



Application of machine learning algorithms to screen potential biomarkers under cadmium exposure based on human urine metabolic profiles



Ting Zeng^{a,b}, Yanshan Liang^{a,b}, Qingyuan Dai^{a,b}, Jinglin Tian^{a,b}, Jinyao Chen^c, Bo Lei^a, Zhu Yang^b, Zongwei Cai^{b,*}

^a Food Science and Technology Program, Beijing Normal University-Hong Kong Baptist University United International College, Zhuhai 519087, China

^b State Key Laboratory of Environmental and Biological Analysis, Department of Chemistry, Hong Kong Baptist University, Hong Kong, China

^c Department of Nutrition, Food Safety and Toxicology, West China School of Public Health, Sichuan University, Chengdu 610041, China

ARTICLE INFO

Article history:

Received 27 December 2021

Revised 23 January 2022

Accepted 4 March 2022

Available online 7 March 2022

Keywords:

Human urine

High-resolution mass spectrometry

Cadmium exposure

Metabolic profiles

Machine learning

ABSTRACT

Exposure to environmental cadmium increases the health risk of residents. Early urine metabolic detection using high-resolution mass spectrometry and machine learning algorithms would be advantageous to predict the adverse health effects. Here, we conducted machine learning approaches to screen potential biomarkers under cadmium exposure in 403 urine samples. In positive and negative ionization mode, 4207 and 3558 features were extracted, respectively. We compared seven machine learning algorithms and found that the extreme gradient boosting (XGBoost) and random forest (RF) classifiers showed better accuracy and predictive performance than others. Following 5-fold cross-validation, the value of area under curve (AUC) was both 0.93 for positive and negative ionization modes in XGBoost classifier. In the RF classifier, AUC were 0.80 and 0.84 for positive and negative ionization modes, respectively. We then identified a biomarker panel based on XGBoost and RF classifiers. The incorporation of machine learning models into urine analysis using high-resolution mass spectrometry could allow a convenient assessment of cadmium exposure.

© 2022 Published by Elsevier B.V. on behalf of Chinese Chemical Society and Institute of Materia Medica, Chinese Academy of Medical Sciences.

Despite the fact that the official regulations of environmental cadmium (Cd) pollution have been in place for over one hundred years, the screening of potential biomarkers related to Cd toxicity is still in need to safeguard human health [1]. The tolerable Cd exposure level is 0.83 µg/kg body weight/day and urinary Cd is 5.24 µg/g creatinine according to the Food and Agriculture Organization/World Health Organization (FAO/WHO) guideline [2]. Appropriately, the safety standard of Cd in rice is 0.2 mg/kg according to the China Food Safety National Standard for Maximum Levels of Contaminants in Foods (GB2762-2017). In many cases, urine is routinely and robustly used to diagnose and monitor health conditions because it is easy-obtained and non-invasive sampling [3,4]. The interpretation of the urinary profile under Cd exposure would be beneficial to the determination of Cd exposure grades based on the screening markers. Mass spectrometry (MS) is able to provide data information including the retention time, ion mass spectral data and peak intensities for thousands of molecules concurrently,

which could deepen our knowledge in the urinary samples. Considering the growing capacities of high-resolution MS (HRMS) to reveal and characterize the pollutant-relevant health status, as well as the accessible nature of urine samples, a bunch of relevant studies has been conducted over the past years [5–7].

Machine learning (ML) algorithms, in parallel, have been successfully applied in clinical laboratories for years, demonstrating their ability to deal with complex data sets [8]. The machine learning approach is a type of artificial intelligence, which contains supervised and unsupervised learning, and has been used to improve the diagnostic capabilities of HRMS on urinary samples for more challenging problems [9]. Briefly, a typical application of ML works as follows. First, it is the infrastructure data collection and pre-processing. The most appropriate machine learning algorithms for the type of data would be selected after loading the data and adjusting the parameters. After training and testing the model on data, cross-validation would be properly used for evaluation through the best-performing machine learning algorithms and the optimized signatures [10,11]. The ratio of the training to testing set was usually set at 70:30 or 80:20, and several compounds were pinpointed to distinguish the exposure status [12]. Altogether, ma-

* Corresponding author.

E-mail address: zwcai@hkbu.edu.hk (Z. Cai).

chine learning algorithms have great potential for automating the interpretation and analysis of the urinary sample under Cd exposure [13,14].

The purpose of this study was to develop ML-based models that used urinary metabolic characteristics to screen potential biomarkers under Cd exposure. We collected 403 urine samples and analyzed them using liquid chromatography-HRMS (LC-HRMS). We then compared seven machine learning algorithms on the LC-HRMS data set, including extreme gradient boosting classifier (XGBoost classifier), logistic regression classifier, support vector classifier (SVC), multi-layer perceptron classifier (MLP classifier), K-nearest neighbor classifier (KNN classifier), Gaussian naive Bayes classifier (GaussianNB classifier) and random forest classifier (RF classifier). Via 5-fold cross-validation, the area under curve (AUC), accuracy, sensitivity, specificity, positive predictive value, and negative predictive value were calculated. Finally, we identified a biomarker panel based on the selected machine learning models. The incorporation of machine learning models into urine analysis by HRMS allows a convenient assessment for Cd exposure.

As for the urinary characteristics of Cd-exposed humans, the present study included 403 participants, among which 280 and 123 belong to the exposure and control groups, respectively. The discriminating criterion was whether the villagers lived in Cd exposure or control areas. The main exposure route of Cd to the villagers was dietary Cd intake. The Cd exposure area was mining-affected and the villagers mainly subsisted on home-grown rice. The Cd concentration of home-grown rice greater than 0.2 mg/kg in the polluted area was reported in the previous study [15]. The villagers in the control area with a Cd concentration of home-grown rice lower than 0.05 mg/kg have a similar lifestyle to the villagers in the polluted area. Table 1 showed the demographic characteristics of the Cd exposure and the control group. Of the 403 participants included in this study, the characteristics such as gender, frequency of smoking, age, residence duration and body mass index between Cd exposure and the control group showed no significant differences ($P > 0.05$). Nevertheless, there were significant differences ($P < 0.05$) in the concentrations of urinary cadmium (UCd), blood cadmium (BCd) and β 2-microglobulin between the Cd exposure and the control groups. Among them, β 2-microglobulin is one of the nephrotoxicity signal molecules. The renal function tests, hepatic function tests and cardiac function tests were conducted. The results of cardiac function tests revealed no significant differences ($P > 0.05$) between the two groups. While some of the clinical traits in the renal function tests (creatinine and blood urea nitrogen) and the hepatic function tests (total bile acid, alkaline phosphatase, gamma-glutamyl transpeptidase, indirect bilirubin, direct bilirubin and total bilirubin) showed a significant difference ($P < 0.05$) between Cd exposure and the control groups. After the acquisition of urine metabolic profiles, 4207 and 3558 features were extracted in positive and negative ionization mode, respectively. Then the partial least squares-discriminant analysis (PLS-DA) and the following metabolic pathway analysis were conducted using MetaboAnalyst (<https://www.metaboanalyst.ca>). As shown in Fig. S1 (Supporting information), the exposure group clustered away from the control group and the peak intensities of the quality control sample were stable during the whole analysis guaranteeing the system stability (RSD < 30%).

Then, these samples were randomly split into two parts at a ratio of 7:3. The forest plot analysis of different machine learning algorithms based on positive and negative ionization mode mass spectra was conducted in terms of classifying the Cd exposure using data from sufficient metabolic features (Fig. 1). In this step, the random subsampling method was applied primarily due to its convenient comparison among multiple algorithms. The XGBoost and RF classifiers outperformed other algorithms when applied to

the data from both positive and negative ionization modes. Following 5-fold cross-validation, accuracy, sensitivity, specificity, positive and negative predictive values of both models were calculated. The AUC for positive and negative ionization modes were both 0.93 in the XGBoost classifier. In the RF classifier, the AUC were 0.80 and 0.84 for positive and negative ionization modes, respectively (Table 2).

As XGBoost and RF classifiers were preferred, the top thirty features calculated from their feature importance were selected from positive and negative ionization modes, as shown in Fig. 2. Among the top thirty features selected from the XGBoost classifier and RF classifier, biomarkers were identified by matching their mass fragmentation patterns with those in the Human Metabolome Database (HMDB, <https://hmdb.ca>). In total, seven biomarkers were found, including S-adenosylmethionine, dehydroepiandrosterone sulfate, propionic acid, L-asparagine, homovanillic acid, deoxyinosine and deoxycytidine (Table 3). These compounds are involved in amino acid metabolism, purine metabolism, steroid hormone biosynthesis, etc. (Fig. S2 in Supporting information) [16–20]. The Spearman correlation matrix of the identified compounds from human urine under Cd exposure was shown in Fig. 3. Correlation were obtained between UCd and BCd ($r = 0.26$), UCd and L-asparagine ($r = -0.22$), deoxycytidine and dehydroepiandrosterone sulfate ($r = -0.32$), deoxycytidine and homovanillic acid ($r = -0.34$), deoxycytidine and S-adenosylmethionine ($r = -0.31$), S-adenosylmethionine and L-asparagine ($r = 0.36$), S-adenosylmethionine and dehydroepiandrosterone sulfate ($r = 0.53$), S-adenosylmethionine and homovanillic acid ($r = 0.73$), homovanillic acid and L-asparagine ($r = 0.37$), homovanillic acid and dehydroepiandrosterone sulfate ($r = 0.51$) as well as dehydroepiandrosterone sulfate and L-asparagine ($r = 0.23$). It is worth noting that additional work is required to identify more endogenous metabolites in urine and to expand the depth and breadth of this altered profile. The current models had scientific merits but were not comprehensive.

Although popular machine learning algorithms have been applied to predict Cd pollution in water springs, agricultural soil and lettuce leaves, to name a few [21–24], our research was the first study using seven different machine learning algorithms to screen potential biomarkers under Cd exposure based on human urine metabolic profiles. After the acquisition of MS data, fitting, coding, debugging, and fine-tuning of the specific machine learning algorithms, the XGBoost and RF classifiers showed better accuracy and predictive performance than others. These two algorithms are both decision tree-based supervised ensemble machine learning methods [25,26]. To avoid overfitting, k -fold cross-validation is commonly used in these two algorithms [27]. The most suitable one for our dataset is 5-fold cross-validation. The principles of various algorithms differ [28,29]. The superior performances of XGBoost and RF classifiers in the present study may attribute to the skewed feature distributions and the unbalanced class distributions (123 control samples and 280 exposure samples) according to previous literature [30]. XGBoost and RF classifiers are both promising in constructing predictive models and, as a result, providing variables of relatively high importance [31]. To gain insights into these variables and several clinicopathologic characteristics, we may find more useful information hidden inside the high-dimensional mass spectrometry dataset. We trained and tested the models with a binary classified dataset (polluted or non-polluted) in this study. In reality, the polluted situations could be classified as non-polluted, low-polluted and high-polluted areas which would be more complex than binary classification. Thus, more human biological samples and optimized algorithms should be used to improve this multi-classification point in future studies. In short, the present study has added a new reference for selecting data-driven machine

Table 1
Characteristics of the populations between Cd exposure and control group at baseline ($n = 403$).

Characteristics	Control group ($n = 123$)	Exposure group ($n = 128$)	P-value	
Gender, n (%) ^a	Male	57 (46.3)	114 (40.7)	0.293
	Female	66 (53.7)	166 (59.3)	
Frequency of smoking, n (%) ^a	Often	40 (32.5)	84 (30.0)	0.775
	Seldom	8 (6.5)	23 (8.2)	
	Never	75 (61.0)	173 (61.8)	
Age, years, median (IQR) ^b	62.0 (51.0, 66.0)	58.0 (51.0, 65.0)	0.135	
Residence duration, years, median (IQR) ^b	50.0 (39.0, 64.0)	51.0 (38.0, 62.0)	0.649	
BMI, median (IQR) ^b	24.2 (21.9, 27.4)	24.4 (22.4, 26.8)	0.786	
UCd, $\mu\text{g/L}$, median (IQR) ^b	1.4 (0.9, 2.1)	4.2 (2.4, 6.7)	<0.001	
BCd, $\mu\text{g/L}$, median (IQR) ^b	2.4 (1.2, 5.9)	4.0 (2.8, 6.7)	<0.001	
β 2-MG, nmol/L, median (IQR) ^b	0.8 (0.05, 3.5)	2.6 (0.3, 9.5)	<0.001	
Renal function tests				
UA, $\mu\text{mol/L}$, median (IQR) ^b	307.0 (253.0, 352.0)	302.0 (252.0, 351.0)	0.783	
Crea, $\mu\text{mol/L}$, median (IQR) ^b	63.9 (57.8, 78.9)	83.6 (56.3, 114.8)	<0.001	
BUN, mmol/L, median (IQR) ^b	0.0 (0.0, 0.0)	0.0 (0.0, 4.5)	<0.001	
Hepatic function tests				
TBA, $\mu\text{mol/L}$, median (IQR) ^b	0.0 (0.0, 0.0)	3.0 (1.9, 4.2)	<0.001	
ALP, U/L, median (IQR) ^b	76.2 (59.2, 87.6)	99.0 (57.0, 137.0)	<0.001	
GGT, U/L, median (IQR) ^b	0.0 (0.0, 0.0)	0.0 (0.0, 28.0)	<0.001	
AST, U/L, median (IQR) ^b	27.5 (22.2, 32.9)	27.0 (23.0, 32.0)	0.526	
ALT, U/L, median (IQR) ^b	26.1 (19.1, 38.0)	25.0 (18.0, 33.0)	0.091	
ALB, g/L, median (IQR) ^b	46.8 (45.3, 48.3)	47.0 (45.0, 48.4)	0.903	
GLB, g/L, median (IQR) ^b	28.0 (26.0, 30.0)	29.0 (25.8, 33.0)	0.012	
A/G, median (IQR) ^b	2.0 (2.0, 2.0)	1.6 (1.4, 1.8)	<0.001	
TP, g/L, median (IQR) ^b	74.8 (72.0, 77.3)	75.4 (71.9, 79.8)	0.082	
IBIL, $\mu\text{mol/L}$, median (IQR) ^b	9.0 (7.0, 12.0)	8.3 (6.3, 10.6)	<0.001	
DBIL, $\mu\text{mol/L}$, median (IQR) ^b	5.2 (4.0, 6.5)	3.8 (2.8, 5.6)	<0.001	
TBIL, $\mu\text{mol/L}$, median (IQR) ^b	14.6 (11.7, 18.6)	12.6 (10.4, 15.5)	<0.001	
Cardiac function tests				
Diastolic blood pressures, mmHg, median (IQR) ^b	82.0 (78.0, 92.0)	84.0 (74.0, 93.0)	0.915	
Systolic blood pressures, mmHg, median (IQR) ^b	134.0 (124.0, 152.0)	142.0 (130.0, 154.0)	0.041	
Heart rate, bpm, median (IQR) ^b	76.0 (70.0, 84.0)	76.0 (70.0, 84.0)	0.550	

Abbreviation: BMI, body mass index; UCd, urinary cadmium; BCd, blood cadmium; β 2-MG, β 2-microglobulin; UA, uric acid; Crea, creatinine; BUN, blood urea nitrogen; TBA, total bile acid; ALP, alkaline phosphatase; GGT, gamma-glutamyl transpeptidase; AST, aspartate aminotransferase; ALT, alanine aminotransferase; ALB, albumin; GLB, globulin; A/G, albumin/globulin; TP, total protein; IBIL, indirect bilirubin; DBIL, direct bilirubin; TBIL, total bilirubin; IQR, interquartile range.

^a Chi-square test was used.

^b Mann-Whitney-U test was used.

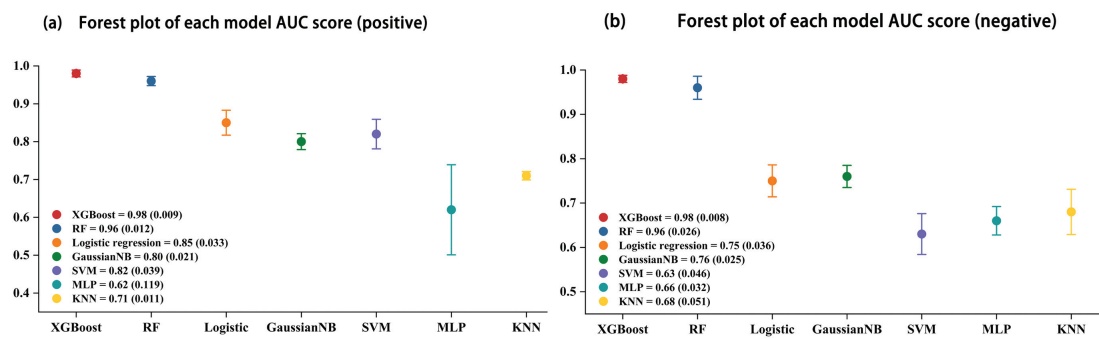


Fig. 1. Forest plot of different machine learning algorithms based on positive (a) and negative (b) ionization mode mass spectra. The red, blue, orange, green, purple, light blue and light yellow represent the results from XGBoost, RF, Logistic regression, GaussianNB, SVM, MLP and KNN classifiers, respectively.

Table 2
Summary of 5-fold cross-validation result of different machine learning models on the dataset.

Models	Dataset	AUC	Accuracy	Sensitivity	Specificity	Positive predictive value	Negative predictive value
XGBoost classifier	Positive ion-mode	0.93	0.93	0.79	0.97	0.88	0.94
	Negative ion-mode	0.93	0.98	0.97	0.99	0.97	0.99
RF classifier	Positive ion-mode	0.80	0.80	1.00	0.76	0.49	1.00
	Negative ion-mode	0.84	0.83	1.00	0.78	0.59	1.00

learning algorithms for a metabolic analysis of urine upon Cd exposure.

Due to the public health concerns about Cd, many metabolomics studies have been conducted. These reports have evidenced that amino acid metabolism, carbohydrate metabolism, purine metabolism, steroid hormone biosynthesis, creatine path-

way, fatty acid oxidation, lipid metabolism, etc. would be influenced by Cd exposure [32–35]. The biomarkers identified in this study were related to the changes in amino acid metabolism (S-adenosylmethionine, L-asparagine and homovanillic acid), purine metabolism (deoxyinosine) and steroid hormone biosynthesis (dehydroepiandrosterone sulfate). By convention, the aberration of

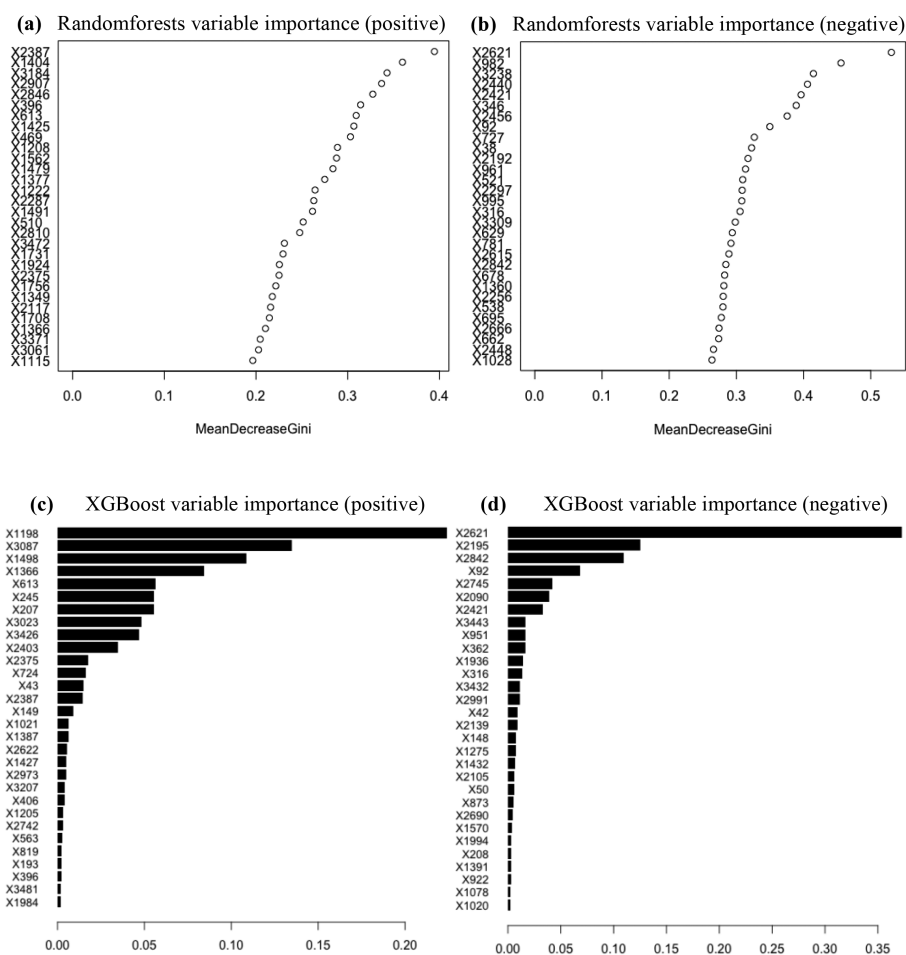


Fig. 2. Top thirty features calculated from their feature importance selected from positive (a) and negative (b) ionization mode mass spectra based on the random forest classifier and positive (c) and negative (d) ionization mode mass spectra based on the extreme gradient boosting classifier. The x-axis represents the results of mean decrease Gini and weight-based feature importance to RF and XGBoost classifiers, respectively.

Table 3

Biomarkers identified from human urine under Cd exposure based on different machine learning models.

Biomarkers	R _t (min)	Molecular weight (Da)	Mode ^a	Error (ppm)	MS/MS fragments	AUC	Model
S-adenosylmethionine	5.35	131.0336	QE (-)	2	71.0494	0.39	RF Classifier
Dehydroepiandrosterone sulfate	7.67	367.1587	QE (-)	1	79.9563	0.49	RF Classifier
Propionic acid	2.33	116.0705	QE (+)	1	57.0337	0.49	RF Classifier
L-Asparagine	6.29	196.0688	QE (+)	2	138.0390	0.43	RF Classifier
Homovanillic acid	3.69	200.0915	QE (+)	1	67.0542, 79.0540	0.44	RF Classifier
Deoxyinosine	0.82	329.0046	QE (+)	1	121.0498	0.50	RF Classifier
Deoxycytidine	8.93	133.5298	QE (+)	5	95.0491	0.43	XGBoost Classifier

^a QE (-) and QE (+) represent the negative and positive ionization modes under Q-Exactive Focus Hybrid Quadrupole-Orbitrap Mass Spectrometer.

amino acid metabolism indicated its energy expenditure under Cd exposure. As an illustration, S-adenosylmethionine, metabolized from the essential amino acid, methionine, is related to liver disease [16]. Concomitantly, purine metabolism is associated with nephrotoxicity [36,37]. The alteration of purine metabolism in this study also indicated the influence of Cd on the kidney. Dehydroepiandrosterone sulfate is a rich circulating steroid in humans and the steroid hormone biosynthesis dedicates to the normal development of the individual. These findings were generally consistent with the previous reports. As shown in the Spearman correlation analysis, the positive correlation coefficient between the concentration of UCd and BCd, two typical indexes for evaluating Cd exposure, was relatively higher than others. It was worth mentioning that mean decrease Gini and weight-based feature importance ranking were applied to RF and XGBoost classifiers,

respectively. Other feature selection methodologies in these two algorithms could also be good alternatives for obtaining metabolic information to distinguish Cd exposure from control. Inferred from the information about hepatic and renal damage collected in the present study (Table 1) and supported by similar findings in the previous study [38–41], we concluded that the liver and kidney were susceptible to Cd exposure regardless of acute and chronic toxicity. As more and more omics data are accumulated [42,43], artificial intelligence-based techniques in combination with omics might be effective in recognizing unknown metabolic biomarkers and pathways, allowing for personalized risk assessments.

In conclusion, we acquired high-resolution mass spectrometry datasets from 403 human urine samples, trained XGBoost and RF classifiers and identified a biomarker panel under Cd exposure. Taking into account the vast potential of big data, prospective risk

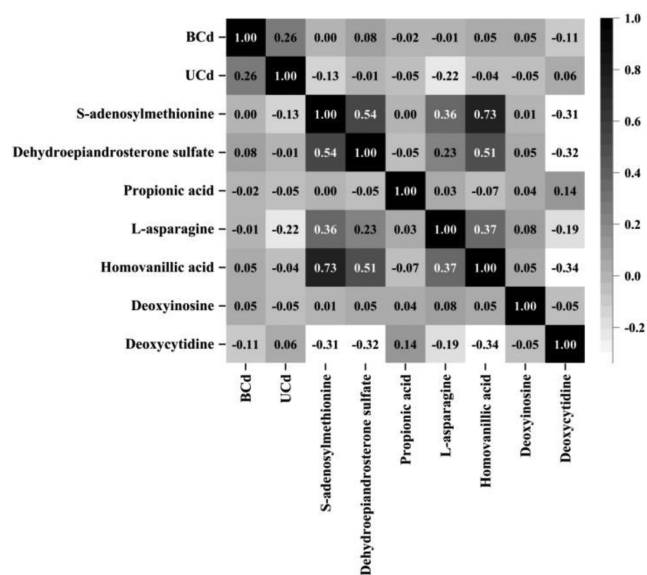


Fig. 3. Spearman correlation matrix of the identified compounds from human urine under Cd exposure.

stratification of machine learning algorithms, more sophisticated models are needed based on the two algorithms.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was supported by the National Key Research and Development Program of China (No. 2017YFC1600500).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ccllet.2022.03.020.

References

[1] S. Himeno, K. Aoshima, Cadmium Toxicity: New Aspects in Human Disease, Rice Contamination, and Cytotoxicity, Springer, Singapore, 2019.

[2] S. Satarug, *Toxics* 6 (2018) 15.
 [3] E.D. Evans, C. Duvallet, N.D. Chu, et al., *Sci. Rep.* 10 (2020) 1–13.
 [4] A.S. Martinez-Vernon, J.A. Covington, R.P. Arasaradnam, et al., *PLoS One* 13 (2018) e0204425.
 [5] Y. Gao, Y. Lu, S. Huang, et al., *Environ. Sci. Technol.* 48 (2014) 6409–6418.
 [6] H. Li, K. Huang, S. Jin, et al., *Int. J. Hyg. Environ. Heal.* 222 (2019) 556–562.
 [7] T. Zeng, Y.S. Liang, J.Y. Chen, et al., *Environ. Int.* 154 (2021) 106646.
 [8] E.H. Wilkes, G. Rumsby, G.M. Woodward, *Clin. Chem.* 64 (2018) 1586–1595.
 [9] H.E. Anderson, T. Liden, B.K. Berger, et al., *Anal. Chim. Acta* 1172 (2021) 338668.
 [10] F. Roux-Dalvai, C. Gotti, M. Leclercq, et al., *Mol. Cell Proteom.* 18 (2019) 2492–2505.
 [11] G.L. Streun, A.E. Steuer, L.C. Ebert, A. Dobay, T. Kraemer, *Clin. Chem. Lab. Med.* 59 (2021) 1392–1399.
 [12] Y.C. Huang, H.H. Chung, E.P. Dutkiewicz, et al., *Anal. Chem.* 92 (2019) 1653–1657.
 [13] U.W. Liebal, A.N.T. Phan, M. Sudhakar, K. Raman, L.M. Blank, *Metabolites* 10 (2020) 243.
 [14] E.H. Wilkes, E. Emmett, L. Beltran, G.M. Woodward, R.S. Carling, *Clin. Chem.* 66 (2020) 1210–1218.
 [15] J. Huo, Z. Huang, R. Li, et al., *PLoS One* 13 (2018) e0201454.
 [16] J.M. Mato, M.L. Martínez-Chantar, S.C. Lu, *Ann. Hepatol.* 12 (2015) 183–189.
 [17] N.N. Pavlova, S. Hui, J.M. Ghergurovich, et al., *Cell Metab.* 27 (2018) 428–438.
 [18] J. Płonka, M. Babiuch, H. Barchanska, *Chemosphere* 286 (2022) 131592.
 [19] M. Chen, B. Zhang, S. Cai, et al., *J. Proteom.* 200 (2019) 134–143.
 [20] J. Neunzig, R. Bernhardt, *PLoS One* 9 (2014) e89727.
 [21] F.K. Abu Salem, M. Jurdi, M. Alkadri, F. Hachem, H.R. Dhaini, *Environ. Sci. Pollut. Res.* 29 (2022) 8253–8268.
 [22] X. Li, T. Geng, W. Shen, J. Zhang, Y. Zhou, *Ecotoxicol. Environ. Saf.* 209 (2021) 111773.
 [23] K. Nakamura, T. Yasutaka, T. Kuwatani, T. Komai, *Chemosphere* 186 (2017) 501–509.
 [24] Z. Xin, S. Jun, T. Yan, et al., *Chemometr. Intell. Lab. Syst.* 200 (2020) 103996.
 [25] L. Wei, Y. Huang, Z. Chen, et al., *Front. Oncol.* 11 (2021) 4148.
 [26] X. Zhu, X. Wang, Y.S. Ok, *J. Hazard. Mater.* 378 (2019) 120727.
 [27] V. Cantoni, R. Green, C. Ricciardi, et al., *Comput. Math Methods Med.* 16 (2021) 5288844.
 [28] X. Jia, B. Hu, B.P. Marchant, et al., *Environ. Pollut.* 250 (2019) 601–609.
 [29] M. Hussain, Z. Chen, M. Lv, et al., *Chin. Chem. Lett.* 31 (2020) 3163–3167.
 [30] Y. Guo, A. Graber, R.N. McBurney, R. Balasubramanian, *BMC Bioinform.* 11 (2010) 1–19.
 [31] S.K. Park, Z. Zhao, B. Mukherjee, *Environ. Health* 16 (2017) 1–17.
 [32] T. Guan, Y. Xin, K. Zheng, et al., *BioMetals* 34 (2021) 33–48.
 [33] X. He, Z. Qi, H. Hou, J. Gao, X.X. Zhang, *J. Hazard. Mater.* 388 (2020) 121791.
 [34] M.Y. Xu, Y.J. Sun, P. Wang, L. Yang, Y.J. Wu, *Ecotoxicol. Environ. Saf.* 195 (2020) 110467.
 [35] L. Zong, J. Xing, L. Shu, et al., *Ecotoxicol. Environ. Saf.* 147 (2018) 26–33.
 [36] G. Xun, C. Yang, Y. Hong, et al., *Sci. Total Environ.* 676 (2019) 72–86.
 [37] L. Xie, Y. Zhao, J. Duan, et al., *Chem. Res. Toxicol.* 33 (2020) 1897–1906.
 [38] Y. Liang, T. Zeng, J. Tian, et al., *Sci. Total Environ.* 798 (2021) 149379.
 [39] M. Valcke, N. Ouellet, M. Dubé, et al., *Toxicol. Lett.* 312 (2019) 148–156.
 [40] M.Y. Xu, P. Wang, Y.J. Sun, Y.J. Wu, *Toxicology* 384 (2017) 50–58.
 [41] T. Zeng, W. Guo, L. Jiang, et al., *Sci. Total Environ.* 801 (2021) 149803.
 [42] L. Xiang, J. Nie, L. Wang, et al., *Chin. Chem. Lett.* 32 (2021) 2197–2202.
 [43] H. Luan, W. Gu, H. Li, et al., *J. Transl. Med.* 19 (2021) 1–10.