

Association of urinary and blood concentrations of heavy metals with measures of bone mineral density loss: A Data mining approach with the results from the National Health and Nutrition Examination Survey (NHANES).

João Paulo B Ximenez, PhD (1); Ariane Zamarioli, PhD (2); Melissa A Kacena, PhD (3); Rommel Melgaço Barbosa, PhD (4); Fernando Barbosa Jr, PhD (1)

- (1) Laboratório de Toxicologia Analítica e de Sistemas, Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP, Brazil;
- (2) Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, SP, Brazil;
- (3) Department of Orthopaedic Surgery, Indiana University School of Medicine, Indianapolis, IN, USA;
- (4) Instituto de Informática, Universidade Federal de Goiás, Goiânia, GO, Brazil.

Correspondence: João Paulo B Ximenez, Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 14040-903, Brazil.

Electronic address: joaopaulo.ximenez@usp.br.

Keywords: Bone Mineral Loss. Metals. Metalloids. Machine Learning. NHANES. ICP-MS.

Wordcount: Abstract: 238, Main text: ~ 7000, Tables: 3, Figures: 3.

This is the author's manuscript of the article published in final edited form as:

Ximenez, J. P. B., Zamarioli, A., Kacena, M. A., Barbosa, R. M., & Barbosa, F. (2021). Association of Urinary and Blood Concentrations of Heavy Metals with Measures of Bone Mineral Density Loss: A Data Mining Approach with the Results from the National Health and Nutrition Examination Survey. *Biological Trace Element Research*, 199(1), 92–101. <https://doi.org/10.1007/s12011-020-02150-7>

ABSTRACT

Osteoporosis and its consequence of fragility fracture represent a major public health problem. Human exposure to heavy metals has received considerable attention over the last decades. However, little is known about the influence of co-exposure to multiple heavy metals on bone density. The present study aimed to examine the association between exposure to metals and bone mineral density (BMD) loss. Blood and urine concentrations of 20 chemical elements were selected from 3 cycles (2005-2010) NHANES (National Health and Nutrition Examination Survey), in which we included white women over 50 years of age and previously selected for BMD testing (N = 1,892). The bone loss group was defined as participants having T-score < -1.0, and the normal group was defined as participants having T-score \geq -1.0. We developed classification models based on support vector machines capable of determining which factors could best predict BMD loss. The model which included the five-best features-selected from the random forest were: age, body mass index, urinary concentration of arsenic (As), cadmium (Cd), and tungsten (W), which have achieved high scores for accuracy (92.18%), sensitivity (90.50%), and specificity (93.35%). These data demonstrate the importance of these factors and metals to the classification since they alone were capable of generating a classification model with a high prediction of accuracy without requiring the other variables. In summary, our findings provide insight into the important, yet overlooked impact that arsenic, cadmium, and tungsten have on overall bone health.

Introduction

Osteoporosis and its consequences represent a significant public health issue with about 40% of white postmenopausal women being affected by osteoporosis, and, with an aging population, this number is expected to increase over the next several years. The fractures most commonly occur in the spine, hip, or wrist, but other bones such as the trochanter, humerus, or ribs can also be affected [1, 2].

Fractures that arise as a consequence of osteoporosis result in considerable morbidities, increased mortality, and increased health-care costs [3]. Low bone mass is a crucial component of fracture risk, but several risk factors, such as age, sex, low body-mass index, previous fractures, asthma, cardiovascular disease, chronic liver disease, advanced chronic kidney disease, diabetes, rheumatoid arthritis, systemic lupus erythematosus, glucocorticoids use, smoking, alcohol, and family history of fractures have to be considered as well [4–7].

Human exposure to toxic metals and metalloids has received considerable attention over the last decades, so much so that the Agency for Toxic Substances and Disease Registry (ATSDR) has classified arsenic, cadmium, lead, and mercury in the top of the priority list of hazardous substances, requiring constant evaluation of human exposure [8]. Exposure to metals and metalloids occurs through various routes, and the time and "dose" of exposure differ widely among individuals. Meanwhile, polluted water, soil, air, smoking, and food are known as main routes of exposure [9–11].

Exposure to toxic metals and metalloids have been reported as a risk factor for fractures and degenerative bone diseases such as osteoporosis [12]. On the other hand, little is known about the influence of co-exposure to multiple metals and metalloids on bone density. Cadmium (Cd) is widely known to have toxic effects on bones, in which *in vivo* and *in vitro* studies have shown that exposure to Cd decreases

bone mineralization, alters bone formation, and increases fracture and osteoporosis risk [13–16]. Commonly in the environment, lead is easy to be absorbed; and, in human adults, trabecular and cortical bones store 90-95% of the lead found within the body [17]. Clinical studies have shown that as the amount of lead accumulated in the body increases, bone density decreases and fracture risk increases [18, 19]. However, research on the association between heavy metals and bone health is scarce, requiring further and more specific investigations.

Data analysis models often associate exposure from a single compound to health outcomes. On the other hand, current developments in data mining techniques enable an analysis of co-exposure to multiple compounds on health outcomes [20]. Therefore, the utilization of such an approach would allow for a better understanding of the health effects associated with co-exposure to multiple compounds [21, 22]. Therefore, the present study was conducted to examine the associations of blood and urinary levels of toxic elements with bone mineral density (BMD) loss in a representative sample, who participated in the 2005-2006, 2007-2008 and 2009-2010 survey cycles of the National Health and Nutrition Examination Survey (NHANES) with the use of a data mining approach.

Materials and methods

Data source

Data were obtained from 3 cycles (2005-2006, 2007-2008 and 2009-2010) of NHANES data. NHANES is a population-based survey of the non-institutionalized U.S. population that includes demographic and laboratory data, interview data, and a physical examination of the subjects [23–25]. In the present study, we included white women in the analysis if they were over 50 years of age and had been selected for

BMD testing (N = 1,892), since this group has a higher prevalence of osteoporosis. The bone loss group was defined as participants having T-score values below -1.0 (T-score < -1.0), and the normal group was defined as participants having T-score values equal to or higher than -1.0 (T-score \geq -1.0).

Metals and metalloids in blood and urine

We considered the NHANES data reported for individual levels in:

1. Whole blood: Cadmium (Cd), Lead (Pb), Mercury (Hg - total and inorganic);
and;
2. Urine: Arsenic (As - total and speciated), Antimony (Sb), Barium (Ba), Beryllium (Be), Cadmium (Cd), Cobalt (Co), Cesium (Cs), Lead (Pb), Mercury (Hg), Molybdenum (Mo), Platinum (Pt), Thallium (Tl), Tungsten (W), and Uranium (U).

Trace elements were measured in clinical specimens by the National Center for Environmental Health Laboratories (CDC, Atlanta, GA) by using inductively coupled plasma-mass spectrometry (ICP-MS) [26]. The creatinine-adjusted levels were considered for urinary concentrations. The levels of Be, Co, Cs, Mo, Pb, Tl, U, arsenous acid, arsenic acid, arsenocholine, monomethylarsonic acid and trimethylarsine oxide in urine were excluded from the data analysis because results were below the limit of detection.

Covariates and potential confounders

Demographic variables (age, body-mass index), medical history (heart attack, diabetes, cancer, asthma, chronic liver disease, rheumatoid arthritis), glucocorticoid

use, heavy alcohol use (eight or more drinks a week), and tobacco use (never, former, and current smoker) were considered as covariates in the multivariate models.

Data Analysis

Feature Selection

Feature selection is a pre-processing step of data mining, which aims for identification and removal of features which are considered unimportant to the classification process. We employed a filter wrapper method built around the random forest classification algorithm to evaluate individually each feature of the feature set and determine their relative importance on T-score [27]. The entire analysis was conducted using R software, version 3.6.2 [28].

Classification algorithms and model evaluation

After removing the irrelevant features, we trained a support vector machines (SVM) algorithm to verify the importance of the remaining variables and to determine which factors could best predict BMD loss. SVM is a classification technique introduced by Cortes and Vapnik [29] which has been popularized in the data mining and classification literature due to its efficiency and empirical success [30, 31]. SVM algorithm aims to obtain an optimal hyperplane with maximum margin to separate the classes of samples applying nonlinear kernel functions to map data into high-dimensional space. In summary, the algorithm computes the decision boundary based on the samples that are nearest to the maximum-margin hyperplane, which are designated support vectors [30, 31].

The construction of the SVM models follows a specific methodology combined with the Random Forest measurements (mean decrease in accuracy). We start by

training the SVM with only one variable which received the highest importance in the Random Forest metric. Next, we added the second-best rated variable to the training set. This process resulted in $K = \{k_1; k_2; \dots; k_m\}$ feature subsets, which are generated and used to train an SVM model, where k_j is composed of the j best features according to the Random Forest measurements, and $m = 30$ is the total number of features in the original feature set. A total of 30 SVM models were built following this scheme.

For each resulting SVM model, we performed a grid search in order to find the best values for the C and σ parameters. The values considered were $C = \{0, 0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 1, 1.5, 2.5\}$ and $\sigma = \{0, 0.01, 0.02, 0.025, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.25, 0.5, 0.75, 0.9\}$.

To evaluate the performance of the classification model, we used k-fold cross validation with $k = 10$. This method splits the data into k subsets and uses $k-1$ fold to train data and one-fold to test data. The relationship between correct and incorrect classifications is organized in a confusion matrix to obtain the measurement performances of accuracy, sensitivity, and specificity. The entire analysis was conducted using R software, version 3.6.2 [28].

Results

Our study subjects consisted of 1,892 white women adults; 793 (41.9%) had the T-score greater than or equal to -1.0 and they were categorized in the normal group, and 1,099 (58.1%) had the T-score lower than -1.0 and were categorized in the bone loss group. Table 1 shows the age, femoral neck BMD, body mass index (BMI), and categorical variables: alcohol drinking, smoking status, arthritis, asthma, cancer history, diabetes, glucocorticoids therapy, heart attack, liver and thyroid disorder of the NHANES participants in two groups divided by the T-score. Participants from the

normal group were, on average, 63.1 ± 8.9 years old, had a BMI of 31.2 ± 6.0 kg/cm², and had a femoral neck BMD of 0.83 ± 0.08 g/cm². Regarding the bone loss group, the mean age was 70.8 ± 9.5 years old, the mean BMI was 26.8 ± 5.3 kg/cm², and the mean femoral neck BMD was 0.63 ± 0.07 g/cm² (SD ± 0.1). The prevalence of cancer (33.7%), liver disorders (5.5%) as well as thyroid disorders (37.8%) were greater in the low BMD group as compared to the normal BMD group.

Table 2 shows geometric means (GMs) and confidence interval 95% (CI95%) of metals and metalloids measured in either whole blood or urine separated by group or combined/overall. Several of the metals in the blood and urine were higher in the bone loss group as compared to the normal group. Indeed, the elements As, Cd and W, had GM (CI95%) urine concentration values of 7.8 (7.3 - 8.3), 0.35 (0.33 - 0.37) and 0.070 (0.067 - 0.074) μ g/L, respectively, for the bone loss group; and 6.6 (6.0 - 7.2), 0.25 (0.24 - 0.28) and 0.057 (0.053 - 0.060) μ g/L, respectively, for normal group. Figure 1 is a graphical representation of the urinary concentration of As, Cd, and W in both groups.

The classification results for the 30 SVM models that were developed are presented in Table 3. The model with the best accuracy (96.46%), sensitivity (95.02%), and specificity (97.47%) was model #19. Here the training subset was formed by Age, BMI, urinary concentration of As, Cd, W, Sb, Ba, Hg, dimethylarsonic acid (DMA), arsenobetaine (AB), Hg, Pt, whole blood concentration of Pb, Hg (total & inorganic), and also arthritis, cancer, thyroid, and former smoker status. However, the model (number #5) which included the five-best features-selected from random forest (Figure 2), where the training subset formed by Age, BMI, urinary concentration of As, Cd, W, also have achieved high scores for accuracy (92.18%), sensitivity (90.50%), and specificity (93.35%). Together, these data demonstrate the importance of these factors

and metals to the classification, since they alone were capable of generating a classification model with a high prediction of accuracy without requiring the other variables. Figure 3 shows the impact of these variables when compared to other models, model #2 has included only age and BMI as variables; model #5 is the model formed by Age, BMI, urinary concentration of As, Cd, W; and model #19 is model which had the best accuracy, sensitivity, and specificity.

Discussion

To the best of our knowledge, this study is the first to evaluate the associations of blood and urinary levels of toxic elements with bone mineral density (BMD) loss in a representative sample of 1,892 individuals with the use of a data mining approach.

In this study, three NHANES cycle databases were mined for general demographic, social, and medical history data from white female participants over 50 years. Additionally, concentrations of 13 metals in blood [Cd, Hg (total and inorganic) and Pb] and urine [As and speciated (arsenobetaine, dimethylarsonic acid), Ba, Cd, Hg, Sb, W] samples were examined and BMD was used to separate the participants into normal (T-score ≥ -1.0) or bone loss (T-score < -1.0) groups. The resulting database underwent SVM modeling to determine which factors could best predict BMD loss.

It is well known that age and BMI are important factors for BMD. Aside from these factors, our modeling process was able to identify that inclusion of three metals (arsenic, cadmium, and tungsten) was also of critical importance in predicting BMD loss. Importantly, participants in the low BMD group had a higher concentration of all these metals in their urine than did the normal BMD group (Figure 1). These findings might clarify a gap regarding the relationship between metal and metalloid exposure

and bone health. Remarkably, higher concentrations of these metals showed significantly higher correlations to lower BMD than did smoking or diabetes, which are well-documented factors leading to bone loss and increased fracture risk [32, 33]. Furthermore, previous investigations have shown impaired bone healing due to smoking and even passive smoking, which highlights the important, yet neglected, impact of heavy metal exposure on bone health [34, 35].

Exposure to arsenic typically results from either consumption via contaminated-arsenic drinking water, soil, and food, or arsenic inhalation in factories [36, 37]. The World Health Organization (WHO) considers that around 200 million people globally are exposed to the metalloid in drinking water at levels above 10 µg/L, the safety threshold [38]. In the present study, the urinary concentration of arsenic was 18.4% higher in the low BMD group compared to the normal BMD group.

Arsenic has the ability to accumulate in bone tissue, likely, competing with the phosphate group to reduce the formation of hydroxyapatite crystals, by instead forming apatite arsenate and other calcium arsenate crystals [39]. Previous studies have shown that exposure to arsenic decreases RANKL and RUNX2 expression, compromising osteoblast maturation, concomitant with reductions in the activity of alkaline phosphatase, as well as the VCAM-I adherence molecule causing a decrease in osteoblastogenesis and osteoblastic activity thus impairing bone remodeling by unbalanced bone turnover [40].

Of interest, clinical studies have shown the use of dental paste containing arsenic trioxide for endodontic treatment of inflamed pulp can cause alveolar bone osteomyelitis and osteonecrosis [41, 42]. In addition, some evidence suggests that Paget's disease, which is caused by an imbalance in bone remodeling, might be associated with arsenic intoxication [43].

Besides, studies have presented a high prevalence of glucose intolerance, diabetes, and metabolic syndrome correlated with arsenic exposure, all of these health disorders are related to high blood glucose levels [21, 44–47]. Remarkably, high glucose concentrations are detrimental to osteocalcin synthesis by the osteoblasts and result in accumulation of AGEs [48], which is linked to higher rates of osteoblast apoptosis and a higher osteoclast resorptive activity [49, 50]. As a result, bone microdamage accumulates, resulting in increased cortical porosity and bone fragility, which may lead to osteoporotic fractures. Moreover, the systemic inflammation associated with these diseases might activate bone resorption, resulting in decreased BMD [51, 52]. β -cell line studies demonstrated the capacity of these cells to methylating inorganic arsenic into monomethylarsenous acid (MMA) and dimethylarsenous acid (DMA). Specifically, MMA can inhibit mitochondrial function and decrease glucose-induced insulin secretion [53–55]. Further, insulin gene expression and transcription factor activities suffer significant effects from arsenic exposure. Arsenic might induce impairment of β -cell function through a decrease of MafA transcriptional activity, such a decrease is an indication of β -cell failure or de-differentiation [56–58].

It has been long known that glucose is a significant substrate for ATP production via glycolysis for osteoblasts and its progenitors [59]. Arsenic has the potential to inhibit ATP production during the process of glycolysis by replacing the phosphate anion with arsenate. This process is called as arsenolysis and might stop feeding osteoblast-mediated bone formation [60–62]. Parallel to glycolysis, glucose-6-phosphate is converted into 6-phosphogluconate using glucose-6-phosphate dehydrogenase via pentose phosphate pathway where NADP⁺ is converted into NADPH which keeps

glutathione in its reduced form. Arsenic can inhibit glucose-6-phosphate dehydrogenase activity and, consequently, reduce glutathione levels [63].

In our study, we found that the urinary concentration of cadmium was 37.6% higher in the low BMD group as compared to that in the normal BMD group. Cadmium is widely distributed in the environment, and the exposure to this metal occurs mainly during the ingestion of food or inhalation of cigarette smoke. It is estimated that over 80% of ingested cadmium comes from cereals, primarily rice and wheat [64].

In vivo studies in experimental animals demonstrate that chronic exposure to cadmium decreases bone volume and increases the percentage of TRAP-positive osteoclast cells in subchondral tibial bone, which can increase bone resorption [65]. Exposure to cadmium may also alter bone formation and mineralization processes since cadmium has been linked to decreased expression of RUNX2, osteocalcin, type I collagen, and alkaline phosphatase, which are markers of osteoblastic differentiation [13].

Clinical studies have shown that even low-level exposure to cadmium through diet and smoking is associated with low BMD and bone fragility in both postmenopausal women and elderly men. Moreover, they also found an association between fractures and cadmium in never-smoker patients, whose central exposure was from their diet. These findings suggest that long-term cadmium exposure has negative consequences on skeletal health [15, 16].

Similar to the arsenic, cadmium might affect the energy metabolism of osteoblasts and osteoblast progenitor cells [62]. Cadmium inhibits enzymes through its high affinity for the free electron pairs in cysteine -SH groups, which are essential in enzyme function. By decreasing the phosphofructokinase activity, cadmium has the potential to limit the glycolysis process in the liver and muscles [66]. We hypothesize

that the same imbalance in energy metabolism might happen in bone differentiation and formation.

Mines and industries are the main occupational sources of human exposure to tungsten, where exposure can be due to pure tungsten, tungsten ore, or tungsten-containing alloys. Workers can be exposed through inhalation or dermal contact of contaminated air. Moreover, tungsten mineral is naturally present in the soil and consumption of contaminated water or air in regions near tungsten mines, industrial sites, or military sites, are also environmental sources. Tungsten particulates in the air can be generated through weathering or emission from industrial and mining sites containing tungsten [67, 68].

The current studies have shown that the urinary concentration of tungsten in the low BMD group was 0.25-fold higher than the normal BMD group. Tungsten has been highlighted as an emerging contaminant, and yet there is limited knowledge of the potential human health risks [69]. *In vivo* evidence suggests that tungsten alters bone homeostasis since young male mice exposed to sodium tungstate (orally for 4 weeks) had enhanced rosiglitazone (PPAR γ ligand)-induced gene expression and adipogenesis [70]. In general, within the bone marrow microenvironment increases in marrow fat usually result in decreases in bone mass. Therefore, the data that tungsten increases adipogenesis may suggest that tungsten increases the commitment of progenitor cells to the adipogenic pathway rather than the osteogenic pathway, which could have significant implications for bone quality.

Bone biology is a complex process consisting of the equilibrium between bone formation and bone resorption. To have a more expansive landscape of how metals and metalloids affect bone remodeling, we have evaluated the effects of multiple metals and metalloids on bone biology by an advanced data mining approach.

Although our model has some limitations, which include the lack of our model validation by applying it to a secondary dataset, and the exclusion of several compounds such as bisphenols, parabens, and phthalates to evaluate the influence of these upon bone health, our findings provide insight into the important impact that arsenic, cadmium, and tungsten have on overall bone health (92.18% of accuracy, 90.50% of sensitivity, and 93.35% of specificity). In general, our data demonstrated the importance of these metals to be classified as risk factors for bone loss since together with age and BMI, they were capable of generating a classification model with a high prediction of accuracy without requiring any other variables.

Declaration of Competing Interest

The authors have no relevant conflicts of interest to disclose.

Funding statements

This study was funded by the São Paulo Research Foundation (FAPESP, no 2018/24069-3).

Author contributions

JPBX, AZ, and FBJ designed the study. JPBX performed the analyses. JPBX, AZ, and FBJ analyzed data and wrote the manuscript. JPBX, AZ, MK, RMB and FBJ interpreted the data and revised the manuscript. All authors approve of the final version of the manuscript

References

1. Cooper C, Campion G, Melton LJ (1992) Hip fractures in the elderly: a world-wide projection. *Osteoporos Int* 2:285–9
2. Rachner TD, Khosla S, Hofbauer LC (2011) Osteoporosis: Now and the future. *Lancet* 377:1276–1287. [https://doi.org/10.1016/S0140-6736\(10\)62349-5](https://doi.org/10.1016/S0140-6736(10)62349-5)
3. Compston JE, Mcclung MR, Leslie WD (2019) Osteoporosis. *Lancet* 393:364–376. [https://doi.org/10.1016/S0140-6736\(18\)32112-3](https://doi.org/10.1016/S0140-6736(18)32112-3)
4. Boyanov M (2009) The diagnosis of osteoporosis and the fracture risk assessment. *Endokrinologiya* 14:30–36
5. ClinRisk. Welcome to the QFracture-2016 risk calculator. <https://qfracture.org/>. Accessed 15 Oct 2019
6. Garvan Institute. Bone Fracture Risk Calculator. <https://www.garvan.org.au/bone-fracture-risk>. Accessed 15 Oct 2019
7. University of Sheffield. Fracture Risk Assessment Tool. <https://www.sheffield.ac.uk/FRAX/>. Accessed 15 Oct 2019
8. Agency for Toxic Substances and Disease Registry (ATSDR) (2017) Substance Priority List. <https://www.atsdr.cdc.gov/SPL/>. Accessed 3 Jan 2020
9. Schümann K, Elsenhans B (2002) The impact of food contaminants on the bioavailability of trace metals. *J. Trace Elem. Med. Biol.* 16:139–144
10. Milnerowicz H, Ściskalska M, Dul M (2015) Pro-inflammatory effects of metals in persons and animals exposed to tobacco smoke. *J. Trace Elem. Med. Biol.* 29:1–10
11. Kim JJ, Kim YS, Kumar V (2019) Heavy metal toxicity: An update of chelating therapeutic strategies. *J. Trace Elem. Med. Biol.* 54:226–231
12. Rodríguez J, Mandalunis PM (2018) A review of metal exposure and its effects

- on bone health. *J Toxicol* 2018:.. <https://doi.org/10.1155/2018/4854152>
13. Brzóska MM, Moniuszko-Jakoniuk J (2005) Disorders in bone metabolism of female rats chronically exposed to cadmium. *Toxicol Appl Pharmacol* 202:68–83. <https://doi.org/10.1016/j.taap.2004.06.007>
 14. Brzóska MM, Moniuszko-Jakoniuk J (2004) Low-level exposure to cadmium during the lifetime increases the risk of osteoporosis and fractures of the lumbar spine in the elderly: Studies on a rat model of human environmental exposure. *Toxicol Sci* 82:468–477. <https://doi.org/10.1093/toxsci/kfh275>
 15. Wallin M, Barregard L, Sallsten G, et al (2016) Low-Level Cadmium Exposure Is Associated with Decreased Bone Mineral Density and Increased Risk of Incident Fractures in Elderly Men: The MrOS Sweden Study. *J Bone Miner Res* 31:732–741. <https://doi.org/10.1002/jbmr.2743>
 16. Engström A, Michaëlsson K, Vahter M, et al (2012) Associations between dietary cadmium exposure and bone mineral density and risk of osteoporosis and fractures among women. *Bone* 50:1372–1378. <https://doi.org/10.1016/j.bone.2012.03.018>
 17. Wani AL, Ara A, Usmani JA (2015) Lead toxicity: A review. *Interdiscip. Toxicol.* 8:55–64
 18. Wong AKO, Beattie KA, Bhargava A, et al (2015) Bone lead (Pb) content at the tibia is associated with thinner distal tibia cortices and lower volumetric bone density in postmenopausal women. *Bone* 79:58–64. <https://doi.org/10.1016/j.bone.2015.05.010>
 19. Levin R, Brown MJ, Kashtock ME, et al (2008) Lead exposures in U.S. children, 2008: Implications for prevention. *Environ. Health Perspect.* 116:1285–1293

20. Bruno A. Rocha, Alexandros G. Asimakopoulos, Masato Honda, Nattane L. da Costa, Rommel M. Barbosa, Fernando Barbosa Jr KK (2017) Advanced data mining approaches in the assessment of urinary concentrations of bisphenols, chlorophenols, parabens and benzophenones in Brazilian children and their association to DNA damage. *Environ Sci Technol*
21. Park SK, Zhao Z, Mukherjee B (2017) Construction of environmental risk score beyond standard linear models using machine learning methods: Application to metal mixtures, oxidative stress and cardiovascular disease in NHANES. *Environ Heal A Glob Access Sci Source* 16:1–17.
<https://doi.org/10.1186/s12940-017-0310-9>
22. Kupsco A, Kioumourtzoglou MA, Just AC, et al (2019) Prenatal Metal Concentrations and Childhood Cardiometabolic Risk Using Bayesian Kernel Machine Regression to Assess Mixture and Interaction Effects. *Epidemiology* 30:263–273. <https://doi.org/10.1097/EDE.0000000000000962>
23. Centers for Disease Control and Prevention (CDC) NC for HS (NCHS). National Health and Nutrition Examination Survey Data, 2009–2010.
<https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2009>. Accessed 3 Jan 2020
24. Centers for Disease Control and Prevention (CDC) NC for HS (NCHS). National Health and Nutrition Examination Survey Data, 2007–2008.
<https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2007>. Accessed 3 Jan 2020
25. Centers for Disease Control and Prevention (CDC) NC for HS (NCHS). National Health and Nutrition Examination Survey Data, 2005–2006.
<https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear>

=2005. Accessed 3 Jan 2020

26. Date AR, Gray AL (1989) Applications of inductively coupled plasma mass spectrometry. Blackie
27. Kursa MB, Rudnicki WR (2010) Feature selection with the boruta package. *J Stat Softw* 36:1–13. <https://doi.org/10.18637/jss.v036.i11>
28. R Core Team (2013) R: A Language and Environment for Statistical Computing. R Found. Stat. Comput.
29. Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20:273–297. <https://doi.org/10.1007/bf00994018>
30. Noble WS (2006) What is a support vector machine? *Nat Biotechnol* 24:1565–1567. <https://doi.org/10.1038/nbt1206-1565>
31. Lo YC, Rensi SE, Torng W, Altman RB (2018) Machine learning in chemoinformatics and drug discovery. *Drug Discov Today* 23:1538–1546. <https://doi.org/10.1016/j.drudis.2018.05.010>
32. Porter SE, Hanley EN (2001) The musculoskeletal effects of smoking. *J. Am. Acad. Orthop. Surg.* 9:9–17
33. Ferrari SL, Abrahamsen B, Napoli N, et al (2018) Diagnosis and management of bone fragility in diabetes: an emerging challenge. *Osteoporos. Int.* 29:2585–2596
34. Castillo RC, Bosse MJ, MacKenzie EJ, et al (2005) Impact of smoking on fracture healing and risk of complications in limb-threatening open tibia fractures. *J Orthop Trauma* 19:151–157. <https://doi.org/10.1097/00005131-200503000-00001>
35. Santiago HAR, Zamarioli A, Sousa Neto MD, Volpon JB (2017) Exposure to Secondhand Smoke Impairs Fracture Healing in Rats. *Clin Orthop Relat Res*

- 475:894–902. <https://doi.org/10.1007/s11999-016-5184-6>
36. Ren X, Mchale CM, Skibola CF, et al (2011) An emerging role for epigenetic dysregulation in arsenic toxicity and carcinogenesis. *Environ. Health Perspect.* 119:11–19
 37. Naujokas MF, Anderson B, Ahsan H, et al (2013) The broad scope of health effects from chronic arsenic exposure: Update on a worldwide public health problem. *Environ. Health Perspect.* 121:295–302
 38. World Health Organization (WHO) (2018) Guidelines for drinking-water quality, 3rd edition: Volume 1 - Recommendations. World Health Organization
 39. Dani SU (2013) Osteoresorptive arsenic intoxication. *Bone* 53:541–545. <https://doi.org/10.1016/j.bone.2013.01.017>
 40. Hu YC, Cheng HL, Hsieh BS, et al (2012) Arsenic trioxide affects bone remodeling by effects on osteoblast differentiation and function. *Bone* 50:1406–1415. <https://doi.org/10.1016/j.bone.2012.03.012>
 41. Marty M, Noirrit-Esclassan E, Diemer F (2016) Arsenic trioxide-induced osteonecrosis treatment in a child: mini-review and case report. *Eur Arch Paediatr Dent* 17:419–422. <https://doi.org/10.1007/s40368-016-0250-z>
 42. Dumlu A, Yalcinkaya S, Olgac V, Güvercin M (2007) Osteomyelitis due to arsenic trioxide use for tooth devitalization. *Int Endod J* 40:317–322. <https://doi.org/10.1111/j.0143-2885.2007.01230.x>
 43. Lever JH (2002) Paget's disease of bone in Lancashire and arsenic pesticide in cotton mill wastewater: A speculative hypothesis. *Bone* 31:434–436. [https://doi.org/10.1016/S8756-3282\(02\)00833-5](https://doi.org/10.1016/S8756-3282(02)00833-5)
 44. Kuo CC, Moon KA, Wang SL, et al (2017) The association of arsenic metabolism with cancer, cardiovascular disease, and diabetes: A systematic

- review of the epidemiological evidence. *Environ. Health Perspect.* 125
45. Kuo CC, Howard B V., Umans JG, et al (2015) Arsenic exposure, arsenic metabolism, and incident diabetes in the strong heart study. *Diabetes Care* 38:620–627. <https://doi.org/10.2337/dc14-1641>
 46. Wang SL, Chiou JM, Chen CJ, et al (2003) Prevalence of non-insulin-dependent diabetes mellitus and related vascular diseases in southwestern arseniasis-endemic and nonendemic areas in Taiwan. *Environ Health Perspect* 111:155–159. <https://doi.org/10.1289/ehp.5457>
 47. Lai MS, Hsueh YM, Chen CJ, et al (1994) Ingested inorganic arsenic and prevalence of diabetes mellitus. *Am J Epidemiol* 139:484–492. <https://doi.org/10.1093/oxfordjournals.aje.a117031>
 48. Saito M, Marumo K (2010) Collagen cross-links as a determinant of bone quality: A possible explanation for bone fragility in aging, osteoporosis, and diabetes mellitus. *Osteoporos. Int.* 21:195–214
 49. Heilmeyer U, Cheng K, Pasco C, et al (2016) Cortical bone laminar analysis reveals increased midcortical and periosteal porosity in type 2 diabetic postmenopausal women with history of fragility fractures compared to fracture-free diabetics. *Osteoporos Int* 27:2791–2802. <https://doi.org/10.1007/s00198-016-3614-7>
 50. Dong XN, Qin A, Xu J, Wang X (2011) In situ accumulation of advanced glycation endproducts (AGEs) in bone matrix and its correlation with osteoclastic bone resorption. *Bone* 49:174–183. <https://doi.org/10.1016/j.bone.2011.04.009>
 51. Rhee EJ, Kim YC, Lee WY, et al (2006) Comparison of insulin resistance and serum high-sensitivity C-reactive protein levels according to the fasting blood

- glucose subgroups divided by the newly recommended criteria for fasting hyperglycemia in 10 059 healthy Koreans. *Metabolism* 55:183–187.
<https://doi.org/10.1016/j.metabol.2005.08.010>
52. Qu Y, Kang MY, Dong RP, Zhao JW (2016) Correlations between abnormal glucose metabolism and bone mineral density or bone metabolism. *Med Sci Monit* 22:824–832. <https://doi.org/10.12659/MSM.895387>
53. Douillet C, Currier J, Saunders J, et al (2013) Methylated trivalent arsenicals are potent inhibitors of glucose stimulated insulin secretion by murine pancreatic islets. *Toxicol Appl Pharmacol* 267:11–15.
<https://doi.org/10.1016/j.taap.2012.12.007>
54. Cui Q, Fu J, Hu Y, et al (2017) Deficiency of long isoforms of Nfe2l1 sensitizes MIN6 pancreatic β cells to arsenite-induced cytotoxicity. *Toxicol Appl Pharmacol* 329:67–74. <https://doi.org/10.1016/j.taap.2017.05.013>
55. Dover EN, Beck R, Huang MC, et al (2018) Arsenite and methylarsonite inhibit mitochondrial metabolism and glucose-stimulated insulin secretion in INS-1 832/13 β cells. *Arch Toxicol* 92:693–704. <https://doi.org/10.1007/s00204-017-2074-y>
56. Shirakawa J, Terauchi Y (2014) Selective and sequential loss of transcriptional factors: A hallmark of β -cell failure in type 2 diabetes? *J. Diabetes Investig.* 5:359–361
57. Gibert Y, Cavarape A, Hesselson D, et al (2019) Braving the Element: Pancreatic β -Cell Dysfunction and Adaptation in Response to Arsenic Exposure. *Front Endocrinol | www.frontiersin.org* 1:344.
<https://doi.org/10.3389/fendo.2019.00344>
58. Sun Q, Yang Q, Xu H, et al (2019) MiR-149 negative regulation of mafA is

- involved in the arsenite-induced dysfunction of insulin synthesis and secretion in pancreatic beta cells. *Toxicol Sci* 167:4. <https://doi.org/10.1093/toxsci/kfy150>
59. Lee WC, Guntur AR, Long F, Rosen CJ (2017) Energy metabolism of the osteoblast: Implications for osteoporosis. *Endocr. Rev.* 38:255–266
 60. Tawfik DS, Viola RE (2011) Arsenate replacing phosphate: Alternative life chemistries and ion promiscuity. *Biochemistry* 50:1128–1134. <https://doi.org/10.1021/bi200002a>
 61. Dixon HBF (1996) The Biochemical Action of Arsonic Acids Especially As Phosphate Analogues. *Adv Inorg Chem* 44:191–227. [https://doi.org/10.1016/S0898-8838\(08\)60131-2](https://doi.org/10.1016/S0898-8838(08)60131-2)
 62. Sabir S, Akash MSH, Fiayyaz F, et al (2019) Role of cadmium and arsenic as endocrine disruptors in the metabolism of carbohydrates: Inserting the association into perspectives. *Biomed. Pharmacother.* 114
 63. A S, A M, A C, DN M (2000) Oxidative Stress in Liver of Mice Exposed to Arsenic-Contaminated Water. *Indian J Gastroenterol* 19:
 64. Reyes-Hinojosa D, Lozada-Pérez CA, Zamudio Cuevas Y, et al (2019) Toxicity of cadmium in musculoskeletal diseases. *Environ Toxicol Pharmacol* 72:103219. <https://doi.org/10.1016/j.etap.2019.103219>
 65. Rodríguez J, Mandalunis PM (2016) Effect of cadmium on bone tissue in growing animals. *Exp Toxicol Pathol* 68:391–397. <https://doi.org/10.1016/j.etp.2016.06.001>
 66. Almeida JA, Novelli ELB, Dal Pai Silva M, Alves Júnior R (2001) Environmental cadmium exposure and metabolic responses of the Nile tilapia, *Oreochromis niloticus*. *Environ Pollut* 114:169–175. [https://doi.org/10.1016/S0269-7491\(00\)00221-9](https://doi.org/10.1016/S0269-7491(00)00221-9)

67. Kraus T, Schramel P, Schaller KH, et al (2001) Exposure assessment in the hard metal manufacturing industry with special regard to tungsten and its compounds. *Occup Environ Med* 58:631–634.
<https://doi.org/10.1136/oem.58.10.631>
68. Bolt AM, Mann KK (2016) Tungsten: an Emerging Toxicant, Alone or in Combination. *Curr Environ Heal reports* 3:405–415.
<https://doi.org/10.1007/s40572-016-0106-z>
69. EPA (2017) Technical Fact Sheet – Tungsten. In: EPA 505-F-070- 005.
70. Bolt AM, Grant MP, Wu TH, et al (2016) Tungsten promotes sex-specific adipogenesis in the bone by altering differentiation of bone marrow-resident mesenchymal stromal cells. *Toxicol Sci* 150:333–346.
<https://doi.org/10.1093/toxsci/kfw008>

Figures legends

Figure 1. Distribution of the concentration level of urine heavy metals across the two groups: normal and bone loss groups based on femoral hip T-score.

Figure 2. Bar plots of relative importance of the compounds and other variables on BMD according to the Random Forest values computed.

Figure 3. Receiver Operating Characteristics Curve (ROC) traces the percentage of true positives to be accurately predicted, by a given logit model, as the prediction probability cutoff is lowered from 1 to 0.

Tables

Table 1. Characteristics of the study population overall and by groups.

	Normal (N = 793)	Bone Loss (N = 1,099)	Overall (N = 1,892)
CONTINUOUS mean (SD)			
Age (years)	63.1 ± 8.9	70.8 ± 9.5	67.6 ± 10.0
BMD (mg/cm ²)	0.8 ± 0.1	0.6 ± 0.1	0.7 ± 0.1
BMI (kg/cm ²)	31.2 ± 6.0	26.8 ± 5.3	28.7 ± 6.0
CATEGORICAL number (%)			
Heavy alcohol use	13 (1.6)	0 (0)	13 (0.7)
Smoking Status			
Never	399 (50)	599 (55)	998 (53)
Former	274 (35)	354 (32)	628 (33)
Current	120 (15)	146 (13)	266 (14)
Arthritis	458 (58)	626 (57)	1084 (57)
Asthma	178 (22)	191 (17)	369 (20)
Cancer historical	162 (20)	370 (34)	532 (28)
Diabetes	52 (7)	41 (4)	93 (5)
Glucocorticoid	38 (5)	58 (5)	96 (5)
Heart attack	55 (7)	99 (9)	154 (8)
Liver disorder	19 (2)	61 (6)	80 (4)
Thyroid disorder	356 (45)	416 (38)	772 (41)

BMD = femoral neck bone mineral density and BMI = body mass index.

Table 2. Geometric means and confidence interval 95% (CI 95%) of metals and metalloids overall and by groups.

	Normal (N = 793)	Bone Loss (N = 1,099)	Overall (N = 1,892)
In whole blood (µg/L)			
Cadmium	0.43 (0.41 - 0.46)	0.5 (0.4 - 0.5)	0.4 (0.4 - 0.5)
Lead	1.5 (1.5 - 1.6)	1.7 (1.7 - 1.8)	1.6 (1.6 - 1.7)
Total Mercury	1.0 (0.9 - 1.1)	0.9 (0.9 - 1.0)	0.9 (0.9 - 1.0)
Inorganic Mercury	0.30 (0.29 - 0.31)	0.31 (0.30 - 0.32)	0.31 (0.30 - 0.31)
In urine (µg/L)			
Antimony	0.047 (0.045 - 0.049)	0.052 (0.050 - 0.055)	0.050 (0.048 - 0.052)
Total Arsenic	6.6 (6.0 - 7.2)	7.8 (7.3 - 8.3)	7.2 (6.9 - 7.6)
Arsenobetaine	1.5 (1.4 - 1.8)	1.5 (1.4 - 1.7)	1.5 (1.4 - 1.7)
Dimethylarsonic acid	2.8 (2.6 - 2.9)	3.2 (3.0 - 3.3)	3.0 (2.9 - 3.1)
Barium	1.2 (1.1 - 1.2)	1.3 (1.2 - 1.3)	1.2 (1.2 - 1.3)
Cadmium	0.25 (0.24 - 0.28)	0.35 (0.33 - 0.37)	0.30 (0.29 - 0.32)
Mercury	0.36 (0.33 - 0.38)	0.41 (0.39 - 0.44)	0.39 (0.38 - 0.41)
Platinum	0.0070 (0.0066 - 0.0074)	0.0080 (0.0074 - 0.0085)	0.0076 (0.0072 - 0.0079)
Tungsten	0.057 (0.053 - 0.060)	0.070 (0.067 - 0.074)	0.064 (0.062 - 0.067)

1 Table 3. Accuracy, sensitivity and specificity values obtained by the SVM model trained on different feature subsets

n ²	Model	Accuracy	Sensitivity	Specificity
01	Age	0.64	0.57	0.69
02	Age + BMI	0.71	0.60	0.79
03	Age + BMI + As	0.79	0.78	0.79
04	Age + BMI + As + Cd	0.87	0.82	0.90
05	Age + BMI + As + Cd + W	0.92	0.90	0.93
06	Age + BMI + As + Cd + W + Pb [#]	0.92	0.89	0.93
07	Age + BMI + As + Cd + W + Pb [#] + Hg [#]	0.94	0.91	0.96
08	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb	0.95	0.93	0.96
09	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#]	0.94	0.92	0.96
10	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba	0.95	0.95	0.94
11	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg	0.95	0.92	0.97
12	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA	0.95	0.93	0.96
13	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA + AB	0.95	0.92	0.97
14	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA + AB + iHg [#]	0.95	0.94	0.96
15	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA + AB + iHg [#] + Pt	0.95	0.94	0.96
16	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA + AB + iHg [#] + Pt + arthritis	0.96	0.93	0.98
17	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA + AB + iHg [#] + Pt + arthritis + cancer	0.96	0.93	0.98
18	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA + AB + iHg [#] + Pt + arthritis + cancer + thyroid	0.96	0.94	0.98
19	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA + AB + iHg [#] + Pt + arthritis + cancer + thyroid + former smoker	0.96	0.95	0.97
20	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA + AB + iHg [#] + Pt + arthritis + cancer + thyroid + former smoker + never smoker	0.96	0.93	0.98
21	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA + AB + iHg [#] + Pt + arthritis + cancer + thyroid + former smoker + never smoker + smoking	0.95	0.93	0.97
22	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA + AB + iHg [#] + Pt + arthritis + cancer + thyroid + former smoker + never smoker + smoking + asthma	0.96	0.94	0.97
23	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA + AB + iHg [#] + Pt + arthritis + cancer + thyroid + former smoker + never smoker + smoking + asthma + liver	0.96	0.95	0.97
24	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA + AB + iHg [#] + Pt + arthritis + cancer + thyroid + former smoker + never smoker + smoking + asthma + liver + heart attack	0.95	0.93	0.9
25	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA + AB + iHg [#] + Pt + arthritis + cancer + thyroid + former smoker + never smoker + smoking + asthma + liver + heart attack + diabetes	0.95	0.94	0.97
26	Age + BMI + As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA + AB + iHg [#] + Pt + arthritis + cancer + thyroid + former smoker + never smoker + smoking + asthma + liver + heart attack + diabetes + heavy drink	0.95	0.94	0.97
27	As	0.65	0.32	0.89
28	As + Cd	0.66	0.45	0.81
29	As + Cd + W	0.75	0.62	0.83
30	As + Cd + W + Pb [#] + iHg [#] + Sb + Cd [#] + Ba + Hg + DMA + AB + iHg [#] + Pt	0.95	0.91	0.97

2

3 BMI = body mass index; DMA = dimethylarsonic acid, AB = arsenobetaine, iHg = total mercury and iHg = inorganic mercury

4 # = whole blood concentration

Groups  Normal  Bone Loss





