scientific reports

Check for updates

OPEN

Association between urinary cadmium levels and increased gallstone disease in US adults

Zhaowei Wu[1,2], Shiming Jiang[1,2], Jinzhi Li[1], Panguo Wang[1] & Yong Chen[1]✉

Heavy metal exposure is acknowledged as a risk factor for poor health. However, the effect of heavy metal exposure on the prevalence of gallstones is still unknown. Therefore, we investigated the relationship between heavy metal concentrations and the prevalence of gallstones among US adults. Multivariate logistic regression indicated that only urinary cadmium was an independent risk factor for gallstones. Compared to the low urine cadmium group, the high cadmium group had a elevated increased risk of gallstone formation. Furthermore, the weighted quantile sum model showed that heavy metal mixtures were not associated with gallstone prevalence. Additionally, urinary cadmium levels were associated with an increased risk of gallstone formation in young individuals, males, Mexican Americans, Non-Hispanic Whites, as well as smokers and drinkers. Moreover, nine machine learning methods were utilized to construct an interpretable predictive model for gallstone prevalence. Among these models, the XGBoost model exhibited the highest performance and was selected for further investigation. Subsequently, shapely additive explanations was used for model interpretation. The results also indicated that urinary cadmium concentrations were the most important variable for gallstones. Thus, our results indicated that long-term chronic cadmium exposure is a risk factor for gallstone prevalence.

Keywords Cadmium, Gallstone, Restricted cubic spline curves, Random forest, Machine learning

Gallstones affect up to 20% of the European population[1]and 10–15% of the US population[2], resulting in high healthcare costs globally. Gallstone disease is a chronic condition, and the risk of gallstone formation increases with age and is greater in females than in males[3]. Dyslipidemia is common among the elderly, with abnormal levels of triglycerides and high-density lipoprotein (HDL) being recognized as risk factors for gallstone formation. High triglyceride levels reduce gallbladder contraction, while HDL promotes the secretion of cholesterol into bile ducts, decreasing the cholesterol saturation in bile. Lower HDL levels hinder the synthesis of hepatic bile acids, which in turn promotes gallstone formation[4,5]. Moreover, Both hyperglycemia and normoglycemic hyperinsulinemia can impair the function of cholecystokinin, leading to decreased gallbladder motility. Therefore, diabetes, which is common among the elderly, is considered a contributing factor to the development of gallstones[6]. The gallstone rate varies widely by race, with a lower rate in African populations and a higher rate in Central and South American Hispanic populations and American-Hispanic populations[7–10]. Multiple studies have indicated that metabolic abnormalities, including obesity[11], hyperinsulinemia[12], and insulin resistance[13], contribute to gallstone formation.

The management of gallstone disease is of continuous interest due to its high global prevalence and associated healthcare costs. Several preventive strategies have been adopted, including regular physical activity, high-fiber and high-calcium diets, and regular abdominal ultrasonography[14]. However, it is essential to identify additional risk factors for gallstones. Heavy metals are typically characterized as metallic elements with a density equal to or greater than 5 g/cm^3[15]. Humans often encounter heavy metals through a variety of sources, including air pollution, household waste, cosmetics, and diet. In recent years, there has been extensive discourse surrounding heavy metal exposure. Several studies have indicated a link between cardiovascular, immune, and digestive system disorders and heavy metal exposure[16–18]. It has been reported that exposure to metal pollution from both natural and manufactured sources is harmful to the human body. For example, cadmium (Cd) is toxic to humans and is found in most human foodstuffs and foods because of its high rates of soil-to-plant transfer[19]. It was reported that exposure to Cd may cause Itai-Itai disease, diabetes, nephropathy, hypertension and cancer[20].

[1]Hepatobiliary Surgery, The First Affiliated Hospital of Chongqing Medical University, Medical College Street, Yuzhong District, Chongqing 404100, China. [2]Zhaowei Wu and Shiming Jiang contributed equally to this work. ✉email: Chenyong61@126.com

However, the impact of heavy metal exposure on gallstone formation remains unclear. Therefore, in this study, we aimed to identify the relationships between the levels of heavy metals in urine and blood and the prevalence of gallstones among National Health and Nutrition Examination Survey (NHANES) 2017–2020 participants.

Material and methods

Population

Data were extracted from the NHANES database from 2017 to 2020, a nationally representative, ongoing, publicly accessible multipurpose survey initiated by the U.S. Centers for Disease Control and Prevention (CDC) to evaluate the health status of the American population. At the outset, the study included 9232 participants aged 20 years and older. Subsequently, those participants were excluded if they were pregnant or if they lacked data on gallstones, blood, or urinary heavy metals. Finally, 2106 participants were selected.

Definition of gallstones and measurement of blood and urinary metal concentrations

The question "Has a doctor ever diagnosed you with gallstones?" was utilized in the administered questionnaire examination to define the presence of gallstones. An exact “yes” response was considered to indicate the presence of a gallstone. Whole-blood and urine samples were stored at a temperature of − 30 °C before being transported to the National Center for Environmental Health for analysis by inductively coupled plasma kinetic reaction cell mass spectrometry, including barium (Ba), Cd, cobalt (Co), cesium (Cs), molybdenum (Mo), manganese (Mn), lead (Pb), antimony (Sb), tin (Sn), thallium (Tl), tungsten (W), uranium (U), nickel (Ni), and mercury (Hg) in urine and Cd, Pb, selenium (Se), Mn and Hg in blood. To ensure result stability, measurement values below the limit of detection (LOD) were excluded. The urinary metal concentrations were adjusted for urine creatinine levels.

Covariates

The following meaningful covariates were selected: age, sex (male/female), race (Mexican American, Non-Hispanic Black, Non-Hispanic White, Other Hispanic, and other races), education level (less than 9 years, 9–11 years, high school graduate/GED, college/AA, and college graduate), family poverty–income ratio (PIR), body mass index (BMI), alcohol consumption status, cigarette smoking status, diabetes status (yes, no, borderline), and hypertension status. BMI was defined as the ratio of body weight in kilograms to the square of height in meters (kg/m^2). The PIR was calculated as the ratio of annual earnings to the poverty threshold, adjusted for family size. Alcohol consumption status was evaluated based on participants' responses to a single-choice question in the questionnaire: "Have you ever consumed any type of alcohol?" Cigarette smoking status was established using the criterion of having smoked at least 100 cigarettes in one's lifetime. Diabetes and hypertension status were identified through self-reported diagnoses received from healthcare professionals.

Statistical analyses

A summary descriptive analysis was carried out on the baseline characteristics of both the gallstone and control groups. The Kruskal–Wallis rank sum test was employed for continuous variables, while Fisher's exact probability test was utilized for count variables.

Initially, two multivariate logistic regression analyses were conducted to investigate the individual effects of blood and urinary heavy metals on gallstone prevalence, adjusting for covariates including age, sex, race, education level, family PIR, BMI, alcohol consumption status, cigarette smoking status, diabetes status, hypertension status, and other heavy metals. The best cutoff point was calculated by pROC and shapely R package. Then, all patients were divided into low urine cadmium group and high urine cadmium group. Multivariate logistic regression analysis was conducted to assess the associations between heavy metal concentrations and gallstone prevalence with no adjustment for any covariates. Furthermore, multivariate logistic regression analysis was conducted adjusting for all covariates, including age, sex, race, education level, family PIR, BMI, alcohol consumption, cigarette smoking status, and diabetes status. Subgroup analysis was subsequently carried out to investigate whether the relationships between urinary concentrations of heavy metals and the prevalence of gallstones were affected by age, sex, race, alcohol consumption status and cigarette smoking status. In addition, restricted cubic spline curves with four knots at the 5th, 35th, 65th, and 95th percentiles were constructed to investigate the relationships between blood and urinary concentrations of heavy metals and the prevalence of gallstones after controlling for confounding factors. In addition, weighted quantile sum (WQS) analysis was conducted to investigate the combined effect of heavy metal mixture concentrations on gallstone prevalence by calculating a weighted linear index in both the positive and negative directions.

All the statistical analyses were conducted using R software (4.3.1). $P < 0.05$ was considered to indicate statistical significance.

Construction of the machine learning model

In this study, the dataset was partitioned into a training set and a testing set, with the former comprising 70% (N = 1685) and the latter comprising 30% (N = 421). Multivariate logistic regression analysis was employed to investigate the relationships between the risk of gallstone formation and blood and urinary heavy metal concentrations while adjusting for covariates. Then, nine machine learning methods (logistics regression, gradient boost machine (GBM), random forest (RF), ridge, lasso, decision tree, naïve Bayes, XGBoost, and support vector machine (SVM) were used to construct predictive models using the training set. Furthermore, the details of each algorithm are shown in Table S1. The performances of the models were analyzed using receiver operating characteristic (ROC) curve and area under the receiver operator curve (AUC) analyses in the training and testing cohorts. Among the nine machine learning models considered, the XGBoost algorithm demonstrated the most superior predictive performance and was selected for further investigation.

Subsequently, Shapley additive explanations (SHAPs) was used for model interpretation following the selection of the XGBoost model by the *shapviz* R packages. SHAP, a game-theoretic approach, was employed to illustrate the decision-making process of the model in predicting the prevalence of gallstones by calculating the SHAP values of variables to compare the difference between the actual prediction and the mean prediction in the XGBoost model. The contribution weight of urinary Cd, age, sex, race, BMI and diabetes status were also analyzed.

Results

Population baseline

The demographic features of the individuals from 2017 to 2020. A total of 2106 individuals were extracted, and gallstones were identified in 195 individuals, representing 9.26% of the study population. Sex, age, race, BMI, diabetes status, and hypertension status were obviously different between the gallstone group and the control group (Table 1, all $P<0.05$). Compared to individuals without gallstones, those in the gallstone group had significantly higher levels of urinary cadmium and blood mercury.

Heavy metal exposure and gallstone formation risk in the logistic regression model

Multivariate regression models were built to assess the associations of blood and urinary concentrations of heavy metals with the prevalence of gallstones. Table 2 shows that only the urinary Cd concentration was significantly associated with gallstone prevalence (Model 1: OR 1.08, 95% CI 1.05–1.12, $P<0.001$; Model 2: OR 1.04, 95% CI 1.01–1.08, $P=0.02$). However, urinary Ba, Cs, Mo, Mn, Pb, Sb, Sn, Tl, W, Ni, and Hg concentrations and blood Cd, Pb, Se, Mn and Hg concentrations were not significantly associated with the risk of gallstone formation (all $P>0.05$). Furthermore, the urinary Cd concentration was converted into high and low groups. Compared to the low urine cadmium group, the high cadmium group had a 1.3-fold increased risk of gallstone formation in Model 1 (95% CI 1.06–1.20, $P<0.05$) and a 1.09-fold increased risk in Model 2 (95% CI 1.02–1.17, $P<0.05$), highlighting a significant association between urinary Cd levels and the risk of gallstone formation (Table 3). It was reported that multicollinearity existed among the variables as the variance inflation factor (VIF) exceeded 10[21]. However, in this study, VIF analysis revealed no evidence of multicollinearity among any of the heavy metals or covariates (Table S1).

Heavy metal exposure and gallstone formation risk in the WQS model

A WQS regression model was constructed to investigate the relationships of the mixtures of the concentrations of thirteen heavy metals in urine and five heavy metals in blood with the prevalence of gallstones. Model 1 was not adjusted for any covariates, and Model 2 was adjusted for covariates. The results indicated that the WQS regression in both the positive and negative directions did not demonstrate significant associations between the urinary and blood heavy metal mixtures, together or alone, and the prevalence of gallstones (all $P>0.05$) (Table 4).

Subgroup analysis and dose–response relationship analysis

Additionally, subgroup analysis was carried out to investigate whether the relationships between urinary Cd concentrations and gallstone prevalence differed according to age, sex, race, BMI, alcohol consumption status and cigarette smoking status. After controlling for confounding factors, the level of urinary Cd was associated with a greater risk of gallstone formation in individuals under 65 years of age, males, Mexican Americans, Non-Hispanic Whites, smokers and drinkers (Table 5).

Restricted cubic spline curves were then utilized to illustrate the relationship between the concentration of urinary Cd and the risk of gallstone formation. A significant positive linear relationship was observed between the concentration of urinary Cd and the prevalence of gallstones after adjusting for covariables, including age, sex, race, BMI, and diabetes status (P for nonlinearity = 0.208) (Fig. 1B). Moreover, no significant relationship was found between the prevalence of gallstones and other heavy metals in the urine or blood (Fig. 1).

Evaluation of the machine learning predictive models

According to the multivariable logistic regression model, urinary Cd, age, sex, race, BMI and diabetes status were identified as being more relevant to gallstone prevalence and were selected for the construction of the predictive models (Table S2). The ROC curves of the nine models in the training and testing sets are shown in Fig. 2. The average AUC values for the logistic regression, gradient boost machine, lasso, decision tree, naïve Bayes, random forest, ridge, support vector machine, and XGBoost models in the training and testing sets were 0.706, 0.728, 0.7015, 0.6325, 0.706, 0.7235, 0.705, 0.726, and 0.749, respectively. Among the nine machine learning models, the XGBoost algorithm demonstrated the most superior predictive performance. Consequently, a predictive model based on the XGBoost algorithm was selected for further analyses. The SHAP person waterfall plot for the XGBoost model was constructed to illustrate the decision-making process for gallstone formation risk by calculating the SHAP values of variables in the XGBoost model. The yellow bar indicates that the variable led to gallstone formation, whereas the brown bar indicates its inhibitory effect. As shown in Fig. 3A, individuals with a urinary cadmium level of 9.01 μg/g, aged 69, female, Mexican American, a BMI of 39.4, and diabetes had positive SHAP values, indicating that these factors were associated with an increased risk of gallstone prevalence. In contrast, individuals with a urinary cadmium level of 0.0629 μg/g, aged 30, male, of other races, a BMI of 27.1, and without diabetes had negative SHAP values, suggesting a protective effect against gallstones (Fig. 3B). The sum of the SHAP values was then calculated to predict the risk of gallstone formation. Since the patient's sum SHAP value (− 2.15) exceeded the SHAP baseline value (− 2.16), it contributed to an increased risk of developing gallstones, and vice versa. Moreover, elevated levels of urinary Cd were associated with increased SHAP values,

Characteristics	Control *N = 1911*	Gallstone *N = 195*	P value
Sex			**< 0.001**
Female	930 (48.7%)	139 (71.3%)	
Male	981 (51.3%)	56 (28.7%)	
Age	50.8	57.9	**< 0.001**
Race			**0.009**
Mexican American	218 (11.4%)	26 (13.3%)	
Non-Hispanic Black	511 (26.7%)	34 (17.4%)	
Non-Hispanic White	682 (35.7%)	89 (45.6%)	
Other Hispanic	177 (9.26%)	21 (10.8%)	
Other Race	323 (16.9%)	25 (12.8%)	
Education level			0.719
Less than 9th grade	636 (33.3%)	67 (34.4%)	
9th–11th grade	102 (5.34%)	24 (12.3%)	
High school graduate/GED or equivalent	495 (25.9%)	55 (28.2%)	
Some college or AA degree	468 (24.5%)	41 (21.0%)	
College graduate or above	102 (5.34%)	8 (4.10%)	
PIR	2.64 (1.63)	2.63 (1.51)	0.976
BMI	29.8	32.7	**< 0.001**
Alcohol consumption status			0.907
No	157 (8.22%)	15 (7.69%)	
Yes	1754 (91.8%)	180 (92.3%)	
Cigarette smoking status			0.666
No	1075 (56.3%)	106 (54.4%)	
Yes	836 (43.7%)	89 (45.6%)	
Diabetes status			**< 0.001**
Borderline	65 (3.40%)	6 (3.08%)	
No	1571 (82.2%)	139 (71.3%)	
Yes	275 (14.4%)	50 (25.6%)	
Hypertension status			**0.002**
No	1179 (61.7%)	98 (50.3%)	
Yes	732 (38.3%)	97 (49.7%)	
Urine Barium (ug/g)	1.63	1.76	0.413
Urine Cadmium (ug/g)	0.32	0.44	**0.019**
Urine Cobalt (ug/g)	0.50	0.55	0.504
Urine Cesium (ug/g)	4.95	5.12	0.427
Urine Molybdenum (ug/g)	41.2	40.2	0.614
Urine Manganese (ug/g)	0.19	0.19	0.992
Urine Lead (ug/g)	0.42	0.40	0.378
Urine Antimony (ug/g)	0.00	0.00	0.294
Urine Tin (ug/g)	0.98	1.26	0.089
Urine Thallium (ug/g)	0.19	0.19	0.982
Urine Tungsten (ug/g)	0.09	0.09	0.502
Urine Nickel (ug/g)	1.63	1.70	0.650
Urine Mercury (ug/g)	0.38	0.37	0.966
Blood Lead (ug/dL)	1.21	1.15	0.427
Blood Cadmium (ug/L)	0.49	0.49	0.966
Blood Mercury, total (ug/L)	1.41	1.09	0.007
Blood Selenium (ug/L)	9.77	10.1	0.175
Blood Manganese (ug/L)	1.21	1.15	0.427

Table 1. Baseline characteristics of the patients in the model. The Kruskal–Wallis rank sum test was employed for continuous variables, while Fisher's exact probability test was utilized for count variables. BMI, body mass index; PIR, family poverty–income ratio. GED, General Educational Development. The urinary heavy metal concentrations were adjusted for urine creatinine levels. Significant values are in bold.

	Model 1		Model 2	
	OR (95% CI)	P value	OR (95% CI)	*P* value
In urine				
Ba	1.00 (1.00–1.01)	0.523	1.00 (1.00–1.00)	0.813
Cd	**1.08 (1.05–1.12)**	**< 0.001**	**1.04 (1.01–1.08)**	**0.02**
Co	1.00 (0.99–1.01)	0.548	1.00 (0.99–1.01)	0.664
Cs	1.00 (1.00–1.01)	0.463	1.00 (0.99–1.00)	0.067
Mo	1.00 (1.00–1.00)	0.684	1.00 (1.00–1.00)	0.148
Mn	1.00 (0.97–1.04)	0.995	0.99 (0.96–1.03)	0.686
Pb	0.99 (0.96–1.02)	0.506	0.97 (0.94–1.01)	0.113
Sb	0.18 (0.00–94.7)	0.593	0.50 (0.00–231)	0.825
Ti	1.01 (1.00–1.01)	0.064	1.00 (1.00–1.01)	0.719
Tl	1.00 (0.92–1.09)	0.981	0.95 (0.87–1.04)	0.303
W	0.99 (0.93–1.05)	0.683	0.98 (0.92–1.04)	0.492
Ni	1.00 (0.99–1.01)	0.638	1.00 (0.99–1.01)	0.874
Hg	1.00 (0.98–1.02)	0.963	1.00 (0.98–1.02)	0.934
In blood				
Pb	1.04 (0.88–1.23)	0.637	1.04 (0.88–1.23)	0.637
Cd	0.96 (0.72–1.27)	0.754	0.96 (0.72–1.27)	0.754
Hg	0.95 (0.86–1.05)	0.300	0.95 (0.86–1.05)	0.300
Se	1.00 (1.00–1.01)	0.666	1.00 (1.00–1.01)	0.666
Mn	1.00 (0.96–1.04)	0.923	1.00 (0.96–1.04)	0.923

Table 2. Multivariable analysis model of the association between urinary and blood heavy metal concentrations and the prevalence of gallstones. Model 1 was no adjustment for any covariates. Model 2 was adjusted for all covariates adjusted for all covariates, including age, sex, race, education level, family PIR, BMI, alcohol consumption, cigarette smoking status, and diabetes status.. Ba, Barium; Cd, Cadmium; Cs, Cesium; Mo, Molybdenum; Mn, Manganese; Pb, Lead; Sb, Antimony; Ti, Tin; Tl, Thallium; W, Tungsten; Ni, Nickel; Hg, Mercury; Se, Selenium. OR: odds ratio; CI: confidence interval. Significant values are in bold.

Characteristics	Model 1 HR (95%CI)	*P*-value	Model 2 HR (95%CI)	*P*-value
Urinary Cd (ug/g)				
Low	Reference	**< 0.001**	Reference	**0.008**
High	1.13 (1.06–1.20)		1.09 (1.02–1.17)	

Table 3. ORs (95% CIs) for gallstone formation associated with a urinary Cd concentration. Model 1 was no adjustment for any covariates. Model 2 was adjusted for all covariates, including age, sex, race, education level, family PIR, BMI, alcohol consumption, cigarette smoking status, and diabetes status. OR: odds ratio; CI: confidence interval. Low: < 0.98 ug/g, High: ≥ 0.98 ug/g. Significant values are in bold.

which in turn were associated with an increased risk of gallstone formation (Fig. 3C, D). The results also indicated that urinary Cd concentrations were the most important variable for the predictive model (Fig. 3E).

Discussion

Gallstones affect up to 20% of the European population[1] and 10–15% of the US population[2], resulting in high healthcare costs globally. Although many gallstones are asymptomatic, severe complications occur in more than 40% of patients with gallstones[22]. Unfortunately, the cause of gallstones is not fully understood. The phrase "family, female, forty, fat, and fertile" summarizes the major risk factors for the development of gallstones. Our study also confirmed that older individuals, females, and those with obesity are at a higher risk of developing gallstones. However, the above factors are difficult to change and lack clear guiding significance for decreasing the prevalence of gallstones. Finding simple and measurable indicators to evaluate the risk of gallstone formation is crucial. In recent years, there has been extensive discussion on the relationship between heavy metal exposure and diseases. Studies have reported that exposure to metal pollution from both natural and man-made sources can be harmful to the human body. However, the impact of heavy metal exposure on gallstone formation remains unclear.

In this study, 13 heavy metals in urine and 5 heavy metals in blood were analyzed via multivariate regression and WQS regression analysis. This indicated that only urinary Cd concentrations were confirmed to be an independent risk factor for gallstones, and neither the presence of any of the above heavy metals (except urine

Heavy metals	Model 1		Model 2	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Urine + Blood				
gWQS (forward)	1.01 (1.00–1.02)	0.056	0.99 (0.98–1.01)	0.253
gWQS (backward)	1.00 (0.99–1.01)	0.984	1.00 (0.97–1.00)	0.151
Urine				
gWQS (forward)	1.01 (1.00–1.02)	0.079	0.99 (0.97–1.00)	0.108
gWQS (backward)	1.00 (0.99–1.01)	0.490	0.98 (0.97–1.00)	0.024
Blood				
gWQS (forward)	1.01 (1.00–1.02)	0.115	1.00 (0.99–1.01)	0.441
gWQS (backward)	1.00 (0.99–1.01)	0.902	1.00 (0.99–1.01)	0.900

Table 4. WQS model between the heavy metal mixture and the prevalence of gallstones. Model 1 was not adjusted for any covariates. Model 2 was adjusted for the covariates. WQS: weighted quantile sum.

Characteristic	OR (95% CI)	*P* value
Sex		
Male	**1.07 (1.01–1.13)**	**0.03**
Female	1.03 (0.98–2.08)	0.237
Age		
< 65	1.00 (0.95–1.05)	0.973
≥ 65	**1.08 (1.02–1.15)**	**0.011**
Race		
Mexican American	**1.38 (1.14–1.68)**	**< 0.001**
Other Hispanic	1.20 (0.99–1.46)	0.064
Non-Hispanic White	**1.07 (1.02–1.44)**	**0.008**
Non-Hispanic Black	0.94 (0.85–1.03)	0.176
Other Race	**0.91 (0.84–0.99)**	**0.031**
Alcohol		
Yes	**1.05 (1.01–1.10)**	**0.011**
No	0.95 (0.84–1.08)	0.424
Cigarette smoking status		
Yes	**1.06 (1.00–1.13)**	**0.064**
No	1.04 (0.99–1.09)	0.113

Table 5. Subgroup analysis. A multivariable analysis model was constructed stratified by selected factors adjusted for age, sex, race, education level, family poverty–income ratio (PIR), BMI, alcohol consumption status, cigarette smoking status, diabetes status, and hypertension status. Significant values are in bold.

Cd) nor the presence of any of the heavy metals were associated with the prevalence of gallstones. The relationship between different urinary concentrations of Cd and the prevalence of gallstones was further analyzed. Compared to the low urine cadmium group, the high cadmium group had a 1.3-fold increased risk of gallstone formation in Model 1 and a 1.09-fold increased risk in Model 2, highlighting a significant association between urinary Cd levels and the risk of gallstone formation. The restricted cubic spline analysis indicated a significant linear correlation between the risk of gallstone formation and the urinary Cd concentration. Moreover, the risk of gallstone formation increases with increasing urinary Cd concentrations. Thus, urinary Cd concentrations were confirmed to be an independent risk factor for gallstones. As the urinary Cd concentration increases, the risk of developing gallstones also increases. Furthermore, neither the mixtures nor the individual heavy metals mentioned above (except urine Cd) were associated with the prevalence of gallstones.

Cd, a nonessential heavy metal and environmental contaminant, is toxic to humans and is found in most human foodstuffs because of its high rates of soil-to-plant transfer[19]. It has been reported that the general population is exposed to Cd through the food supply[23]. Exposure to Cd can lead to the development of cancer[24], cardiovascular disease[25]and renal toxicity[26]. Elevated levels of blood Cd were positively associated with gallbladder cancer[27]. In contrast, Jia Wang et al. reported no significant association between blood Cd levels and gallstone prevalence[28], which is consistent with our findings. The literature has reported that high urinary Cd levels are associated with hypertension status[29]. To our knowledge, this is the first study to elucidate the relationship between urinary Cd and gallstone prevalence. We speculated that the reason why the impact of urinary Cd differs from that of blood Cd on gallstone formation lies in the fact that urinary Cd more accurately reflects long-term exposure to Cd, while blood Cd only reflects recent exposure to Cd[20].

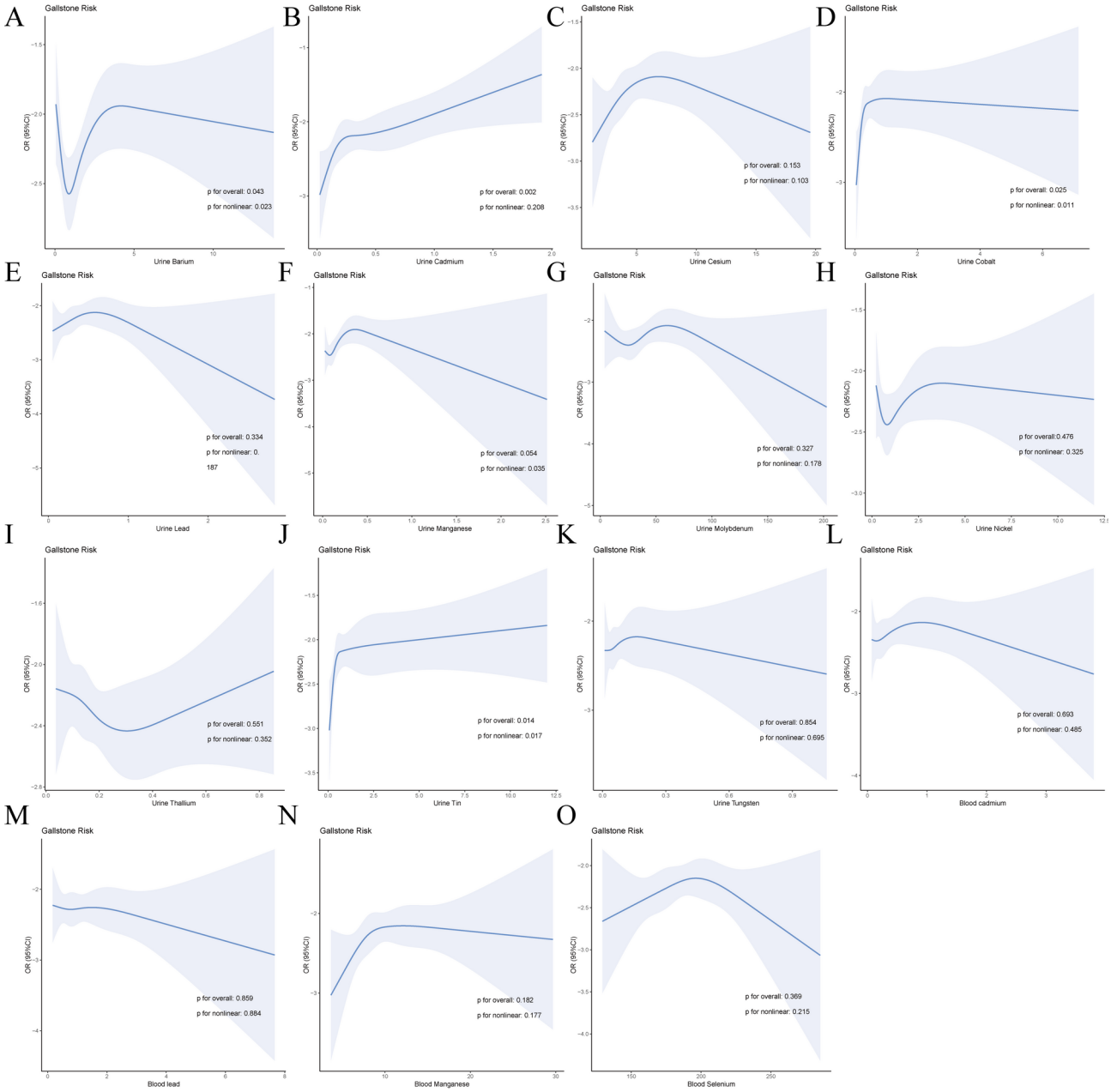

Fig. 1. Dose–response relationship analysis between urinary and blood heavy metal concentrations and the risk of gallstone formation based on restricted cubic spline curves for urinary Ba, Ca, Cs, Mo, Mn, Pb, Sb, Sn, Tl, W, Ni, and Hg concentrations and blood Cd, Pb, Se, Mn and Hg concentrations. A significant positive linear relationship was observed between the level of urinary Cd and the prevalence of gallstones after adjusting for covariables.

Subgroup analysis indicated that urinary Cd concentrations were associated with a greater risk of gallstone formation in individuals younger individual, males, Mexican Americans, Non-Hispanic Whites, smokers, and drinkers. This may be due to inconsistent concentrations of urinary Cd among different subgroups. Satarug et al.[30]demonstrated that renal Cd accumulation was greater in younger age groups than in older age groups. This could be attributed to the greater dietary intake of Cd among younger individuals[31,32]. Moreover, the concentration of Cd in smokers was greater than that in nonsmokers[33]. The use of Cd pigments in the enamels of decorated drinking glasses has also been identified as a potential source of Cd exposure[34]. In addition, the proportion of males who smoke and drink alcohol is greater than that of females. Thus, urinary Cd concentrations were significantly associated with a greater risk of gallstone formation in males. These findings may help explain why younger individual, males, individuals with a BMI > 35, smokers, and drinkers are associated with a greater risk of gallstone formation. However, more prospective studies are necessary to elucidate the role of Mexican-

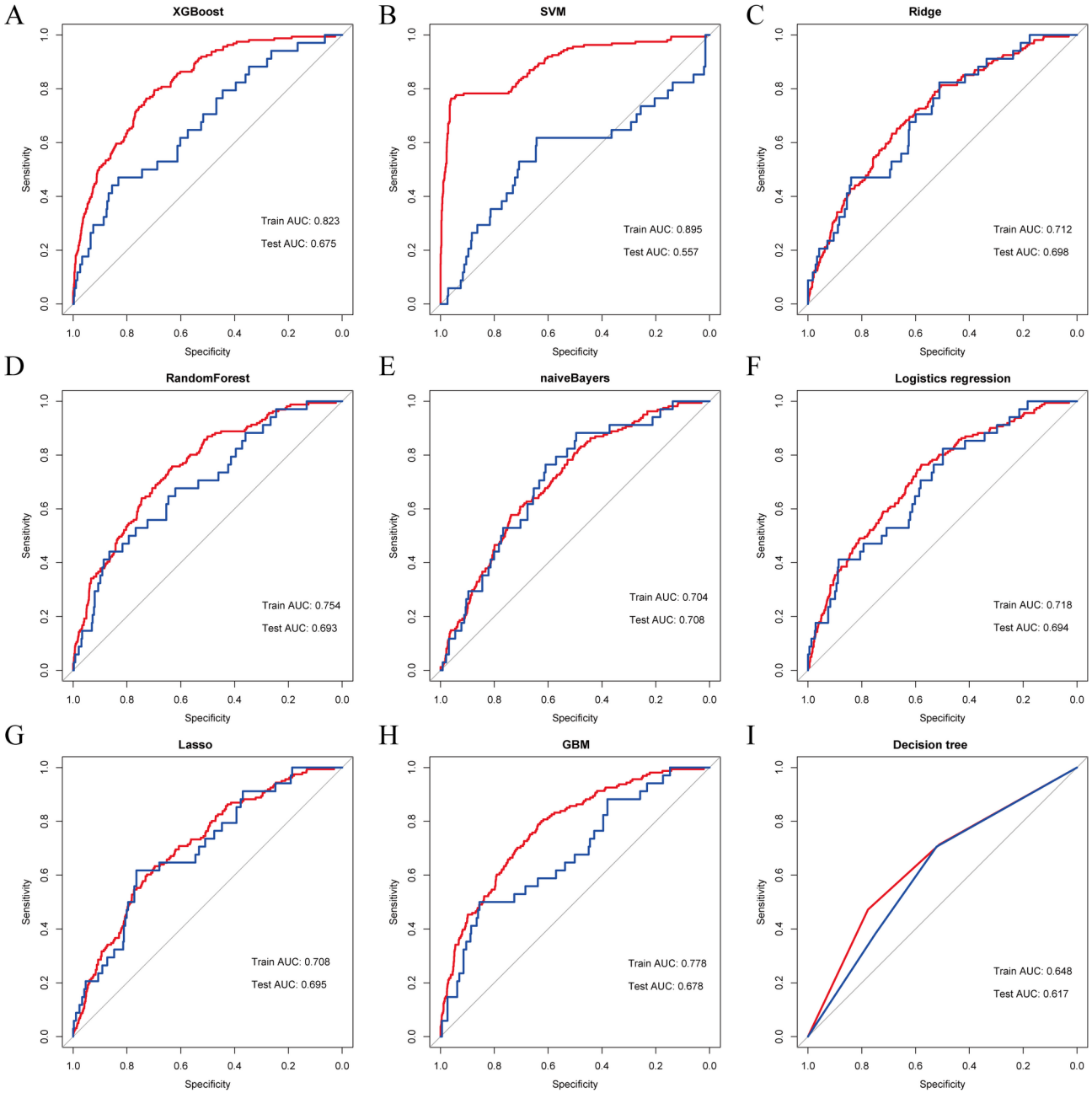

Fig. 2. The ROC curves of the nine machine learning models in the training and testing sets. All models were constructed by urinary Cd, age, sex, race, BMI, and diabetes status. (**A**) XGBoost. (**B**) SVM. (**C**) Ridge. (**D**) Random forest. (**E**) Naive Bayes. (**F**) Logistics regression. (**G**) Lasso. (**H**) GBM. (**I**) Decision tree.

Americans and Non-Hispanic White individuals in the increased risk of gallstone formation associated with urinary Cd.

Moreover, machine learning algorithms combined with an interpretable pipeline were utilized to investigate the potential predictive capacity of heavy metal concentrations for gallstones. Among the nine machine learning models considered, the XGBoost algorithm demonstrated the best predictive performance. Rapidly developed machine learning interpretability methods have been used to enhance the interpretability of black-box models. The PDP approach has been remarkably successful in revealing underlying functional relationships and identifying nonlinear associations between variables and outcomes[35]. A higher urinary Cd concentration was correlated with a stronger ability to predict gallstone formation. The results also indicated that urinary Cd concentrations were the most important variable for the predictive model. Additionally, SHAP was employed to illustrate the decision-making process for gallstone formation risk, revealing that the predictive model assigns greater importance to higher urinary Cd concentrations. This predictive model could aid in devising personalized individual care plans tailored to heavy metal exposure profiles.

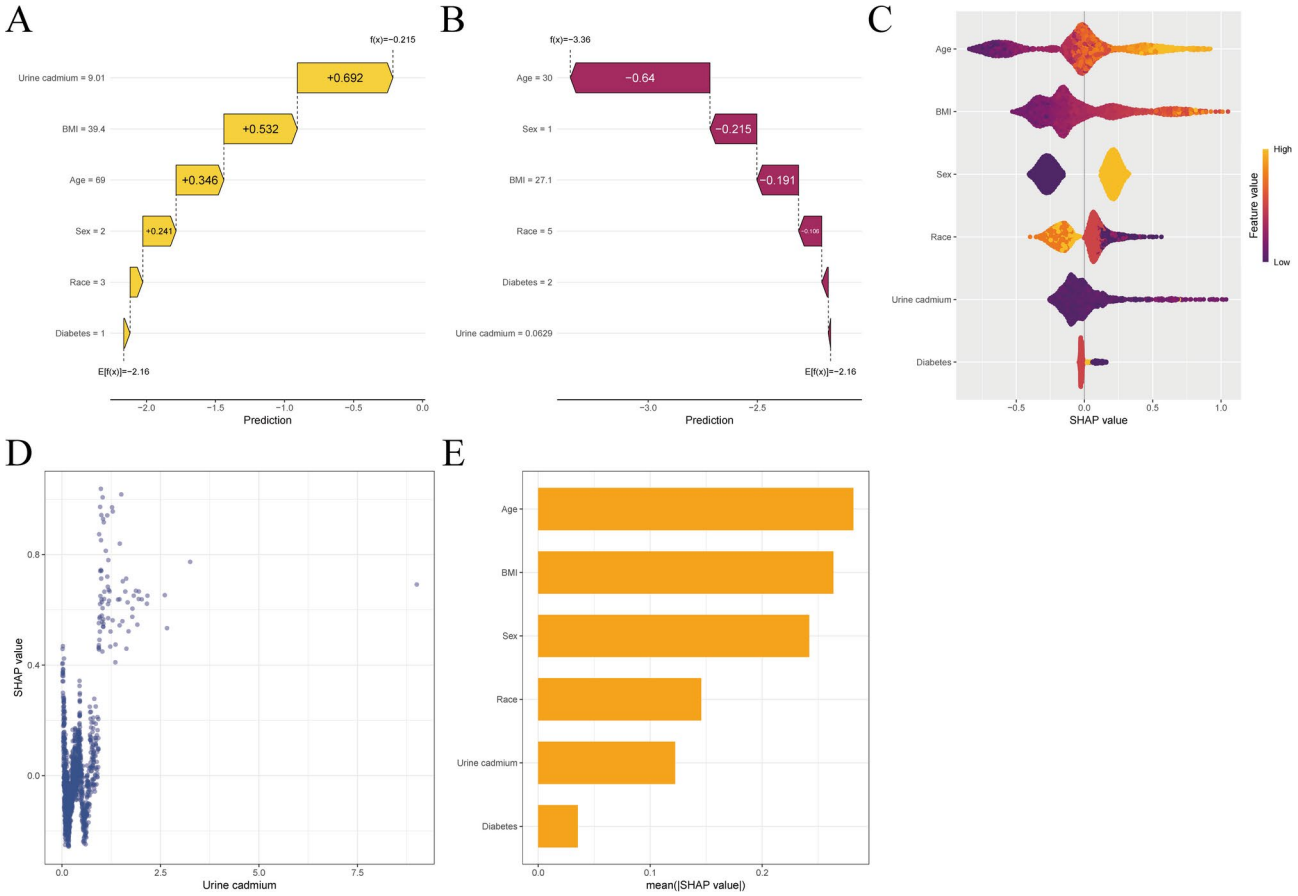

Fig. 3. Evaluation and comparison of the XGBoost model by SHAPs. (**A**, **B**) The SHAP person waterfall plot was constructed to illustrate the decision-making process for gallstone formation risk. The yellow bar indicates that the variable led to gallstone formation, whereas the brown bar indicates its inhibitory effect. (**C**) The contributions of six variables in the XGBoost predictive model were determined by calculating the SHAP. (**D**) Distribution diagram of the SHAP values of urinary Cd. (**E**) The importance of urinary Cd, age, sex, race, BMI and diabetes status in the XGBoost model.

There are some advantages in the present study. First, to our knowledge, this is the first study to elucidate the relationship between urinary Cd and gallstone prevalence. In addition, several potentially confounding variables were adjusted to enhance the reliability of the results, and various statistical methods were employed to ensure the robustness of the findings. Third, all the data were extracted from the NHANES dataset, which included a nationally representative sample of the population. The sample weights provided by the NHANES were utilized to extrapolate the results to the national population. Finally, we developed a XGBoost predictive model using the urinary Cd concentration, age, sex, race, and BMI. This model may assist in predicting the risk of gallstone formation. However, there are some limitations in this study. First, in this study, we adopted a cross-sectional design, and a prospective study is necessary to validate our findings. Second, the NHANES database lacks information on Cd sources and bioavailability. Therefore, it remains uncertain whether various uncontrollable factors, such as wastewater and cosmetics, could have influenced the results.

Conclusions

In summary, our study indicated that urinary cadmium concentration was significantly linearly associated with gallstone prevalence. Compared to the low urinary cadmium group, the high cadmium group had a increased risk of gallstone formation, highlighting the significant link between urinary cadmium levels and gallstone risk. Subgroup analysis revealed that higher urinary cadmium concentrations were associated with an increased risk of gallstones in younger individuals, males, those with a BMI > 35, Mexican Americans, Non-Hispanic Whites, smokers, and drinkers. Moreover, nine machine learning models were employed to construct the predictive model, and among these models, the XGBoost model exhibited the highest performance. People with an increased urinary Cd concentration should attempt to identify the cause, change their lifestyle or dietary habits, and try to decrease their urinary Cd concentrations, which may reduce the risk of gallstone formation.

Data availability

The datasets used for these analyses are publicly available (https://www.cdc.gov/nchs/nhanes/index.html). The code and original data will be provided as needed (Supplementary file 1 and Supplementary file 2).

Received: 20 September 2024; Accepted: 29 April 2025
Published online: 08 May 2025

References

1. Gutt, C., Schläfer, S. & Lammert, F. The treatment of gallstone disease. *Dtsch. Arztebl. Int.* **117**(9), 148–158 (2020).
2. Lim, J. et al. Obesity, adiposity, and risk of symptomatic gallstone disease according to genetic susceptibility. *Clin. Gastroenterol. Hepatol.* **20**(5), e1083–e1120 (2022).
3. Lammert, F. et al. Gallstones. *Nat. Rev. Dis. Primers* **2**, 16024 (2016).
4. Smelt, A. H. Triglycerides and gallstone formation. *Clin. Chim. Acta* **411**(21–22), 1625–1631 (2010).
5. Halloran, L. G., Schwartz, C. C., Vlahcevic, Z. R., Nisman, R. M. & Swell, L. Evidence for high-density lipoprotein-free cholesterol as the primary precursor for bile-acid synthesis in man. *Surgery* **84**(1), 1–7 (1978).
6. Nakeeb, A. et al. Insulin resistance causes human gallbladder dysmotility. *J. Gastrointest. Surg.* **10**(7), 940–948 (2006) (**discussion 8–9**).
7. Everhart, J. E. & Ruhl, C. E. Burden of digestive diseases in the United States part III: Liver, biliary tract, and pancreas. *Gastroenterology* **136**(4), 1134–1144 (2009).
8. Miquel, J. F. et al. Genetic epidemiology of cholesterol cholelithiasis among Chilean Hispanics, Amerindians, and Maoris. *Gastroenterology* **115**(4), 937–946 (1998).
9. Carey, M. C. & Paigen, B. Epidemiology of the American Indians' burden and its likely genetic origins. *Hepatology* **36**(4 Pt 1), 781–791 (2002).
10. Tsai, C. J., Leitzmann, M. F., Willett, W. C. & Giovannucci, E. L. Macronutrients and insulin resistance in cholesterol gallstone disease. *Am. J. Gastroenterol.* **103**(11), 2932–2939 (2008).
11. Tsai, C. J., Leitzmann, M. F., Willett, W. C. & Giovannucci, E. L. Central adiposity, regional fat distribution, and the risk of cholecystectomy in women. *Gut* **55**(5), 708–714 (2006).
12. Chait, A., Bierman, E. L. & Albers, J. J. Low-density lipoprotein receptor activity in cultured human skin fibroblasts. Mechanism of insulin-induced stimulation. *J. Clin. Invest.* **64**(5), 1309–1319 (1979).
13. Weikert, C. et al. Presence of gallstones or kidney stones and risk of type 2 diabetes. *Am. J. Epidemiol.* **171**(4), 447–454 (2010).
14. Portincasa, P. et al. Management of gallstones and its related complications. *Expert Rev. Gastroenterol. Hepatol.* **10**(1), 93–112 (2016).
15. Yao, X. et al. Stratification of population in NHANES 2009–2014 based on exposure pattern of lead, cadmium, mercury, and arsenic and their association with cardiovascular, renal and respiratory outcomes. *Environ. Int.* **149**, 106410 (2021).
16. Pan, Z., Gong, T. & Liang, P. Heavy metal exposure and cardiovascular disease. *Circ. Res.* **134**(9), 1160–1178 (2024).
17. Zheng, K. et al. Epidemiological evidence for the effect of environmental heavy metal exposure on the immune system in children. *Sci. Total Environ.* **868**, 161691 (2023).
18. Kaur, R. & Rawal, R. Influence of heavy metal exposure on gut microbiota: Recent advances. *J. Biochem. Mol. Toxicol.* **37**(12), e23485 (2023).
19. Clemens, S. Toxic metal accumulation, responses to exposure and mechanisms of tolerance in plants. *Biochimie* **88**(11), 1707–1719 (2006).
20. Satarug, S., Garrett, S. H., Sens, M. A. & Sens, D. A. Cadmium, environmental exposure, and health outcomes. *Environ. Health Perspect.* **118**(2), 182–190 (2010).
21. Liu, M. et al. Predicted fat mass and lean mass in relation to all-cause and cause-specific mortality. *J. Cachexia Sarcopenia Muscle* **13**(2), 1064–1075 (2022).
22. Völzke, H. et al. Independent risk factors for gallstone formation in a region with high cholelithiasis prevalence. *Digestion* **71**(2), 97–105 (2005).
23. Florez-Garcia, V. A. et al. Cadmium exposure and risk of breast cancer: a meta-analysis. *Environ. Res.* **219**, 115109 (2023).
24. Filippini, T. et al. Cadmium exposure and risk of breast cancer: A dose-response meta-analysis of cohort studies. *Environ. Int.* **142**, 105879 (2020).
25. Chowdhury, R. et al. Environmental toxic metal contaminants and risk of cardiovascular disease: Systematic review and meta-analysis. *BMJ* **362**, k3310 (2018).
26. Young, J. L. & Cai, L. Implications for prenatal cadmium exposure and adverse health outcomes in adulthood. *Toxicol. Appl. Pharmacol.* **403**, 115161 (2020).
27. Lee, M. H. et al. A metallomic approach to assess associations of serum metal levels with gallstones and gallbladder cancer. *Hepatology* **71**(3), 917–928 (2020).
28. Wang, J. et al. The association between blood heavy metals and gallstones: a cross-sectional study. *Sci. Total Environ.* **904**, 166735 (2023).
29. Yang, J., Lo, K. & Yang, A. Trends in urinary and blood cadmium levels in U.S. adults with or without comorbidities, 1999–2018. *Nutrients* **14**(4), 802 (2022).
30. Satarug, S., Baker, J. R., Reilly, P. E., Moore, M. R. & Williams, D. J. Cadmium levels in the lung, liver, kidney cortex, and urine samples from Australians without occupational exposure to metals. *Arch. Environ. Health* **57**(1), 69–77 (2002).
31. Horiguchi, H. et al. Comprehensive study of the effects of age, iron deficiency, diabetes mellitus, and cadmium burden on dietary cadmium absorption in cadmium-exposed female Japanese farmers. *Toxicol. Appl. Pharmacol.* **196**(1), 114–123 (2004).
32. Kikuchi, Y. et al. Uptake of cadmium in meals from the digestive tract of young non-smoking Japanese female volunteers. *J. Occup. Health* **45**(1), 43–52 (2003).
33. Jain, R. B. Concentrations of cadmium, lead, and mercury in blood among US cigarettes, cigars, electronic cigarettes, and dual cigarette-e-cigarette users. *Environ. Pollut.* **251**, 970–974 (2019).
34. Turner, A. Cadmium pigments in consumer products and their health risks. *Sci. Total Environ.* **657**, 1409–1418 (2019).
35. Yu, Q. et al. Study becomes insight: ecological learning from machine learning. *Methods Ecol. Evol.* **12**(11), 2117–2128 (2021).

Acknowledgements

We appreciate the contributions of all staff and participants to the U.S. National Health and Nutrition Examination Survey (NHANES).

Author contributions

Writing the manuscript: Zhaowei Wu; Data extraction and statistical analysis: Zhaowei Wu and Jinzhi Li; Reviewing and editing: Yong Chen; Project administration: Yong Chen; Conceptualization and supervision: Pan-

guo Wang and Shiming Jiang. All authors have read and agreed to the published version of the manuscript.

Funding
This study was funded by the National Natural Science Foundation of China under Grant No. 81871261.

Declarations

Competing interests
The authors declare no competing interests.

Ethics and approval
The study protocol was approved by the NCHS Research Ethics Review Board (ERB), and all participants provided written informed consent prior to participation. (https://www.cdc.gov/nchs/nhanes/irba98.htm).

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-00648-5.

Correspondence and requests for materials should be addressed to Y.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.