Research Article | Open Access

Journal of Environmental Exposure Assessment

Liu *et al. J. Environ. Expo. Assess.* **2026**, *5*, 31 **DOI:**10.20517/jeea.2026.33

Sex-specific effects of prenatal metal exposure on child anogenital distance: an advanced computational analysis

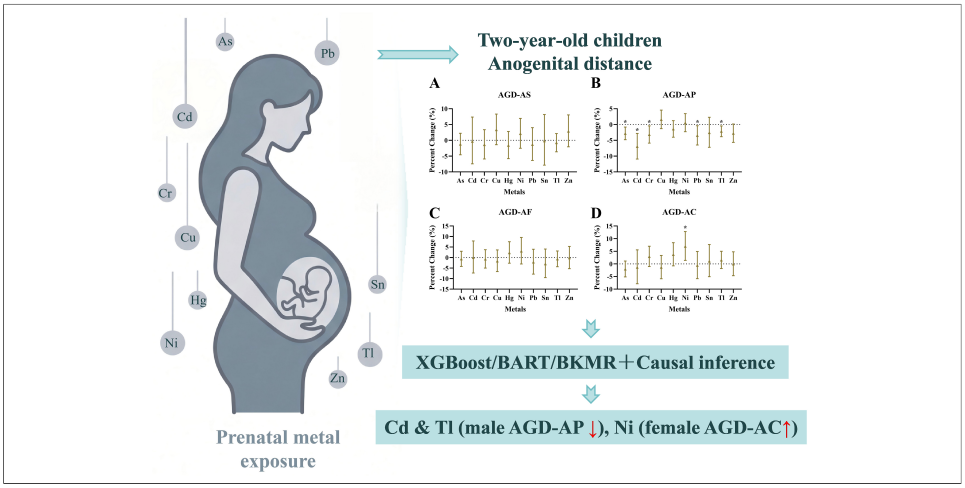
Jiangtao Liu[1], Shulan Li[2], Yuanyuan Li[3], Wei Xia[3], Hongxiu Liu[1], Shunqing Xu[1]

Keywords:
Heavy metals, anogenital distance, prenatal exposure, machine learning, sex-specific effects

Citation: Liu, J.; Li, S.; Li, Y.; Xia, W.; Liu, H.; Xu, S. Sex-specific effects of prenatal metal exposure on child anogenital distance: an advanced computational analysis. *J. Environ. Expo. Assess.* **2026**, *5*, 31. https://dx.doi.org/10.20517/jeea.2026.33

Received: 11 Jun 2026
First Decision: 23 Jul 2026
Revised: 30 Aug 2026
Accepted: 8 Sep 2026
Published: 23 Sep 2026

Academic Editor:
Stuart Harrad
Copy Editor:
Pei-Yun Wang
Production Editor:
Pei-Yun Wang



Abstract

Whether prenatal metal mixtures exert sex-dependent effects on anogenital distance (AGD) - a biomarker of prenatal androgen action - remains unknown. To address this gap, we conducted a prospective cohort study of 248 mother-infant pairs, measuring 10 metals in first-trimester maternal plasma and AGD at age 2 years. To evaluate the associations and identify the most important metal contributors, we combined single-metal linear regression with three complementary advanced computational approaches (extreme gradient boosting, Bayesian additive regression trees, Bayesian kernel machine regression). We further applied propensity score stratification and overlap weighting to strengthen causal inference. The analyses revealed sex-specific patterns: in males, cadmium (Cd), thallium (Tl), and arsenic (As) were inversely associated with anopenile distance (AGD-AP), with Cd showing the strongest effect (-6.96%, P_{FDR} = 0.013), followed by As (-2.83%, P_{FDR} = 0.033) and Tl (-2.16%, P_{FDR} = 0.040). In females, nickel (Ni) was suggestively

[1]School of Environmental Science and Engineering, Hainan University, Haikou 570228, Hainan, China.
[2]Wuhan Children's Hospital (Wuhan Maternal and Child Healthcare Hospital), Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430016, Hubei, China.
[3]Key Laboratory of Environment and Health, Ministry of Education & Ministry of Environmental Protection, and State Key Laboratory of Environmental Health, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei, China.

Correspondence to: Prof. Shunqing Xu, Prof. Hongxiu Liu, School of Environmental Science and Engineering, Hainan University, Haikou 570228, Hainan, China. E-mail: xus@hainanu.edu.cn; lhx@hainanu.edu.cn

positively associated with anoclitoral distance (AGD-AC) (6.94%, P = 0.014), though not significant after false discovery rate correction. Machine learning identified Tl and Cd as the top predictors in males, and As in females. These associations were supported by propensity score-based sensitivity analyses, particularly for male Tl and Cd. No significant associations were found for AGD-AS/AGD-AF in either sex. These findings demonstrate that prenatal metal exposure exerts sex-specific opposing effects on AGD, with Cd and Tl in males showing the most robust inverse associations. Our results highlight the value of integrating machine learning with traditional epidemiological approaches and underscore the importance of sex-stratified analyses in mixture exposure analysis.

INTRODUCTION

Heavy metals are ubiquitous environmental contaminants. Industrial emissions, agricultural practices, vehicular traffic, and household sources continually release metals such as chromium (Cr), nickel (Ni), copper (Cu), zinc (Zn), arsenic (As), cadmium (Cd), tin (Sn), mercury (Hg), thallium (Tl), and lead (Pb) into air, water, soil, and food chains[1-3]. Once released, these elements persist in the environment and accumulate in biological systems, where they have been linked to a range of chronic and acute health conditions[4,5].

Pregnant women and developing fetuses are particularly vulnerable to metal exposure, as metals can cross the placental barrier and interfere with critical developmental processes[6,7]. For instance, Cd, Pb, and As are all associated with adverse fetal outcomes, including reduced birth weight, increased risk of preterm birth, and impaired fetal growth[8]. The first trimester - when organogenesis occurs and the fetal organs and systems are forming and maturing rapidly - is a period of heightened vulnerability to environmental insults, including heavy metal exposure[9-12]. During this critical phase, the fetal body lacks fully developed detoxification mechanisms, making it more susceptible to the adverse effects of metal accumulation[13,14]. Consequently, early pregnancy is widely recognized as a critical window for metal-induced developmental toxicity, where even low-level metal exposure may lead to long-term, irreversible impairments in fetal development.

Anogenital distance (AGD), measured from the anus to the genital tubercle, is a sexually dimorphic trait sensitive to prenatal endocrine disruption, strongly influenced by embryonic androgen action[15-18]. AGD is increasingly used as a noninvasive biomarker in human epidemiology: shorter male AGD correlates with lower fertility, reduced semen quality, lower testosterone levels in adulthood, and birth defects like hypospadias[19,20]; in females, longer AGD is associated with polycystic ovary syndrome[21,22], whereas shorter AGD correlates with endometriosis[21], primary ovarian insufficiency[23], and altered female sexual function[24]. Importantly, the literature frequently reports that AGD is affected by environmental pollutants and can reflect prenatal hormone perturbations.

Existing epidemiological evidence on metals and reproductive development remains limited[25,26]. One key study provided preliminary evidence: Huang *et al.* found that prenatal serum Pb and Cr were inversely associated with AGD in 2- to 3-year-old boys, while antimony and strontium showed positive associations[26]; Celik *et al.* found that neonatal urinary Hg, maternal urinary manganese (Mn), and cord blood Mn were negatively correlated with penile length, and neonatal urinary Hg was also inversely associated with AGD[25]. Despite these initial reports, current literature has largely concentrated on adverse birth outcomes such as preterm birth and fetal growth restriction, and research on AGD as an endocrine-sensitive endpoint remains sparse. Collectively, these observations underscore the need to further explore the relationships between prenatal metal exposure and AGD to address key knowledge gaps in this field.

To address these gaps, we conducted a prospective cohort study of 248 mother-infant pairs. First-trimester maternal plasma was analyzed for 10 metals (As, Cd, Cr, Cu, Hg, Ni, Pb, Sn, Tl, Zn). At age 2 years, we assessed the sexually dimorphic, endocrine-sensitive AGD. Single-metal linear regression was first performed

to identify metals significantly associated with AGD. We employed three advanced computational approaches - extreme gradient boosting (XGBoost)[27], Bayesian additive regression trees (BART)[28], and Bayesian kernel machine regression (BKMR)[29,30] - to evaluate variable importance and mixture effects. To assess the robustness of the single-metal associations to residual confounding, we also performed propensity score stratification and overlap weighting as causal sensitivity analyses. Our primary aim was to compare mixture associations with individual metal patterns and thereby contextualize metal-induced endocrine disruption.

EXPERIMENTAL

Study population

The study participants were derived from a prospective birth cohort established at the Wuhan Women's and Children's Health Care Center[31-33]. Eligible participants met the following criteria: (1) singleton pregnancy with gestational age less than 16 weeks at enrollment; (2) permanent residence in Wuhan; (3) provided a maternal plasma sample during the first trimester; (4) delivered a live-born infant without congenital anomalies at the study hospital; and (5) completed a follow-up visit at two years of age, during which AGD were measured. The present analysis included 248 mother-infant pairs with available data on first-trimester plasma, maternal and child characteristics, and AGD measurements at 2 years of age, with children born between August 2018 and May 2020.

The study protocol was reviewed and approved by the Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology ([2012] (14#)) and the Institutional Review Board of Wuhan Medical & Healthcare Center for Women and Children (No. 2010009 and 2016003) prior to the initiation of maternal recruitment. At enrollment, written informed consent was obtained from all participating mothers. This consent covered telephone follow-ups, questionnaire surveys, and the collection of biological samples during routine prenatal care, delivery hospitalization, and subsequent child development follow-up. During the follow-up phase, children and their parents were contacted by telephone and invited to attend physical examinations and questionnaire surveys. For the 2-year-old child assessments (including AGD measurements and questionnaire administrations), written informed consent was obtained from the parents or legal guardians of all participating children prior to any study procedures. All parents and guardians were fully informed that participation was voluntary and that they could withdraw their child from the study at any time without any negative consequences.

Plasma metal concentration analysis

Maternal venous blood was collected during the first trimester, with a median gestational age of 13 weeks. Plasma samples were analyzed for 10 metals - Cr, Ni, Cu, Zn, As, Cd, Sn, Hg, Tl, and Pb - using inductively coupled plasma mass spectrometry (ICP-MS, Agilent 7900 Series, Agilent Technologies, USA). The analytical procedure followed established methods[31,34]. In detail, a 90-μL aliquot of plasma was mixed with 90 μL of nitric acid and allowed to react for 24 h. The mixture was subsequently diluted with 1.8 mL of deionized water and centrifuged. The supernatant was then introduced into the ICP-MS system for quantification. Quality assurance included blank samples and certified reference materials (ClinChek® Plasma Control for Trace Elements, Level I and II) in each batch. To correct for matrix effects and instrumental drift, rhodium (1 μg/L) was added as an internal standard to all samples prior to measurement (Agilent Technologies, 2023). We generated a multi-point calibration curve using certified standard solutions. Limits of detection (LODs) for the 10 metals are summarized in Supplementary Table 1. For concentrations below the LOD, values were replaced with LOD/2 for statistical analyses.

Assessment of AGD

AGD was measured in 2-year-old children following standardized procedures[16,17,20,35]. Children were placed in a supine position on a flat examination table with the lower body exposed. The hips were positioned at the edge of the table, and the thighs were flexed to approximately 60-90 degrees relative to the trunk, ensuring a consistent posture. The center of the anus was marked with a disposable skin marker as a reference point. Two AGD measures were obtained using a digital caliper (Deli, China; precision 0.1 mm): two AGDs were measured: (1) anoscrotal distance (AGD-AS) in males/anofourchette distance (AGD-AF) in females, defined as the distance from the anus to the posterior base of the scrotum (males) or to the posterior fourchette (females); and (2) anopenile distance (AGD-AP) in males/anoclitoral distance (AGD-AC) in females, defined as the distance from the anus to the anterior base of the penis (males) or to the clitoral glans (females). Each measurement was performed twice consecutively. To keep the child calm and maintain the correct posture, the examiner used gentle verbal encouragement or other age-appropriate distraction techniques as needed, with assistance from the parent if required. All measurements were carried out by trained examiners who were blinded to the metal exposure status of the child.

Data collection on maternal and child characteristics

Demographic and clinical information was obtained through two approaches: structured interviews conducted by trained research nurses and abstraction from electronic medical records. Maternal factors included age (categorized as ≤ 24, 25-34, and ≥ 35 years), pre-pregnancy body mass index (PBMI, underweight < 18.5, normal 18.5-23.9, overweight 24-27.9, obese ≥ 28), education level (junior high or below, high school, college or above), household income (low: < 50,000 RMB; middle: 50,000-199,999 RMB; high: ≥ 200,000 RMB), parity (primiparous, second, third or more), and gestational age at delivery (preterm < 37 weeks, term ≥ 37 weeks). Child factors included sex (male, female) and age at AGD measurement (continuous, in years), and birth weight (low birth weight: < 2,500 g; normal birth weight: 2,500-3,999 g; high birth weight: ≥ 4,000 g.

Statistical analyses

Descriptive statistics

Descriptive statistics were computed to summarize the study variables. Baseline demographic, maternal, and infant characteristics were described by infant sex. Continuous variables were summarized using medians with interquartile ranges (IQRs), and categorical variables as frequencies and percentages. Prenatal concentrations of the 10 metals were summarized by infant sex using selected percentiles (2.5th, 25th, 50th, 75th, and 97.5th), with detection rates also reported. AGD measures were similarly summarized using the same percentiles. Differences between sexes were assessed using Wilcoxon rank-sum tests for continuous measures and chi-square or Fisher's exact tests for categorical variables. Spearman rank correlation coefficients were computed to evaluate pairwise associations among the 10 metals and between AGD outcomes (i.e., AGD-AP *vs.* AGD-AS in males; AGD-AC *vs.* AGD-AF in females). Scatter plots for the AGD–AGD correlations are presented to visualize these relationships.

Linear regression models

Linear regression models were constructed to evaluate associations between each of the 10 metals (Cr, Ni, Cu, Zn, As, Cd, Sn, Hg, Tl, Pb) and the outcomes (AGD-AS, AGD-AF, AGD-AP, AGD-AC). To address right-skewed distributions, metal concentrations were $\log_2$-transformed (with 0.001 added to avoid zeros). The $\log_2$-transformed values were then linearly transformed to Z-scores (mean = 0, standard deviation = 1), a process that preserves the distributional shape of the $\log_2$-transformed data. All outcome variables were natural log-transformed to improve residual normality. Categorical covariates were entered as factors.

A hierarchical adjustment strategy was applied using three sequential models: Model 1 (Crude) was unadjusted; Model 2 (Demographic) was adjusted for age at outcome measurement (continuous, in years), because age is a fundamental determinant of health outcomes and a well-established confounder in perinatal studies[36,37]; Model 3 was further adjusted for maternal age, gestational age, PBMI, parity, maternal education, and family income. Covariates for Model 3 were selected a priori based on established literature and causal considerations to identify a sufficient set of confounders that would minimize bias without over-adjusting for mediators or colliders. Specifically, maternal age was included to control for age-related sociodemographic and biological risks[36,38]; gestational age was adjusted to account for the direct effect of prematurity on neonatal outcomes[39]; PBMI and parity were incorporated to address metabolic status and obstetric history, both of which independently influence pregnancy outcomes[40,41]; and maternal education and family income were added as socioeconomic proxies to capture health literacy, lifestyle, and healthcare access, which are consistently associated with perinatal health[42,43]. All covariates were treated as continuous or categorical as specified.

All models were repeated separately for male and female infants. To account for multiple comparisons, *P* values were adjusted using the Benjamini–Hochberg false discovery rate (FDR) method, applied separately by outcome and sex.

For each association, the standardized regression coefficient (β) corresponds to a one-standard-deviation increase in the $\log_2$-transformed metal concentration. To express effects on a more interpretable scale, β was converted to represent an IQR increase in the $\log_2$-transformed metal concentration. The IQR was calculated as the difference between the 75th and 25th percentiles of each metal's $\log_2$-transformed distribution in the study population. The conversion was performed using the formula[44]:

$$\beta_{IQR} = \beta \times \frac{IQR_{log_2}}{SD_{log_2}}$$

where IQR_{log_2} is the IQR of the $\log_2$-transformed metal and SD_{log_2} is its standard deviation. The percent change in the outcome per IQR increase was then computed as[44]:

$$\text{Percent change} = (\exp(\beta_{IQR}) - 1) \times 100\%$$

with 95% confidence intervals (CIs) derived from the corresponding bounds of β. Absolute changes in the original outcome units were estimated by applying these percentages to the mean outcome values.

To test whether the associations varied by sex, we added a metal × sex interaction term to the fully adjusted linear regression models. The interaction *P* value was used to assess effect modification by child sex. Positive interaction estimates indicate stronger associations in females; negative estimates indicate stronger associations in males. Interaction analyses were performed on the overall cohort (n = 248) using the same covariate set as Model 3.

To further characterize the exposure-response relationship, metal concentrations were categorized into quartiles (Q1 to Q4) based on their distributions in the study population, with the lowest quartile (Q1) serving as the reference category. Linear regression models were fitted using the same fully adjusted covariate set as in the single-metal analyses (Model 3), with the quartile variable entered as a categorical predictor. The percent changes in each outcome and their 95%CIs were estimated for Q2, Q3, and Q4 relative to Q1. To assess linear trends across increasing exposure levels, we performed a trend test by assigning the median concentration of each quartile category as a continuous score and including this score as a continuous predictor in the fully adjusted regression model. The resulting *P* for trend reflects the dose-response

relationship across increasing exposure levels, and the corresponding percent change per one-quartile increase and its 95%CI were derived. To visualize the dose-response relationships, scatter plots were generated for metals with *P* for trend < 0.05. Linear regression lines with 95%CI bands were fitted to illustrate the associations, with both axes $\log_2$-transformed.

Advanced computational methods

To complement the traditional regression and mixture analyses, we applied three advanced computational approaches - XGBoost, BART, and BKMR - to evaluate variable importance, non-linear dose-response relationships, and mixture effects. All three approaches were performed on the full set of 10 metals, using the same covariate set as Model 3. All analyses were performed separately for males and females.

XGBoost is a scalable, gradient-boosted decision tree algorithm implemented in the *xgboost* R package[27]. We trained separate models for males and females, using the same set of predictors (metals plus covariates in Model 3). Hyperparameters were tuned to balance model fit and generalizability: number of trees = 100, learning rate (eta) = 0.1, maximum tree depth = 3, subsample ratio = 0.8, and column subsampling (colsample_bytree) = 0.8. To assess model generalizability, we performed five-fold cross-validation and evaluated root mean square error (RMSE) across folds to see whether overfitting affected the variable importance rankings. Early stopping with a patience of 10 rounds was applied to prevent overfitting. The objective was set to "reg:squarederror" (i.e., regression with squared error loss) for our continuous outcome (natural log-transformed AGD). Feature importance was assessed using Gain, which measures the average improvement in accuracy (or reduction in loss) when a feature is used as a split point. Gain values are scale-independent and can be compared across features.

BART is a Bayesian non-parametric ensemble method that sums the predictions of many regression trees, implemented in the *BART* R package[28]. We fitted BART using 100 trees, 500 burn-in iterations, and 500 posterior draws. The prior parameters were left at default values (e.g., prior on tree depth and terminal node variance), which are known to work well for small-to-moderate sample sizes[28]. Variable importance was quantified by the split frequency - the total number of times a variable was selected as a splitting rule across all trees and Markov chain Monte Carlo (MCMC) draws. This metric reflects how often a variable contributes to the ensemble's predictive structure. Because BART automatically handles non-linearities and interactions, no explicit tuning of interaction terms was required.

BKMR is a flexible, non-parametric Bayesian regression method for estimating the joint effect of a mixture, implemented in the *bkmr* R package[30]. We performed BKMR in two complementary ways. First, to evaluate the overall mixture effect, we included all 10 metals in a single BKMR model without hierarchical variable selection and estimated the change in AGD when all metals were jointly increased from the 10th to the 90th percentile of their joint distribution, using the 50th percentile (median) as the reference point. Second, metals were grouped for hierarchical variable selection based on their pairwise Spearman correlation coefficients, using a threshold of $\rho \geq 0.5$ to define moderately or strongly correlated metals[29]. Groups were formed only when metals met this correlation threshold; if no correlations reached this threshold within a sex, no groups were formed. We reported group-level posterior inclusion probabilities (PIPs) to assess the importance of each metal group. In addition, individual PIPs were extracted from both the ungrouped and grouped BKMR models: plain PIPs from the model without hierarchical variable selection, and individual PIPs (grouped) from the model with grouping. For groups containing more than one metal, the individual PIPs from the grouped model reflect the relative contribution of each metal within its group. The BKMR models were run using a Gaussian kernel with 25,000 MCMC iterations, a burn-in of 5,000, and a thinning interval of 5. All models were fitted separately for males and females. The exposure variables were Z-standardized prior to analysis (as described above), while covariates were included as linear terms.

Sensitivity analyses

To further control for potential confounding and assess the robustness of the associations, we performed propensity score stratification[45,46]. For each metal, exposure was dichotomized into high *vs.* low groups using the sex-specific median concentration. A logistic regression model including all covariates from the fully adjusted model (Model 3) was used to estimate the propensity score (probability of being in the high-exposure group). All 10 metals were analyzed separately by sex. The study population was then divided into five strata according to the quintiles of the propensity score, ensuring that within each stratum the distribution of covariates was balanced between exposure groups. Within each stratum, a linear regression model comparing AGD between high- and low-exposure groups was fitted without further covariate adjustment. Stratum-specific effect estimates were pooled using inverse-variance weighting.

As an additional sensitivity analysis, we applied overlap weighting, which focuses on the subpopulation with good overlap in propensity scores[47]. Exposure was dichotomized as described above. Propensity scores were estimated using the same logistic regression model. Overlap weights were calculated as $w = PS \times (1 - PS)$, where PS is the propensity score. These weights assign higher weights to individuals with propensity scores near 0.5 and down-weight those with extreme scores, thereby targeting the average treatment effect in the overlap population (ATO). Weighted linear regression was then used to estimate the association between high exposure and AGD. Robust standard errors were obtained via the survey package. The same sex- and metal-specific subsets were analyzed. The β coefficients represent the mean difference in natural log-transformed AGD between the high-exposure and low-exposure groups.

To evaluate the robustness of the observed associations, we conducted two sensitivity analyses: birth weight adjustment and extreme value treatment. All analyses were performed on all 10 metals separately by sex, using the same covariate set as Model 3. For the birth weight adjustment analysis, missing birth weight values were imputed using sex-specific medians. We compared models with and without birth weight as an additional covariate. For the extreme value treatment, metal concentrations were processed using two approaches at the 5th and 95th percentiles: (1) truncation, excluding observations below the 5th or above the 95th percentile; and (2) winsorization, replacing extreme values with the 5th or 95th percentile values while retaining all observations. These methods allow assessment of whether the associations are driven by extreme observations.

All statistical analyses were performed using R software (version 4.5.2). Key R packages included *tidyverse*, *survey*, *xgboost*, *BART*, and *bkmr*. Figures were generated using GraphPad Prism 10 (Dotmatics, Boston, MA, USA) and R. A two-sided P value < 0.05 was considered statistically significant.

RESULTS AND DISCUSSION

Study population characteristics

A total of 248 mother-infant pairs were included in the analyses. Table 1 presents the baseline characteristics stratified by infant sex. Maternal factors including maternal age, education level, household income, pre-pregnancy BMI, and gestational age were comparable between male and female infants (all $P > 0.05$). However, parity differed significantly by sex, with primiparous mothers more common in females than in males (87.5% *vs.* 70.6%, $P = 0.004$). Regarding child characteristics, birth weight differed marginally between sexes ($P = 0.050$), with low birth weight (< 2,500 g) present only in females (3.6%); infant age was slightly but significantly older in females (median 2.0 *vs.* 2.0 years, $P = 0.041$). Both AGD measures were significantly larger in males than in females: AGD-AS/AGD-AF (37.2 mm *vs.* 23.9 mm, $P < 0.001$) and AGD-AP/AGD-AC (83.6 mm *vs.* 56.7 mm, $P < 0.001$). Detailed distributions of AGD measures, including selected percentiles (2.5th, 25th, 50th, 75th, and 97.5th), are provided in Supplementary Table 2. Spearman correlation analysis revealed a moderate positive correlation between the two AGD measures in both sexes: ρ = 0.493 ($P < 0.001$) for males (AGD-AP *vs.* AGD-AS) and ρ = 0.43 ($P < 0.001$) for females (AGD-AC *vs.* AGD-AF) [Supplementary Figure 1].

Table 1. Baseline characteristics of the study participants, stratified by infant sex

Characteristic	Male (*n* = 136)	Female (*n* = 112)	*P* value
Maternal factors			
Maternal age [*n*, (%)]			0.364
Young maternal group (≤ 24 years)	4 (2.9)	5 (4.5)	
Prime reproductive group (25-34 years)	111 (81.6)	96 (85.7)	
Older (≥ 35 years)	21 (15.4)	11 (9.8)	
Maternal education level [*n*, (%)]			0.134
Junior high or below	9 (6.6)	4 (3.6)	
High school	10 (7.4)	16 (14.3)	
College or above	117 (86.0)	92 (82.1)	
Household income [*n*, (%)]			0.338
Low	19 (14.0)	12 (10.7)	
Middle	104 (76.5)	83 (74.1)	
High	13 (9.6)	17 (15.2)	
Maternal PBMI [*n*, (%)]			0.249
Underweight (< 18.5)	22 (16.2)	17 (15.2)	
Normal (18.5-23.9)	86 (63.2)	82 (73.2)	
Overweight (24-27.9)	22 (16.2)	11 (9.8)	
Obese (≥ 28)	6 (4.4)	2 (1.8)	
Gestational age [*n*, (%)]			0.522
Preterm (< 37 weeks)	3 (2.2)	5 (4.5)	
Term (≥ 37 weeks)	133 (97.8)	107 (95.5)	
Parity [*n*, (%)]			**0.004**
Primiparous	96 (70.6)	98 (87.5)	
Second	37 (27.2)	14 (12.5)	
Third or more	3 (2.2)	0 (0.0)	
Child factors			
Birth weight [*n*, (%)]			**0.050**
Low birth weight (< 2,500 g)	0 (0.0)	4 (3.6)	
Normal (2,500-3,999 g)	120 (88.2)	102 (91.1)	
High birth weight (≥ 4,000 g)	8 (5.9)	3 (2.7)	
Missing	8 (5.9)	3 (2.7)	
Infant age (median [IQR])	2.0 [2.0, 2.0]	2.0 [2.0, 2.1]	**0.041**
AGD-AS/AGD-AF (median [IQR])	37.2 [33.1, 42.5]	23.9 [21.3, 26.6]	**< 0.001**
AGD-AP/AGD-AC (median [IQR])	83.6 [77.4, 91.2]	56.7 [51.4, 62.9]	**< 0.001**

Data are presented as *n* (%) for categorical variables and median [IQR] for continuous variables. *P* values for categorical variables were derived from the chi-square test or Fisher's exact test, as appropriate; a single *P* value is reported for each categorical variable, comparing male and female groups. *P* values for continuous variables were derived from the Wilcoxon rank-sum test. *P* values < 0.05 were considered statistically significant. The bold values in the table indicate statistically significant results ($P < 0.05$). PBMI: Pre-pregnancy body mass index; IQR: interquartile range; AGD-AS: anogenital distance (anus-scrotum) for males; AGD-AF: anogenital distance (anus-fourchette) for females; AGD-AP: anogenital distance (anus-penis) for males; AGD-AC: anogenital distance (anus-clitoris) for females; g: grams.

Prenatal metal exposure characteristics

Concentrations of the 10 metals in first-trimester maternal plasma are summarized by infant sex in Supplementary Table 3. Detection rates exceeded 90% for all metals except Cd (males: 61.8%; females: 69.6%) and Sn (males: 78.7%; females: 65.2%), which remained above 60%. No statistically significant differences

were observed between male and female infants for any metal (all $P > 0.05$), although Zn and As showed borderline differences ($P = 0.055$ and $P = 0.053$, respectively).

Supplementary Figure 2 shows Spearman's correlation matrices for ten plasma metals (As, Cd, Cr, Cu, Hg, Ni, Pb, Sn, Tl, Zn) from first-trimester maternal samples stratified by infant sex. Both subgroups showed strong positive inter-metal correlations. For male infants, the top-ranked correlations were Cd-Zn (ρ = 0.65) and Cd-Pb (ρ = 0.63); Tl was moderately correlated with As (0.53), Pb (0.43), and Cr (0.48). Cu and Hg barely correlated with other metals, indicating separate exposure sources. The correlation structure was comparable among mothers carrying female infants yet exhibited stronger metal covariation, dominated by Pb-Zn (ρ = 0.71), Cd-Zn (0.66), As-Cd (0.60), and As-Pb (0.60). A prominent sex-specific distinction was observed: Ni displayed weak negative correlations with As (-0.23) and Cd (-0.21) in the female group, while Ni was only positively correlated with Cr (0.27) in the male group. Overall, these correlation patterns indicate a stable core metal cluster among As, Cd, Pb, and Zn in both sexes, with stronger covariation in the female subgroup, and highlight Ni's sex-specific correlation profile, which may reflect differences in exposure sources or metabolic handling.

Associations with AGD

In sex-stratified linear regression models, we found distinct metal-AGD associations that differed between males and females, with adverse associations predominantly observed in males [Figure 1 and Supplementary Tables 4-7]. In males, three metals - Cd, Tl, and As - were significantly associated with reduced AGD-AP [Figure 1 and Supplementary Table 5]. Cd showed the strongest effect, with a -6.96% change (95%CI: -9.22 to -4.63; $P = 0.001$) per IQR increase in Cd concentration, supported by a clear dose-response trend ($P = 0.011$) [Supplementary Table 7]. Tl (-2.16%, 95%CI: -3.89 to -0.40; $P = 0.014$) and As (-2.83%, 95%CI: -4.87 to -0.75; $P = 0.007$) per IQR increase showed smaller but consistent effects [Figure 1 and Supplementary Table 5]. All three metals remained significant after FDR correction ($P_{FDR} = 0.013$, 0.040, and 0.033, respectively). The biological plausibility of these associations is supported by a common biological pathway: they interfere with testicular steroidogenesis - Cd and As by downregulating key enzymes (Star, Cyp17a1, Hsd3b)[48-50] and Tl by inducing Sertoli cell oxidative stress - leading to reduced testosterone biosynthesis and impaired androgen signaling during the male programming window[51].

Pb (-3.48%, 95%CI: -6.48 to -0.39; $P = 0.028$) and Cr (-3.19%, 95%CI: -5.88 to -0.42; $P = 0.025$) were significantly associated with AGD-AP, but were not significant after FDR correction (Pb: $P_{FDR} = 0.056$; Cr: $P_{FDR} = 0.053$) [Figure 1 and Supplementary Table 5]. The 2020 Guangxi cohort also reported inverse associations for Pb and Cr with AGD in boys[26], consistent with their roles as androgen receptor antagonists (Pb)[52,53] and disruptors of testicular steroidogenesis (Cr)[54,55]. The attenuated of significance for these two metals may reflect limited statistical power rather than absence of effect.

Zn showed a non-significant inverse trend, while Cu, Hg, Ni, and Sn were not associated with AGD-AP. Dose-response analyses supported the main findings [Supplementary Tables 6 and 7]: Cd, Pb, As, Cr, and Tl all showed significant negative trends (all P for trend < 0.05) for AGD-AP, with Cd and Tl showing the strongest. The dose-response relationships were visually confirmed in scatter plots of $\log_2$-transformed metal concentrations against natural log-transformed AGD-AP [Supplementary Figure 3], where linear regression lines with 95% confidence bands illustrated the monotonic decreasing trends for these metals. In contrast, none of the metals were significantly associated with AGD-AS (all P and $P_{FDR} > 0.05$), suggesting that AGD-AP is a more sensitive biomarker of prenatal metal-induced androgen disruption in males. Overall, these consistent dose-response patterns, combined with FDR-adjusted significance for As, Cd, and Tl, strengthen the inference that prenatal exposure to these metals may exert biologically meaningful effects on male genital development.

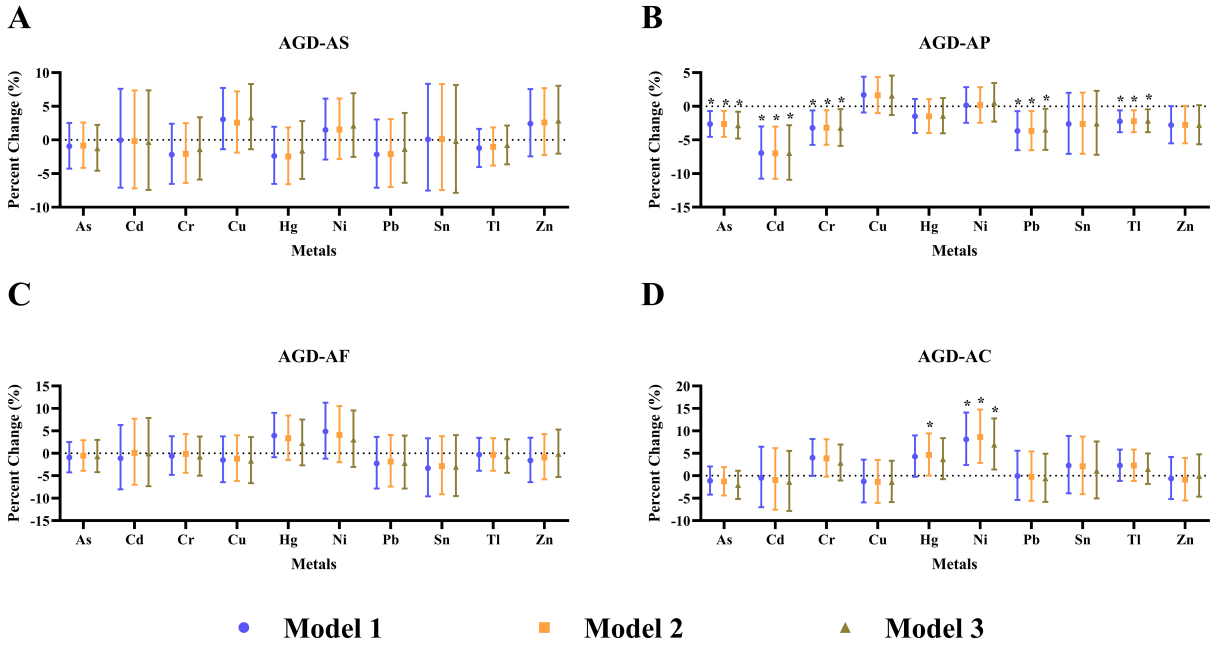


Figure 1. Sex-stratified associations between first-trimester plasma metal concentrations and AGD in 2-year-old children. Panels show percent change (with 95%CI) in AGD-AS/AGD-AF (A and C) and AGD-AP/AGD-AC (B and D) per IQR increase in each metal concentration, estimated using three sequential models: Model 1 (unadjusted), Model 2 (adjusted for children's age), and Model 3 (fully adjusted for children's age, maternal age, gestational age, PBMI, parity, maternal education, and family income). Metal concentrations were $\log_2$-transformed and Z-standardized; the percent change corresponds to an increase from the 25th to the 75th percentile of the $\log_2$-transformed concentration. Outcome variables were natural log-transformed. Statistically significant associations (two-sided unadjusted $P < 0.05$) are marked with an asterisk (*). Error bars represent 95%CI. Sample sizes: males $n = 136$; females $n = 112$. AGD: Anogenital distance; CI: confidence interval; AGD-AS: anogenital distance (anus-scrotum) for males; AGD-AF: anogenital distance (anus-fourchette) for females; AGD-AP: anogenital distance (anus-penis) for males; AGD-AC: anogenital distance (anus-clitoris) for females; IQR: interquartile range; PBMI: pre-pregnancy body mass index.

In females, only Ni showed a suggestive positive association with AGD-AC in the fully adjusted model, corresponding to a 6.94% increase (95%CI: 1.40-12.79; $P = 0.014$) per IQR increase in Ni concentration, although the FDR-adjusted P value did not reach statistical significance ($P_{FDR} = 0.140$). No other metals were significantly associated with AGD-AC in females (all P and $P_{FDR} > 0.05$). The suggestive Ni-AGD-AC association was supported by a significant dose-response trend (P for trend = 0.005) [Supplementary Table 7], and the positive dose-response relationship was visually confirmed in Supplementary Figure 3 for Ni in females. In female rats, Ni exposure decreased estradiol and caused ovarian damage[56]; in girls, prenatal Ni was associated with slower breast development[57], suggesting anti-estrogenic actions. This effect could theoretically reduce estrogenic tone, making androgen-mediated AGD elongation more apparent in females.

To formally evaluate whether the associations differed by sex, we included a metal × sex interaction term in the fully adjusted models [Supplementary Table 8]. Significant interactions were observed for Ni ($P_{interaction} = 0.011$), Tl ($P_{interaction} = 0.015$), Cr ($P_{interaction} = 0.002$), and Hg ($P_{interaction} = 0.019$) for AGD-AP/AC, with positive interaction estimates indicating stronger associations in females than in males. For Ni, the interaction reflected its positive association in females (6.94%, $P = 0.014$) and null association in males. For Tl and Cr, the significant interactions reflected their inverse associations in males (Tl: -2.16%, $P = 0.014$; Cr: -3.19%, $P = 0.025$) that were attenuated to null in females. For Hg, despite non-significant associations in both sexes, the positive interaction estimate suggests a potential sex-dependent pattern warranting further investigation. Overall, these interaction results confirm that sex is an important modifier of metal-AGD associations, with effects generally more pronounced in males for AGD-AP and in females for AGD-AC.

Table 2. Importance of metals from three computational models for AGD by sex

Metal	XGBoost (Gain)	BART (Split Frequency)	BKMR Plain PIP	BKMR Group PIP	BKMR Individual PIP (grouped)
Male					
Tl	0.127	2,361	0.1360	0.2825	0.1725
Cd	0.094	4,178	0.0355	0.0613	0.0450
As	0.031	2,867	0.0675	0.2825	0.1100
Pb	0.060	3,733	0.0138	0.0613	0.0148
Cr	0.104	2,061	0.0070	0.0073	0.0073
Ni	0.113	1,670	0.0050	0.0183	0.0183
Cu	0.092	2,411	0.0000	0.0068	0.0068
Zn	0.026	3,164	0.0003	0.0613	0.0015
Sn	0.037	2,784	0.0013	0.0025	0.0025
Hg	0.077	2,368	0.0000	0.0000	0.0000
Female					
As	0.112	3,287	0.1840	0.1963	0.1830
Ni	0.080	3,517	0.0158	0.0203	0.0203
Cd	0.023	3,652	0.0073	0.1963	0.0060
Pb	0.078	2,602	0.0033	0.1963	0.0058
Zn	0.032	2,555	0.0020	0.1963	0.0015
Cr	0.081	1,486	0.0033	0.0055	0.0055
Sn	0.083	3,967	0.0070	0.0060	0.0060
Hg	0.089	3,181	0.0005	0.0143	0.0143
Cu	0.051	2,687	0.0010	0.0038	0.0038
Tl	0.041	2,154	0.0003	0.0008	0.0008

XGBoost Gain measures the average information gain; higher values indicate greater contribution to model predictive performance. BART Split Frequency is the total number of times a variable was selected as a splitting rule across all trees and MCMC draws. BKMR Plain PIP: Individual posterior inclusion probability from the ungrouped BKMR model. BKMR Group PIP: Probability that a metal group was selected in the grouped BKMR model. BKMR Individual PIP (grouped): Individual posterior inclusion probability for each metal estimated from the grouped BKMR model. For single-metal groups, this value represents the marginal inclusion probability of the metal in the grouped model. Metals are ordered by their relevance based on the combined evidence from all three models. Sample sizes: males n = 136; females n = 112. AGD: Anogenital distance; XGBoost: extreme gradient boosting; BART: Bayesian additive regression trees; BKMR: Bayesian kernel machine regression; PIP: posterior inclusion probability; MCMC: Markov Chain Monte Carlo.

Advanced computational analysis

To complement the single-metal linear regression analyses and assess the robustness of the findings, we applied three complementary computational models - XGBoost, BART, and BKMR [Table 2]. These models are sensitive to different aspects of the data: XGBoost captures linear and weak non-linear additive effects[27], BART is more flexible for complex non-linearities and interactions[28,58], and BKMR with hierarchical variable selection handles correlated exposures by grouping metals based on Spearman correlation ($\rho \geq 0.5$ for exploratory analysis) and estimating group-level and BKMR individual PIP (grouped)[29]. XGBoost models were evaluated by five-fold cross-validation to assess generalizability; stable predictive errors across folds (mean RMSE: 0.118 ± 0.012 for males and 0.176 ± 0.021 for females) indicated that variable importance rankings were not substantially affected by overfitting.

In males, Tl showed the highest XGBoost Gain (0.127) and the highest BKMR Individual PIP (grouped) (0.1725). Within the As–Tl group (group PIP = 0.2825), Tl dominated the group effect (0.1725 *vs.* 0.1100 for As). Cd had the highest BART split frequency (4,178) and strong XGBoost Gain (0.094). Within the Cd–Zn–Pb group (group PIP = 0.0613), Cd was the dominant contributor (0.0450 *vs.* 0.0015 for Zn and

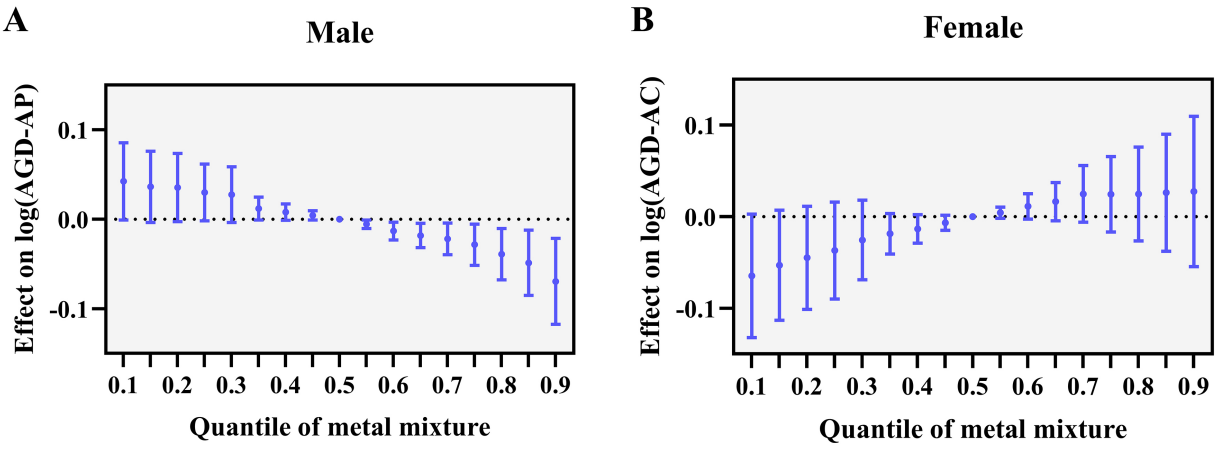


Figure 2. Overall mixture effects of 10 metals on AGD from BKMR stratified by sex. (A) Males: AGD-AP; (B) Females: AGD-AC. The solid line represents the estimated effect on the natural log-transformed AGD (95%CI). The dashed horizontal line at zero indicates no effect. The x-axis represents the quantile of the metal mixture, where 0.5 is the reference point (all metals fixed at their median concentrations). Values below 0.5 indicate joint decreases and values above 0.5 indicate joint increases in all metal concentrations. All models were adjusted for child's age, maternal age, gestational age, PBMI, parity, maternal education and household income. Sample sizes: males *n* = 136; females *n* = 112. AGD: Anogenital distance; BKMR: Bayesian kernel machine regression; AGD-AP: anogenital distance (anus-penis) for males; AGD-AC: anogenital distance (anus-clitoris) for females; CI: confidence interval; PBMI: pre-pregnancy body mass index.

0.0148 for Pb). These findings independently confirm Cd and Tl as the key metals driving the inverse AGD-AP associations observed in linear regressions.

Notably, Pb exhibited high BART frequency (3,733) but low BKMR Individual PIP (grouped) (0.0148), suggesting its apparent single-metal effect largely reflects co-exposure with Cd. This is consistent with the understanding that Cd and Pb are common co-contaminants from shared environmental sources[59,60]. The BKMR Individual PIP (grouped) results indicate that Cd drives the joint mixture effect, with Pb contributing little additional independent signal once Cd is accounted for. This explains why Pb was significant in single-metal models but not robust in co-exposure-controlled analyses.

In females, As was the most important metal across all models, with the highest BKMR Plain PIP (0.1840), highest XGBoost Gain (0.112), and substantial BART split frequency (3,287). Within the As–Zn–Cd–Pb group (group PIP = 0.1963), As overwhelmingly drove the group effect (PIP: 0.183 *vs.* < 0.006 for the others). This confirms the importance of As in the mixture, though its single-metal association was non-linear (only Q2 *vs.* Q1 was significant, $P = 0.008$). Ni showed moderate importance (BART Split Frequency: 3,517, BKMR Plain PIP: 0.0158), providing independent support for its positive linear association (Model 3: 6.94%, $P = 0.014$), though not significant after FDR correction.

BART and XGBoost showed some divergence for certain metals. For example, Cd in females had a high BART split frequency but a low XGBoost Gain, suggesting potential non-linear effects that BART is more sensitive to[28,58]. This divergence highlights the value of using multiple complementary approaches to characterize the exposure-response relationship.

In males, the BKMR overall mixture effect was significant and negative, indicating that joint exposure to the 10 metals was associated with reduced AGD-AP [Figure 2]. In females, the overall mixture effect was not significant, suggesting that opposing metal effects may cancel each other out [Figure 2].

Table 3. Propensity score-based sensitivity analyses for associations between each metal and AGD-AP (males)/AGD-AC (females), stratified by sex

Metal	Propensity score stratification		Overlap weighting	
	β (95%CI)	*P*	β (95%CI)	*P*
Male				
Cr	-0.0372 (-0.0792, 0.0049)	0.083	-0.0419 (-0.0841, 0.0003)	0.054
Ni	0.0163 (-0.0259, 0.0585)	0.449	0.0065 (-0.0352, 0.0482)	0.76
Cu	0.0060 (-0.0381, 0.0501)	0.789	0.0228 (-0.0212, 0.0668)	0.312
Zn	**-0.0464 (-0.0887, -0.0042)**	**0.031**	**-0.0478 (-0.0886, -0.0071)**	**0.023**
As	-0.0361 (-0.0781, 0.0059)	0.092	-0.0386 (-0.0797, 0.0025)	0.068
Cd	-0.0367 (-0.0757, 0.0023)	0.065	**-0.0548 (-0.0955, -0.0142)**	**0.009**
Sn	-0.0354 (-0.0756, 0.0048)	0.084	-0.0266 (-0.0679, 0.0146)	0.207
Hg	-0.0324 (-0.0732, 0.0084)	0.120	-0.0235 (-0.0654, 0.0183)	0.272
Tl	**-0.0587 (-0.0981, -0.0194)**	**0.003**	**-0.0627 (-0.1026, -0.0228)**	**0.003**
Pb	-0.0333 (-0.0747, 0.0081)	0.115	-0.0338 (-0.0751, 0.0075)	0.111
Female				
Cr	0.0085 (-0.0594, 0.0764)	0.807	0.0277 (-0.0398, 0.0951)	0.423
Ni	0.0587 (-0.0058, 0.1232)	0.074	0.0638 (-0.0017, 0.1292)	0.059
Cu	0.0021 (-0.0611, 0.0653)	0.948	-0.0031 (-0.0701, 0.0639)	0.928
Zn	0.0249 (-0.0474, 0.0973)	0.499	0.0317 (-0.0374, 0.1008)	0.370
As	-0.0473 (-0.1167, 0.0220)	0.181	0.0033 (-0.0641, 0.0708)	0.923
Cd	-0.0215 (-0.0913, 0.0483)	0.546	-0.0185 (-0.0877, 0.0507)	0.601
Sn	0.0003 (-0.0657, 0.0662)	0.994	0.0361 (-0.0334, 0.1055)	0.311
Hg	0.0222 (-0.0423, 0.0866)	0.501	0.0570 (-0.0091, 0.1231)	0.094
Tl	0.0110 (-0.0551, 0.0770)	0.745	0.0201 (-0.0484, 0.0885)	0.567
Pb	-0.0070 (-0.0761, 0.0622)	0.843	0.0058 (-0.0612, 0.0728)	0.865

Covariates (child's age, maternal age, gestational age, PBMI, parity, maternal education, and household income) were included in the logistic regression model for estimating propensity scores. Exposure was dichotomized at the sex-specific median for each metal. For propensity score stratification, stratum-specific estimates were pooled using inverse-variance weighting across five strata. For overlap weighting, weights were calculated as $PS \times (1 - PS)$, targeting the ATO. β represents the mean difference in natural log-transformed AGD between the high-exposure group (above the sex-specific median) and the low-exposure group (below the sex-specific median), estimated from propensity score stratification (inverse-variance weighted average across strata) and overlap weighting (weighted linear regression), respectively. Statistical significance ($P < 0.05$) is indicated in bold. Sample sizes: males $n = 136$; females $n = 112$. The bold values in the table indicate statistically significant results ($P < 0.05$). AGD-AP: Anogenital distance (anus-penis) for males; AGD-AC: anogenital distance (anus-clitoris) for females; CI: confidence interval; PBMI: pre-pregnancy body mass index; PS: propensity score; ATO: average treatment effect in the overlap population; AGD: anogenital distance.

In summary, the computational models converged on a consistent narrative: Tl and Cd in males, and As in females are the most robustly associated metals with AGD, while other metals either showed model-dependent importance or consistently low importance. These results strengthen the inference from traditional regression analyses and highlight the value of integrating multiple analytical approaches to disentangle the effects of correlated environmental exposures.

Sensitivity analyses

To assess robustness against residual confounding, model specification, and extreme observations, we conducted three sensitivity analyses: propensity score-based methods, birth weight adjustment, and extreme value treatment [Table 3, Supplementary Tables 9 and 10].

Propensity score stratification and overlap weighting confirmed that male Tl remained significantly associated with AGD-AP across both methods (β = -0.0587 and -0.0627, both $P < 0.01$). The discrepancy for

Cd between stratification and overlap weighting may partly reflect the lower detection frequency of Cd (61.76% for males and 69.64% for females), which reduces the precision of estimates within strata in the stratification approach, whereas overlap weighting, by retaining full sample information and using continuous weights, improves estimation efficiency[61,62]. Zn became significant in both propensity score methods (P = 0.031 and 0.023), whereas As and Pb were attenuated and no longer significant (P > 0.05). In females, Ni showed borderline positive associations (P = 0.074 and 0.059).

Given that birth weight may lie on the causal pathway between prenatal metal exposure and AGD - which could introduce overadjustment bias - we compared models with and without birth weight adjustment. Comparison of models with and without birth weight adjustment yielded nearly identical estimates for all metals (e.g., male As: -0.0305 *vs.* -0.0302; female Ni: 0.0487 *vs.* 0.0482) [Supplementary Table 9], indicating that birth weight does not materially confound or mediate the observed associations[61].

Extreme value treatment (truncation and winsorization at 5th/95th percentiles) confirmed that male Cd, male Tl, and female Ni remained significant across all three approaches [Supplementary Table 10]. In contrast, male As and Pb lost significance after truncation (As: P = 0.007 *vs.* 0.358; Pb: P = 0.028 *vs.* 0.284) but remained significant in winsorized models (As: P = 0.013; Pb: P = 0.034), suggesting these associations may be partially driven by extreme observations. Female Hg became significant only after truncation (P = 0.105 *vs.* 0.015), hinting at a potential non-linear effect.

Collectively, these sensitivity analyses consistently identified male Tl, male Cd, and female Ni as the most robust associations. The associations for male As and Pb showed method dependence and warrant cautious interpretation. These findings support the robustness of the primary conclusions regarding sex-specific effects of prenatal metal exposure on AGD.

Limitations

Several limitations should be acknowledged. First, our sample size (n = 248) is modest, which may have limited power to detect smaller mixture effects or weaker interactions. However, we were still able to detect significant single-metal associations and consistent dose-response trends, suggesting that power was sufficient for the main findings. Second, metal concentrations were measured only once in the first trimester, which may not capture full exposure variability across pregnancy. Nevertheless, the first trimester (8-14 weeks) represents the critical window for genital development[13,14], and single-spot measurements in this period have been widely used and accepted in similar studies as a reasonable proxy[63,64]. Third, AGD was measured at age 2 years rather than at birth. However, AGD is known to track stably from infancy through early childhood[65,66], and our measurements were performed by trained examiners blinded to exposure status using standardized protocols, minimizing bias. Moreover, postnatal growth is unlikely to substantially alter the ranking of AGD within the cohort, supporting the validity of our exposure-outcome assessment. Fourth, the present analysis included a subset of the original cohort with complete data on plasma metals and AGD measurements. The included participants covered the full range of key characteristics of the broader cohort[31-33]. Importantly, the consistency of our findings across multiple analytical approaches - including linear regression, dose-response analysis, propensity score methods, and machine learning - further supports the robustness of the observed associations and suggests that selection bias, if present, is unlikely to have materially affected the conclusions.

CONCLUSIONS

In this prospective cohort study, prenatal metal exposures exerted sex-specific opposing effects on AGD. Tl in males was the most robust association, remaining significant across all sensitivity analyses. Cd in males and Ni in females showed consistent associations in single-metal and dose-response analyses, with additional

support from overlap weighting for Cd and borderline support from propensity score-based methods for Ni. Machine learning independently identified Tl and Cd as the top male predictors and As as the top female predictor. The associations for As and Pb were method-dependent and warrant cautious interpretation. Our findings underscore sex-specific susceptibility as a critical modifier of metal toxicity and highlight the value of integrating machine learning with traditional epidemiological approaches to characterize environmental exposure-response relationships. Future studies should elucidate the underlying sex-specific mechanisms and validate these findings in larger, more diverse populations.

DECLARATIONS

Acknowledgments

We thank all participants and investigators for their invaluable contributions to the successful completion of this birth cohort study.

Authors' contributions

Writing - original draft, data curation, formal analysis, investigation, methodology, funding acquisition, visualization: Liu, J.
Investigation, validation, visualization: Li, S.
Supervision, validation, project administration: Li, Y.
Resources, supervision, project administration: Xia, W.
Writing - review and editing, conceptualization, supervision: Liu, H.
Writing - review and editing, conceptualization, funding acquisition, project administration, resources, supervision: Xu, S.

Availability of data and materials

The data presented in this study are available on request from the corresponding author due to ethical and privacy restrictions.

AI and AI-assisted tools statement

During the preparation of this work, the authors used Nano Banana Pro (released November 2025) solely for creating the silhouette of the maternal figure in the Graphical Abstract. Additionally, DeepSeek (Version V3, released May 2025) was used for language polishing and readability improvement of the manuscript. These tools did not influence the study design, data collection, analysis, interpretation, or the scientific content of the work. All authors take full responsibility for the accuracy, integrity, and final content of the manuscript.

Financial support and sponsorship

This study was funded by the National Natural Science Foundation of China (grant 22236001), the China Postdoctoral Science Foundation (grant 2025M770743), Hainan Provincial Natural Science Foundation of China (grant 826QN0572), and the Postdoctoral Research Project of Hainan Province (grant 2025-38).

Conflicts of interest

Li, Y. is an Editorial Board Member of the journal *Journal of Environmental Exposure Assessment*. Liu, H. is a Youth Editorial Board Member of the journal *Journal of Environmental Exposure Assessment*. They were not involved in any stage of editorial processing, including reviewer selection, manuscript handling, and decision-making. The other authors declare no conflicts of interest.

Ethical approval and consent to participate

The study protocol was reviewed and approved by the Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology ([2012] (14#)) and the Institutional Review Board of Wuhan Medical & Healthcare Center for Women and Children (No. 2010009 and 2016003) prior to the initiation of maternal recruitment. At enrollment, written informed consent was obtained from all participating mothers. This consent covered telephone follow-ups, questionnaire surveys, and the collection of biological samples during routine prenatal care, delivery hospitalization, and subsequent child

development follow-up. During the follow-up phase, children and their parents were contacted by telephone and invited to attend physical examinations and questionnaire surveys. For the 2-year-old child assessments (including AGD measurements and questionnaire administrations), written informed consent was obtained from the parents or legal guardians of all participating children prior to any study procedures. All parents and guardians were fully informed that participation was voluntary and that they could withdraw their child from the study at any time without any negative consequences.

Consent for publication

Not applicable.

Copyright

Supplementary Materials

Supplementary Materials

REFERENCES

1. Jalil, A.; Noor, M. J.; Iram, S.; Qaisar, I. A systematic review: global trend in heavy metal concentration in vegetables, dietary exposure, and health risk assessment. *Environ. Monit. Assess.* **2026**, *198*, 252. DOI PubMed
2. Jiao, L.; Huo, L.; Stein, A.; An, Y.; Zhang, P. A framework for unveiling multi-pathway mechanisms of soil heavy metal pollution: evidence from five major urban agglomerations in China. *J. Environ. Manage.* **2026**, *401*, 128952. DOI
3. Xiao, N.; Wang, Q.; Wang, Y.; Yao, Y.; Shen, C. Four decades of global cereal lead contamination patterns. *Ecotoxicol. Environ. Saf.* **2026**, *309*, 119687. DOI PubMed
4. Duarte, N. A. A.; Santos, L. C. D.; Maraslis, F. T.; et al. Blood-based transcriptomic profiles associated with chronic mercury exposure in Amazonian riverine populations. *J. Hazard. Mater.* **2026**, *504*, 141273. DOI PubMed
5. Jomova, K.; Alomar, S. Y.; Nepovimova, E.; Kuca, K.; Valko, M. Heavy metals: toxicity and human health effects. *Arch. Toxicol.* **2025**, *99*, 153-209. DOI PubMed PMC
6. Kinjo, Y.; Shibata, E.; Askew, D. J.; Japan Environment and Children's Study Group. Association of placental weight at birth with maternal whole blood concentration of heavy metals (cadmium, lead, mercury, selenium, and manganese): The Japan Environment and Children's Study (JECS). *Environ. Int.* **2024**, *188*, 108725. DOI PubMed
7. Zhou, J.; Tong, J.; Chen, S.; et al. Placental transfer efficiency of metals and its association with cognitive development in offspring: roles of placental inflammatory cytokines and oxidative stress markers. *J. Hazard. Mater.* **2025**, *500*, 140380. DOI PubMed
8. Yu, G.; Meng, X.; Tan, Y. Q.; et al. Gestational exposure to metals and small vulnerable newborns: a systematic review and meta-analysis. *Environ. Int.* **2026**, *208*, 110052. DOI PubMed
9. Adibi, J. J.; Zhao, Y.; Koistinen, H.; et al. Molecular pathways in placental-fetal development and disruption. *Mol. Cell. Endocrinol.* **2024**, *581*, 112075. DOI PubMed PMC
10. Andonotopo, W.; Bachnas, M. A.; Dewantiningrum, J.; et al. Endocrine disrupting chemicals: translating mechanisms into perinatal risk assessment. *J. Perinat. Med.* **2025**, *53*, 1238-52. DOI PubMed
11. Grandjean, P.; Landrigan, P. J. Developmental neurotoxicity of industrial chemicals. *Lancet* **2006**, *368*, 2167-78. DOI PubMed
12. Barouki, R.; Gluckman, P. D.; Grandjean, P.; Hanson, M.; Heindel, J. J. Developmental origins of non-communicable disease: implications for research and public health. *Environ. Health.* **2012**, *11*, 42. DOI PubMed PMC
13. Gu, J.; Jin, Z.; Wu, G.; et al. Correlation between essential and toxic elements in maternal blood during early pregnancy and atrial septal defects/ventricular septal defects/patent ductus arteriosus in offspring. *Birth. Defects. Res.* **2024**, *116*, e2282. DOI PubMed
14. Barrett, E. S.; Rivera-Núñez, Z.; Getz, K.; et al. Protective role of the placental efflux transporter BCRP/ABCG2 in the relationship between prenatal cadmium exposure, placenta weight, and size at birth. *Environ. Res.* **2023**, *225*, 115597. DOI PubMed PMC
15. Barrett, E. S.; Sharghi, S.; Thurston, S. W.; et al. Associations of exposure to air pollution during the male programming window and mini-puberty with anogenital distance and penile width at birth and at 1 year of age in the multicenter U.S. TIDES cohort. *Environ. Health. Perspect.* **2023**, *131*, 117001. DOI PubMed PMC
16. Cajachagua-Torres, K. N.; Salvi, N. B.; Seok, E.; et al. Prenatal phthalate exposure and anogenital distance in infants at 12 months. *Environ. Int.* **2025**, *199*, 109472. DOI PubMed PMC
17. Barrett, E. S.; Sathyanarayana, S.; Mbowe, O.; et al. First-trimester urinary bisphenol A concentration in relation to anogenital distance, an androgen-sensitive measure of reproductive development, in infant girls. *Environ. Health. Perspect.* **2017**, *125*, 077008. DOI PubMed PMC

18. Bornehag, C. G.; Carlstedt, F.; Jönsson, B. A.; et al. Prenatal phthalate exposures and anogenital distance in Swedish boys. *Environ. Health. Perspect.* **2015**, *123*, 101-7. DOI PubMed PMC
19. Nelson, W.; Liu, D. Y.; Yang, Y.; Zhong, Z. H.; Wang, Y. X.; Ding, Y. B. In utero exposure to persistent and nonpersistent endocrine-disrupting chemicals and anogenital distance. A systematic review of epidemiological studies†. *Biol. Reprod.* **2020**, *102*, 276-91. DOI PubMed
20. Andreasen, S. M.; Gehrt, L.; Hagen, C. P.; et al. Correlation of anogenital distance from childhood to age 9 years-a prospective population-based birth cohort-the Odense Child Cohort. *Hum. Reprod. Open.* **2024**, *2024*, hoae050. DOI PubMed PMC
21. Zamani, P.; Hemati, Z.; Kelishadi, R.; Kolahdozan, S.; Dianatinasab, M.; Keikha, M. Association between anogenital distance as a noninvasive index in the diagnosis and prognosis of reproductive disorder: a systematic review. *Int. J. Reprod. Biomed.* **2023**, *21*, 599-618. DOI PubMed PMC
22. Barrett, E. S.; Hoeger, K. M.; Sathyanarayana, S.; et al. Anogenital distance in newborn daughters of women with polycystic ovary syndrome indicates fetal testosterone exposure. *J. Dev. Orig. Health. Dis.* **2018**, *9*, 307-14. DOI PubMed PMC
23. Dural, O.; Kurbanova, T.; Yasa, C.; et al. Idiopathic primary ovarian insufficiency is associated with anogenital distance, a marker for prenatal environment. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **2021**, *258*, 304-8. DOI PubMed
24. Kulaksiz, D.; Toprak, T. The relationship between anogenital distance and female sexual life. *Minerva. Obstet. Gynecol.* **2023**, *75*, 311-5. DOI PubMed
25. Celik, M.; Iyigun, I.; Yalcin, S. S.; et al. Perinatal exposure to heavy metals and trace elements of preterm neonates in the NICU: a toxicological study using multiple biomatrices. *Toxics* **2025**, *13*, 898. DOI PubMed PMC
26. Huang, L.; Huang, S.; Luo, X.; et al. Associations of prenatal exposure to multiple metals with testicular volume and anogenital distance in infant boys: a longitudinal cohort study. *Environ. Int.* **2020**, *143*, 105900. DOI PubMed
27. Chen, T.; Guestrin, C. XGBoost: a scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Min*. Association for Computing Machinery; 2016. pp. 785-94. DOI
28. Chipman, H. A.; George, E. I.; McCulloch, R. E. BART: Bayesian additive regression trees. *Ann. Appl. Stat.* **2010**, *4*, 266-98. DOI
29. Bobb, J. F.; Valeri, L.; Claus Henn, B.; et al. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. *Biostatistics* **2015**, *16*, 493-508. DOI PubMed PMC
30. Bobb, J. F.; Claus Henn, B.; Valeri, L.; Coull, B. A. Statistical software for analyzing the health effects of multiple concurrent exposures via Bayesian kernel machine regression. *Environ. Health.* **2018**, *17*, 67. DOI PubMed PMC
31. Xie, Y.; Xiao, H.; Zheng, D.; et al. Associations of prenatal metal exposure with child neurodevelopment and mediation by perturbation of metabolic pathways. *Nat. Commun.* **2025**, *16*, 2089. DOI PubMed PMC
32. Hu, J.; Papandonatos, G. D.; Zheng, T.; et al. Prenatal metal mixture exposure and birth weight: a two-stage analysis in two prospective cohort studies. *Eco. Environ. Health.* **2022**, *1*, 165-71. DOI PubMed PMC
33. Chen, X.; Liu, H.; Zhou, A.; et al. Fetal weight growth trajectories and childhood development: a population-based cohort study. *Sci. Bull.* **2024**, *69*, 3404-14. DOI PubMed
34. Jia, Z.; Zhang, H.; Yu, L.; et al. Prenatal lead exposure, genetic factors, and cognitive developmental delay. *JAMA. Netw. Open.* **2023**, *6*, e2339108. DOI PubMed PMC
35. Sathyanarayana, S.; Grady, R.; Redmon, J. B.; et al.; TIDES Study Team. Anogenital distance and penile width measurements in The Infant Development and the Environment Study (TIDES): methods and predictors. *J. Pediatr. Urol.* **2015**, *11*, 76.e1-6. DOI PubMed PMC
36. Olapeju, B.; Hong, X.; Wang, G.; et al. Birth outcomes across the spectrum of maternal age: dissecting aging effect versus confounding by social and medical determinants. *BMC. Pregnancy. Childbirth.* **2021**, *21*, 594. DOI PubMed PMC
37. Blomberg, M.; Birch Tyrberg, R.; Kjølhede, P. Impact of maternal age on obstetric and neonatal outcome with emphasis on primiparous adolescents and older women: a Swedish Medical Birth Register Study. *BMJ. Open.* **2014**, *4*, e005840. DOI PubMed PMC
38. Vandekerckhove, M.; Guignard, M.; Civadier, M. S.; Benachi, A.; Bouyer, J. Impact of maternal age on obstetric and neonatal morbidity: a retrospective cohort study. *BMC. Pregnancy. Childbirth.* **2021**, *21*, 732. DOI PubMed PMC
39. Wilcox, A. J.; Weinberg, C. R.; Basso, O. On the pitfalls of adjusting for gestational age at birth. *Am. J. Epidemiol.* **2011**, *174*, 1062-8. DOI PubMed PMC
40. Liu, X.; Du, J.; Wang, G.; Chen, Z.; Wang, W.; Xi, Q. Effect of pre-pregnancy body mass index on adverse pregnancy outcome in north of China. *Arch. Gynecol. Obstet.* **2011**, *283*, 65-70. DOI PubMed
41. Li, S.; Li, J.; Du, L.; et al. Association between parity and adverse maternal and neonatal outcomes: a population-based cross-sectional study. *Front. Med.* **2025**, *12*, 1697655. DOI PubMed PMC
42. Li, L.; Wu, Y.; Yang, Y.; Wu, Y.; Zhuang, Y.; You, D. Maternal educational inequalities about adverse pregnancy outcomes observed in a rural area of a province of China during a time period (2010-2018). *J. Epidemiol. Community. Health.* **2022**, *76*, 458-65. DOI PubMed

43. Yen, F. S.; Wei, J. C.; Wu, Y. L.; et al. Impact of family income on the development of gestational diabetes mellitus and the associated birth outcomes: a nationwide study. *J. Diabetes. Investig.* **2025**, *16*, 51-9. DOI PubMed PMC
44. Barrera-Gómez, J.; Basagaña, X. Models with transformed variables: interpretation and software. *Epidemiology* **2015**, *26*, e16-7. DOI PubMed
45. Linden, A. A comparison of approaches for stratifying on the propensity score to reduce bias. *J. Eval. Clin. Pract.* **2017**, *23*, 690-6. DOI PubMed
46. Thomas, L. E.; Thomas, S. M.; Li, F.; Matsouaka, R. A. Addressing substantial covariate imbalance with propensity score stratification and balancing weights: connections and recommendations. *Epidemiol. Methods.* **2023**, *12*, 20220131. DOI PubMed PMC
47. Li, F.; Thomas, L. E.; Li, F. Addressing extreme propensity scores via the overlap weights. *Am. J. Epidemiol.* **2019**, *188*, 250-7. DOI PubMed
48. Wen, J.; Yao, X.; Lu, R.; et al. Associations between seminal plasma metals/metalloids and semen quality: a cross-sectional study. *Ecotoxicol. Environ. Saf.* **2024**, *279*, 116472. DOI PubMed
49. Chedrese, P. J.; Piasek, M.; Henson, M. C. Cadmium as an endocrine disruptor in the reproductive system. *Immunol. Endocr. Metab. Agents. Med. Chem.* **2006**, *6*, 27-35. DOI
50. Alamdar, A.; Tian, M.; Huang, Q.; et al. Enhanced histone H3K9 tri-methylation suppresses steroidogenesis in rat testis chronically exposed to arsenic. *Ecotoxicol. Environ. Saf.* **2019**, *170*, 513-20. DOI PubMed
51. Branch, F. M.; Perry, M. J.; Chen, Z.; Louis, G. M. B. Metal(loid)s and human semen quality: the LIFE Study. *Reprod. Toxicol.* **2021**, *106*, 94-102. DOI PubMed PMC
52. Culty, M.; Albert, O. Antiandrogens and male reproduction. *Encyclopedia. of. Reproduction.* , 2026. pp. 813-22. DOI
53. Chaturvedi, N. K.; Kumar, S.; Negi, S.; Tyagi, R. K. Endocrine disruptors provoke differential modulatory responses on androgen receptor and pregnane and xenobiotic receptor: potential implications in metabolic disorders. *Mol. Cell. Biochem.* **2010**, *345*, 291-308. DOI PubMed
54. Li, H.; Liu, J.; Huang, L.; et al. Hexavalent chromium inhibits testosterone synthesis in bovine testicular leydig cells through mitochondrial damage via BNIP3. *Reprod. Biol.* **2026**, *26*, 101181. DOI PubMed
55. Shobana, N.; Kumar, M. K.; Navin, A. K.; Akbarsha, M. A.; Aruldhas, M. M. Prenatal exposure to excess chromium attenuates transcription factors regulating expression of androgen and follicle stimulating hormone receptors in Sertoli cells of prepuberal rats. *Chem. Biol. Interact.* **2020**, *328*, 109188. DOI PubMed
56. Kong, L.; Tang, M.; Zhang, T.; et al. Nickel nanoparticles exposure and reproductive toxicity in healthy adult rats. *Int. J. Mol. Sci.* **2014**, *15*, 21253-69. DOI PubMed PMC
57. Ashrap, P.; Sánchez, B. N.; Téllez-Rojo, M. M.; et al. In utero and peripubertal metals exposure in relation to reproductive hormones and sexual maturation and progression among girls in Mexico City. *Environ. Res.* **2019**, *177*, 108630. DOI PubMed PMC
58. Hill, J.; Linero, A.; Murray, J. Bayesian additive regression trees: a review and look forward. *Ann. Rev. Stat. Appl.* **2020**, *7*, 251-78. DOI
59. Kou, M.; Wang, Z.; Wang, M.; et al. Cadmium and lead ions interaction: a model for assessing combined heavy metal contamination in soils for food safety risk evaluation. *Ecotoxicol. Environ. Saf.* **2025**, *298*, 118324. DOI PubMed
60. Obeng-Gyasi, E.; Obeng-Gyasi, B. Association of combined lead, cadmium, and mercury with systemic inflammation. *Front. Public. Health.* **2024**, *12*, 1385500. DOI PubMed PMC
61. Schisterman, E. F.; Cole, S. R.; Platt, R. W. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology* **2009**, *20*, 488-95. DOI PubMed PMC
62. Thomas, L. E.; Li, F.; Pencina, M. J. Overlap weighting: a propensity score method that mimics attributes of a randomized clinical trial. *JAMA* **2020**, *323*, 2417-8. DOI PubMed
63. Rosen, E. M.; Stevens, D. R.; McNell, E. E.; et al. Variability and longitudinal trajectories of phthalate and replacement biomarkers across pregnancy in the human placenta and phthalates study. *Environ. Sci. Technol.* **2023**, *57*, 13036-46. DOI PubMed PMC
64. Bartell, S. M.; Purdue, M. P.; Rhee, J.; Nøst, T. H.; Rusiecki, J.; Steenland, K. How well does a single blood sample represent long-term exposure for epidemiological studies of PFOA among men in the general population? *Environ. Int.* **2024**, *192*, 109056. DOI PubMed
65. Thankamony, A.; Pasterski, V.; Ong, K. K.; Acerini, C. L.; Hughes, I. A. Anogenital distance as a marker of androgen exposure in humans. *Andrology* **2016**, *4*, 616-25. DOI PubMed PMC
66. Cortes, D.; Fischer, M. B.; Hildorf, A. E.; et al. Anogenital distance in a cohort of 169 infant boys with uni- or bilateral cryptorchidism including 18 boys with vanishing testes. *Hum. Reprod.* **2024**, *39*, 689-97. DOI PubMed PMC