–176. https://doi.org/10.5152/alphapsychiatry.2021.21307.
Faraone, S.V., Banaschewski, T., Coghill, D., Zheng, Y., Biederman, J., Bellgrove, M.A., et al., 2021. The World Federation of ADHD international consensus statement: 208 evidence-based conclusions about the disorder. Neurosci. Biobehav. Rev. 128, 789–818. https://doi.org/10.1016/j.neubiorev.2021.01.022.
Faraone, S.V., Bellgrove, M.A., Brikell, I., Cortese, S., Hartman, C.A., Hollis, C., et al., 2024. Attention-deficit/hyperactivity disorder. Nat. Rev. Dis. Primers 10 (1), 11. https://doi.org/10.1038/s41572-024-00495-0.
Forns, J., Dadvand, P., Foraster, M., Alvarez-Pedrerol, M., Rivas, I., López-Vicente, M., et al., 2016. Traffic-related air pollution, noise at school, and behavioral problems in barcelona schoolchildren: a cross-sectional study. Environ. Health Perspect. 124 (4), 529–535. https://doi.org/10.1289/ehp.1409449.
Fox, J., Monette, G., 1992. Generalized collinearity diagnostics. J. Am. Stat. Assoc. 87 (417), 178–183. https://doi.org/10.1080/01621459.1992.10475190.
Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G., et al., 2013. Cohort profile: the avon longitudinal study of parents and children: ALSPAC mothers cohort. Int. J. Epidemiol. 42 (1), 97–110. https://doi.org/10.1093/ije/dys066.
Garcia-Rosales, A., Vitoratou, S., Faraone, S.V., Rudaizky, D., Banaschewski, T., Asherson, P., et al., 2021. Differential utility of teacher and parent–teacher combined information in the assessment of attention deficit/hyperactivity disorder symptoms. Eur. Child Adolesc. Psychiatry 30 (1), 143–153. https://doi.org/10.1007/s00787-020-01509-4.
Gustavson, K., Ystrom, E., Stoltenberg, C., Susser, E., Surén, P., Magnus, P., et al., 2017. Smoking in pregnancy and child ADHD. Pediatrics 139 (2), e20162509. https://doi.org/10.1542/peds.2016-2509.
Gutierrez-Ortiz, C., Hossain, B., Dessenne, C., Aguayo, G.A., Ruiz-Castell, M., 2025. Role of the exposome in mental disorders: a scoping review protocol. BMJ Open 15 (8), e101575. https://doi.org/10.1136/bmjopen-2025-101575.
Hempel, S., Miles, J.N., Booth, M.J., Wang, Z., Morton, S.C., Shekelle, P.G., 2013. Risk of bias: a simulation study of power to detect study-level moderator effects in meta-analysis. Syst. Rev. 2 (1), 107. https://doi.org/10.1186/2046-4053-2-107.
Hjortebjerg, D., Andersen, A.M.N., Christensen, J.S., Ketzel, M., Raaschou-Nielsen, O., Sunyer, J., et al., 2016. Exposure to road traffic noise and behavioral problems in 7-year-old children: a cohort study. Environ. Health Perspect. 124 (2), 228–234. https://doi.org/10.1289/ehp.1409430.
Hothorn, T., Hornik, K., Zeileis, A., 2006. Unbiased recursive partitioning: a conditional inference framework. J. Comput. Graph. Stat. 15 (3), 651–674. https://doi.org/10.1198/106186006X133933.
Huang L, Wang Y, Zhang L, Zheng Z, Zhu T, Qu Y, et al. Maternal Smoking and Attention-Deficit/Hyperactivity Disorder in Offspring: A Meta-analysis. Pediatrics. 2018 Jan 1; 141(1). doi:10.1542/peds.2017-2465.
Jansone, K., Eichler, A., Fasching, P.A., Kornhuber, J., Kaiser, A., Millenet, S., et al., 2023. Association of maternal smoking during pregnancy with neurophysiological and ADHD-related outcomes in school-aged children. Int. J. Environ. Res. Public Health 20 (6), 4716. https://doi.org/10.3390/ijerph20064716.
Julvez, J., Gignac, F., Fernández-Barrés, S., Romaguera, D., Sala-Vila, A., Ranzani, O.T., et al., 2021. Walnuts, long-chain polyunsaturated fatty acids, and adolescent brain development: protocol for the walnuts smart snack dietary intervention trial. Front. Pediatr. 8, 9. https://doi.org/10.3389/fped.2021.593847.
Kim, J.H., Kim, J.Y., Lee, J., Jeong, G.H., Lee, E., Lee, S., et al., 2020. Environmental risk factors, protective factors, and peripheral biomarkers for ADHD: an umbrella review. Lancet Psychiatry 7 (11), 955–970. https://doi.org/10.1016/S2215-0366(20)30312-6.
Knyazhansky, M., Shrot, T., 2025. ADHD diagnostic tools across ages: traditional and digital approaches. Front. Psych. 14, 16. https://doi.org/10.3389/fpsyt.2025.1668070.
Lampa, E., Lind, L., Lind, P.M., Bornefalk-Hermansson, A., 2014. The identification of complex interactions in epidemiology and toxicology: a simulation study of boosted regression trees. Environ. Health 13 (1), 57. https://doi.org/10.1186/1476-069X-13-57.
Langley, K., Heron, J., Smith, G.D., Thapar, A., 2012. Maternal and paternal smoking during pregnancy and risk of ADHD symptoms in offspring: testing for intrauterine effects. Am. J. Epidemiol. 176 (3), 261–268. https://doi.org/10.1093/aje/kwr510.
Little R, Rubin D. Statistical Analysis with Missing Data, Third Edition [Internet]. Wiley; 2019. (Wiley Series in Probability and Statistics). Available from: https://onlinelibrary.wiley.com/doi/book/10.1002/9781119482260 doi:10.1002/9781119482260.
Madley-Dowd, P., Hughes, R., Tilling, K., Heron, J., 2019. The proportion of missing data should not be used to guide decisions on multiple imputation. J. Clin. Epidemiol. 110, 63–73. https://doi.org/10.1016/j.jclinepi.2019.02.016.
Meijering S. Random Forest code [Internet]. 2026. Available from: https://github.com/Sanne-Meijering/Equal-Life-Random-Forest.
Mikeš, O., Sáňka, O., Rafajová, A., Vlaanderen, J., Chen, J., Hoek, G., et al., 2023. Development of historic monthly land use regression models of SO2, NOx and suspended particulate matter for birth cohort ELSPAC. Atmos. Environ. 301, 119688. https://doi.org/10.1016/j.atmosenv.2023.119688.
Mitchell, J.T., Howard, A.L., Belendiuk, K.A., Kennedy, T.M., Stehli, A., Swanson, J.M., et al., 2019. Cigarette smoking progression among young adults diagnosed with ADHD in childhood: a 16-year longitudinal study of children with and without ADHD. Nicotine Tob. Res. 21 (5), 638–647. https://doi.org/10.1093/ntr/nty045.
Moore, S., Paalanen, L., Melymuk, L., Katsonouri, A., Kolossa-Gehring, M., Tolonen, H., 2022. The association between ADHD and environmental chemicals—a scoping review. Int. J. Environ. Res. Public Health 19 (5), 2849. https://doi.org/10.3390/ijerph19052849.
Musullulu, H., 2025. Evaluating attention deficit and hyperactivity disorder (ADHD): a review of current methods and issues. Front. Psychol. 24, 16. https://doi.org/10.3389/fpsyg.2025.1466088.
Nonkovic, N., Marceau, K., McGeary, J.E., Ramos, A.M., Palmer, R.H.C., Heath, A.C., et al., 2024. Maternal smoking during pregnancy is associated with DNA methylation in early adolescence: a sibling comparison design. Dev. Psychol. 60 (9), 1639–1654. https://doi.org/10.1037/dev0001747.
Oudin, A., Frondelius, K., Haglund, N., Källén, K., Forsberg, B., Gustafsson, P., et al., 2019. Prenatal exposure to air pollution as a potential risk factor for autism and ADHD. Environ. Int. 133, 105149. https://doi.org/10.1016/j.envint.2019.105149.
Pingault, J.B., Barkhuizen, W., Wang, B., Hannigan, L.J., Eilertsen, E.M., Corfield, E., et al., 2023. Genetic nurture versus genetic transmission of risk for ADHD traits in the Norwegian mother, father and child cohort study. Mol. Psychiatry 28 (4), 1731–1738. https://doi.org/10.1038/s41380-022-01863-6.
Pratt, G.C., Parson, K., Shinoda, N., Lindgren, P., Dunlap, S., Yawn, B., et al., 2014. Quantifying traffic exposure. J. Eposure Sci. Environ. Epidemiol. 24 (3), 290–296. https://doi.org/10.1038/jes.2013.51.
Rengasamy, D., Rothwell, B.C., Figueredo, G.P., 2021. Towards a more reliable interpretation of machine learning outputs for safety-critical systems using feature importance fusion. Appl. Sci. 11 (24), 11854. https://doi.org/10.3390/app112411854.
Rocco, I., Corso, B., Bonati, M., Minicuci, N., 2021. Time of onset and/or diagnosis of ADHD in European children: a systematic review. BMC Psychiatry 21 (1), 575. https://doi.org/10.1186/s12888-021-03547-x.
Sciberras, E., Mulraney, M., Silva, D., Coghill, D., 2017. Prenatal risk factors and the etiology of ADHD—review of existing evidence. Curr. Psychiatry Rep. 19 (1), 1. https://doi.org/10.1007/s11920-017-0753-2.
Sharma, A., Couture, J., 2014. A review of the pathophysiology, etiology, and treatment of Attention-Deficit Hyperactivity Disorder (ADHD). Ann. Pharmacother. 48 (2), 209–225. https://doi.org/10.1177/1060028013510699.
Skröder, H., Pettersson, H., Norlén, F., Gustavsson, P., Rylander, L., Albin, M., et al., 2021. Occupational exposure to whole body vibrations and birth outcomes – a nationwide cohort study of Swedish women. Sci. Total Environ. 751, 141476. https://doi.org/10.1016/j.scitotenv.2020.141476.
Sonuga-Barke, E.J.S., Becker, S.P., Bölte, S., Castellanos, F.X., Franke, B., Newcorn, J.H., et al., 2023. Annual research review: perspectives on progress in <scp>ADHD</scp> science – from characterization to cause. J. Child Psychol. Psychiatry 64 (4), 506–532. https://doi.org/10.1111/jcpp.13696.
Stekhoven, D.J., Bühlmann, P., 2012. MissForest—non-parametric missing value imputation for mixed-type data. Bioinformatics 28 (1), 112–118. https://doi.org/10.1093/bioinformatics/btr597.
Strandell, A., Nurmio, K., Seifert-Dähnn, I., Skumlien Furuseth, I., Wolf, R., Åström, S., et al., 2024. Air pollution inequality and its temporal trends in Nordic countries. Int. J. Environ. Stud. 81 (5), 2203–2226. https://doi.org/10.1080/00207233.2024.2403952.
Strobl, C., Boulesteix, A.L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. BMC Bioinf. 9 (1), 307. https://doi.org/10.1186/1471-2105-9-307.
Strobl, C., Malley, J., Tutz, G., 2009. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychol. Methods 14 (4), 323–348. https://doi.org/10.1037/a0016973.
Šulc L. Logistic regression code [Internet]. 2026. Available from: https://github.com/Libor-Sulc/Equal-Life-Logistic-Regression.

Torvik, F.A., Eilertsen, E.M., McAdams, T.A., Gustavson, K., Zachrisson, H.D., Brandlistuen, R., et al., 2020. Mechanisms linking parental educational attainment with child ADHD, depression, and academic problems: a study of extended families in the Norwegian mother, father and child cohort study. J. Child Psychol. Psychiatry 61 (9), 1009–1018. https://doi.org/10.1111/jcpp.13197.

Ullmann, R.K., Sleator, E.K., Sprague, R.L., 1984. A new rating scale for diagnosis and monitoring of ADD children. Psychopharmacol. Bull. 160–4.

VanderWeele, T.J., Ding, P., 2017. Sensitivity analysis in observational research: introducing the E-Value. Ann. Intern. Med. 167 (4), 268–274. https://doi.org/10.7326/M16-2607.

van Eijsden, M., Vrijkotte, T.G., Gemke, R.J., van der Wal, M.F., 2011. Cohort profile: the Amsterdam Born Children and their Development (ABCD) study. Int. J. Epidemiol. 40 (5), 1176–1186. https://doi.org/10.1093/ije/dyq128.

van Kamp, I., Persson Waye, K., Kanninen, K., Gulliver, J., Bozzon, A., Psyllidis, A., et al., 2022. Early environmental quality and life-course mental health effects: the equal-life project. Environ. Epidemiol. 6 (1), e183.

Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. J. Stat. Softw. 36 (3). https://doi.org/10.18637/jss.v036.i03.

Vittinghoff, E., McCulloch, C.E., 2007. Relaxing the rule of ten events per variable in logistic and cox regression. Am. J. Epidemiol. 165 (6), 710–718. https://doi.org/10.1093/aje/kwk052.

Voronin I, Ouellet-Morin I, Petitclerc A, Morneau-Vaillancourt G, Brendgen M, Dione G, et al. Intergenerational transmission of genetic risk for hyperactivity and inattention. Direct genetic transmission or genetic nurture? JCPP Advances. 2024 Jun 4;4(2). doi:10.1002/jcv2.12222.

Who, 2017. Urban green spaces: a brief for action. WHO.

Wijga, A.H., Kerkhof, M., Gehring, U., de Jongste, J.C., Postma, D.S., Aalberse, R.C., et al., 2014. Cohort profile: the prevention and incidence of asthma and mite allergy (PIAMA) birth cohort. Int. J. Epidemiol. 43 (2), 527–535. https://doi.org/10.1093/ije/dys231.

Wild, C.P., 2005. Complementing the genome with an "Exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiol. Biomarkers Prev. 14 (8), 1847–1850. https://doi.org/10.1158/1055-9965.EPI-05-0456.

Willcutt, E.G., 2012. The prevalence of DSM-IV attention-deficit/hyperactivity disorder: a meta-analytic review. Neurotherapeutics 9 (3), 490–499. https://doi.org/10.1007/s13311-012-0135-8.

Williams, T., Barclay, I., Bevan-Jones, R., Livingston, L.A., Agha, S.S., Ford, T., et al., 2025. Reflections on the manifestation of attention-deficit hyperactivity disorder in girls from young adults with lived experiences: a qualitative study. Br. J. Psychiatry 227 (5), 775–782. https://doi.org/10.1192/bjp.2025.10376.

Wright MN, Ziegler A. ranger : A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. J Stat Softw. 2017;77(1). doi:10.18637/jss.v077.i01.

Yuchi, W., Brauer, M., Czekajlo, A., Davies, H.W., Davis, Z., Guhn, M., et al., 2022. Neighborhood environmental exposures and incidence of attention deficit/hyperactivity disorder: a population-based cohort study. Environ. Int. 161, 107120. https://doi.org/10.1016/j.envint.2022.107120.

Zhao, Y.K., Li, M., Shi, T.T., Feng, M.M., Hu, L.L., 2024. Association of premature birth and maternal education level on attention deficit hyperactivity disorder in children: a meta-analysis. World J Psychiatry. 14 (12), 1956–1970. https://doi.org/10.5498/wjp.v14.i12.1956.

Zhao, J., He, T., Wang, F., Liu, W., 2024. Association of prenatal and postnatal exposure to air pollution with clinically diagnosed attention deficit hyperactivity disorder: a systematic review. Front. Public Health 24, 12. https://doi.org/10.3389/fpubh.2024.1396251.