Online Supplementary Material

Urinary Metals and Incident Diabetes in Midlife Women: Study of Women's Health Across the Nation (SWAN)

Xin Wang[a], Carrie A. Karvonen-Gutierrez[a], William H. Herman[b], Bhramar Mukherjee[c], Siobán D. Harlow[a], Sung Kyun Park[a,d]

[a] Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI

[b] Department of Internal Medicine, University of Michigan, Ann Arbor, MI

[c] Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI

[d] Department of Environmental Health Sciences, School of Public Health, University of Michigan, Ann Arbor, MI

Corresponding author: Sung Kyun Park, Department of Epidemiology, University of Michigan, M5541 SPH II, 1415 Washington Heights, Ann Arbor, Michigan 48109-2029. Fax: (734) 936-2084. Phone: (734) 936-1719. E-mail: sungkyun@umich.edu.

Supplementary Methods

Urinary metals

Urine specimens were collected prior to 11 am. First morning voided urine was collected. Aliquoted specimens were frozen and stored in ultra-low freezers at -80 ˚C until they were later analyzed for the metal content. All specimens were collected and stored in the SWAN Repository (http://swanrepository.com/) using a systematic protocol. Baseline concentrations of the following 20 metals including total arsenic, barium, beryllium, cadmium, cobalt, chromium, cesium, copper, mercury, manganese, molybdenum, nickel, lead, antimony, tin, thallium, uranium, vanadium, tungsten, and zinc were measured in these urine samples using high-resolution inductively coupled plasma-mass spectrometry (ICP-MS) (Thermo Scientific iCAP RQ, Waltham, MA) following the CDC method 3018.3,[1] with modifications for the expanded metals panel, by the Applied Research Center of NSF International (Ann Arbor, Michigan). All standards, quality controls (QCs), blanks, rinse solution and urine samples were diluted 10-fold in a diluent consisting of 2% HNO_3 solution containing the internal standards and gold. The samples were analyzed in two analysis modes - standard (default) for the majority of the metals, and kinetic energy discrimination (KED) for vanadium, chromium, arsenic, molybdenum and cadmium. The following QC procedures was conducted in parallel with sample analyses: (a) second source standards and spike and surrogate recoveries were tested periodically; (b) linearity and drift checks were performed with each sample batch; and (c) metal internal standards were used on each sample. Each sample run contained a minimum of 4 calibration standards and a blank. The coefficients of variation were 2.4-34.8% for the low QC pools; 1.6-4.0% for the high QC pools; and 1.8-4.0% for the laboratory fortified blank. The limits of detection (LODs) of each metal were determined during the method validation by running a dilution matrix blank 10

times and then calculating the standard deviation of the instrument response. The limit of detection was then defined by calculating three times the standard deviation.

Inverse probability weighting

In our study, selection bias may exist, as selection into SWAN the Multi-Pollutant Substudy were probably affected by metal exposures, their related diabetes risk factors, or potential confounders before or at the time of enrollment. On the other hand, selective loss to follow-up or other forms of attrition that occur after metal measurements may also bias estimates of associations between metals and diabetes if continuation in the follow-up is influenced by metal exposures and risk factors of diabetes. We addressed these two types of bias by using the inverse probability weighting (IPW).

I. Selective participation in the SWAN multi-pollutant substudy

We used Repository samples available from the third SWAN follow-up visit (visit 03, 1999-2000) for metal measurements in our analysis. Women enrolled in SWAN were between the age of 42 to 52 years at the SWAN baseline (visit 00, 1996-1997), which marked a time of increased risk for diabetes.[2] Some of women who were at high risk of diabetes at the SWAN baseline have been censored before visit 03. Thus, participants susceptible to developing diabetes at the time of metal measurements were possibly different from the source population. On the other hand, at visit 03, only a subpopulation with 1,400 SWAN participants, but not all participants remained in the cohort had urine samples stored in the SWAN Biorepository assayed for metal concentration determinations. In this way, the analysis based on these 1,400 participants is likely to be susceptible to bias attributable to the selective participation in the substudy as shown in the directed acyclic graphs (DAG) (**supplementary Figure 2**).

In the **supplementary Figure 2**, DM represents incidence of diabetes. E_0 and E_3 represent measures of urinary metals at SWAN baseline (visit 00) and third follow-up visit (visit 03). We measure metals only at visit 03, so E_0 is unobserved. Considering the environmental exposure to metals at one time point is often reasonably correlated with the exposure at other time points, we consider an effect emanating from E_0 to E_3 and terminating in DM to represent a causal effect of metals on diabetes for the purposes of identifying potential bias in our DAG. RF represents metal induced health effects which may affect continuation in the SWAN up to visit 03, substudy participation, and diabetes. L represents both time-fixed and time-varying covariates which may influence both diabetes risks and selection. S_3 with a box drawn around represents remaining uncensored and free of diabetes up to visit 03. S_U with a box drawn around represents urinary metals substudy participation. Selection bias can be found in the DAG. For example, conditioning on S_U opens the path $RF_3 \rightarrow S_U \leftarrow L_3$, introducing an association between E_3 and DM ($E_3 \rightarrow RF_3 \rightarrow S_U \leftarrow L_3 \rightarrow DM$) which is not causal. At the same time, conditioning on S_U blocks some of the association that goes from E_3 to diabetes through RF_3, because conditioning on S_U partially conditions on RF_3.

IPW was used to alleviate the potential bias resulting from the selection into the SWAN multi-pollutant substudy.[3] IPW uses information available for participants with and without metal measurements to weight observations from participants with metal measurements, so that the weighted subpopulation is representative of all SWAN participants in the original cohort who were continuing in the cohort and were free of diabetes at the time of metal measurements (visit 03). Probability of continuation in the follow-up study up to visit 03 and probability of selection into the substudy given that participants were not censored at visit 03 were modeled separately. We estimated the probability of continuing in the study up to visit 03 by using pooled logistic

regression,[4] conditional on covariates (RF, and L) and on being uncensored at the previous visit, which equals to $\prod_{k=1}^{i} \Pr[C_{ik} = 0 \mid C_{i(k-1)} = 0, L_{i(k-1)}, RF_{i(k-1)}]$, where k represents the kth visit (01-03) and C is the censoring indicator. Given the large number of possible predictors among the relative to the number of persons who dropped out of the study, we used forward selection to inform the variables included in the final models, including age, race/ethnicity, study site, education level, marital status, husband's employment status, smoking, menopausal status, self-rated health, and diagnose of heart attack or angina. The reciprocal of this cumulative probability (W_1) is the weight of remaining free of diabetes and in the study for individual *i* at visit 03. For the probability of selection into the substudy given that participants were not censored at visit 03, we used a single logistic regression model to predict the probability,[5] which equals to $\Pr[U_i = 1 \mid C_{i,3} = 0, L_{i,3}, RF_{i,3}]$, where U indicates the selection into substudy. Variables included in the final logistic model were determined through forward selection, including age, study site, education level, smoking, menopausal status, total cholesterol level, low density lipoprotein cholesterol level, triglyceride level, and hypertension. The reciprocal of this probability (W_2) is the weight of being selected into the substudy at visit 03. Finally, we calculated a combined weight $W_{substudy} = W_1 \times W_2$, as the inverse of the probability of the conjunction of these two events.

II. Selective attrition after metal measurements

We hypothesized that women with higher concentrations of toxic metals would experience higher risk of diabetes during 15 years of follow-up after metal measurements. However, given the toxicity of metals such as arsenic, those with high concentrations of toxic metals who remained in the cohort might have other beneficial characteristics (healthier) that protected them from developing diabetes. This is because the risk factors for diabetes especially

those health conditions predict the censoring or attrition after the metal measurements. Studies of toxic metals that are also themselves associated with substantial attrition through the related adverse health outcomes correlated with diabetes. In this way, the selection induces an association between metals and diabetes, even if there is no true effect (see DAG in **supplementary Figure 3**).

Same symbols (E, RF, L, DM) as those in **supplementary Figure 2** were used in **supplementary Figure 3** to represent the same type of variables in the DAG. S with a box drawn around represents continuation in the SWAN study at each visit after visit 03. S is a collider on which we condition through the restriction of our analysis to those remained in the cohort at each visit. Therefore, statistical associations, for example, between E_3 and L_3, RF_3 and L_3, are induced via conditioning on S_4. E_0 is then connected to DM through paths that do not emanate from E_3, such as $E_3 \rightarrow RF_3 \rightarrow S_4 \leftarrow L_3 \rightarrow DM$, which is noncausal. On the other hand, if continuation in SWAN is at least partly driven by RF, then the continuation in the cohort effectively conditions on RF, resulting in bias from conditioning on an intermediate between metal and diabetes.

Similar to the strategy we used to address the selective participation, IPW was used to reduce potential bias resulting from the selective attrition. The intuition behind these weights is that participants with characteristics similar to the observations missing due to attrition are upweighted, so as to represent their original contribution as well as their missing contributions. We modeled and estimate the probability of continuing in the study after visit 03 by using pooled logistic regression, conditional on covariates (RF and L) and on being uncensored at the previous visit, which equals to $\prod_{k=4}^{j} \Pr[C_{ik} = 0 \mid C_{i(k-1)} = 0, E_i, L_{i(k-1)}, RF_{i(k-1)}, Z_i]$, where k represents the k^{th} visit (04-15) and C is the censoring indicator. Age, study site, SWAN visit number,

household income, smoking, use of hormone, self-rated health, BMI (linear and quardratic terms), and waist circumference (linear and quardratic terms) were included in the final logistic model after forward selection. The reciprocal of this cumulative probability of continuing is the non-stabilized weight ($W_{attrition}$). And the weight was applied at the level of observations within individuals.

III. Combine IPWs for selective participation and selective attrition

We calculated the total $W = W_{substudy} \times W_{atrrition}$ for each participant in the metals-diabetes analysis, as the inverse of the probability of being selected into the SWAN multi-pollutant substudy from the original SWAN cohort and of being uncensored up to a given study visit after metal measurements. To note, including L in the calculation of the weight is not sufficient to control for confounding when evaluating associations between metals and diabetes incidence, and as such, the potential confounders were adjusted as covariates in the Cox proportional hazards model in our primary analysis.[3,4]

Supplementary Table 1. Hazard ratios for diabetes for two-fold increase in specific gravity corrected-urinary metal concentrations.

Metals	**Initial model[a]**		**Full Model[b]**	
	HR (95% CI)	**P-value**	**HR (95% CI)**	**P-value**
Arsenic	1.06 (0.98, 1.15)	0.14	1.19 (1.09, 1.30)	<.0001
Barium	0.97 (0.87, 1.08)	0.56	0.95 (0.84, 1.07)	0.38
Cadmium	0.99 (0.89, 1.10)	0.82	0.94 (0.85, 1.06)	0.29
Cobalt	0.95 (0.83, 1.09)	0.47	0.99 (0.86, 1.13)	0.87
Cesium	1.09 (0.89, 1.33)	0.39	1.18 (0.96, 1.46)	0.12
Copper	1.04 (0.89, 1.21)	0.63	0.94 (0.80, 1.11)	0.44
Mercury	0.83 (0.75, 0.91)	0.02	0.91 (0.81, 1.01)	0.47
Manganese	1.11 (0.95, 1.33)	0.27	1.11 (0.91, 1.34)	0.31
Molybdenum	0.92 (0.80, 1.06)	0.25	1.01 (0.87, 1.17)	0.88
Nickel	1.06 (0.91, 1.24)	0.43	1.12 (0.95, 1.31)	0.18
Lead	1.11 (0.98, 1.26)	0.10	1.19 (1.04, 1.36)	0.01
Antimony	1.02 (0.90, 1.15)	0.78	1.05 (0.92, 1.19)	0.50
Tin	1.10 (1.01, 1.20)	0.04	1.13 (1.03, 1.24)	0.01
Thallium	1.04 (0.94, 1.15)	0.44	1.02 (0.91, 1.14)	0.71
Zinc	1.48 (1.26, 1.72)	<.0001	1.28 (1.08, 1.50)	0.004

Note: all models were constructed by Cox proportional hazards model.

[a] Initial model: adjustment for age, race/ethnicity, study sites.

[b] Full model: initial model with additional adjustment for education, household income, body mass index (baseline level), waist circumference (baseline level), smoking status, alcohol consumption, physical activity score, total energy intake, menopausal status, and use of hormone. In full model, seafood and rice intake was additionally adjusted for arsenic, cadmium, and mercury models; zinc intake from diets and supplements was additionally adjusted for zinc model.

Supplementary Table 2. Hazard ratios for diabetes for two-fold increase in specific gravity corrected-urinary metal concentrations after additional adjustments.

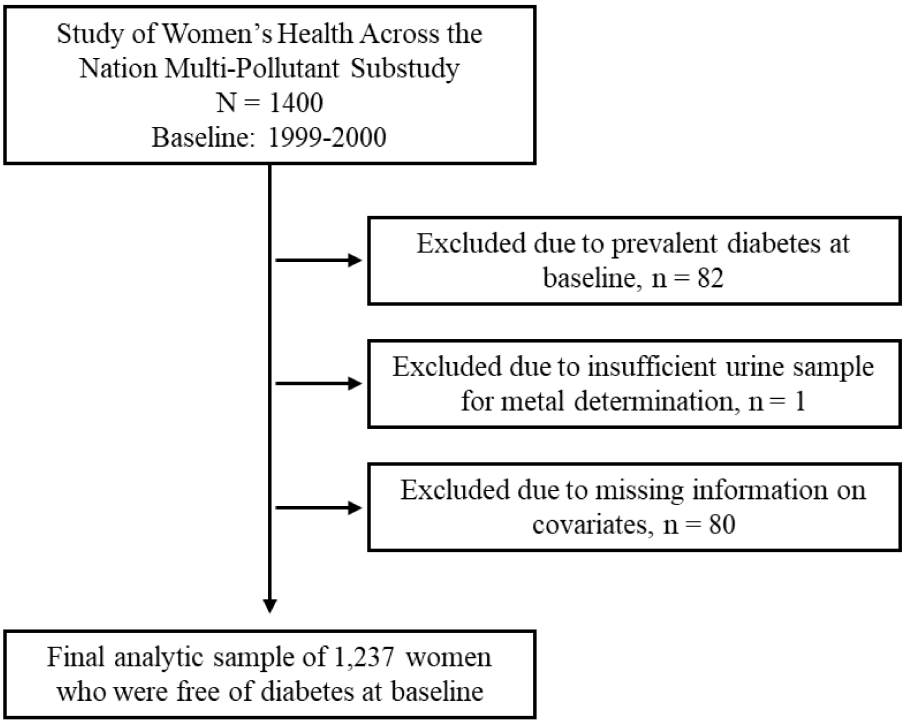
Metals	Full model[a]		Full model with additional adjustment for family history of diabetes and baseline levels of blood pressure and lipids[b]		Full model with additional adjustment for family history of diabetes, baseline levels of blood pressure, lipids, and fasting glucose[c]	
	HR (95% CI)	P-value	HR (95% CI)	P-value	HR (95% CI)	P-value
Arsenic	1.19 (1.10, 1.30)	<0.0001	1.24 (1.14, 1.35)	<0.0001	1.20 (1.10, 1.32)	<0.0001
Barium	0.96 (0.85, 1.09)	0.53	0.99 (0.87, 1.12)	0.83	0.97 (0.85, 1.10)	0.58
Cadmium	0.96 (0.86, 1.07)	0.42	0.96 (0.86, 1.07)	0.47	0.97 (0.87, 1.08)	0.59
Cobalt	1.01 (0.88, 1.15)	0.90	1.00 (0.87, 1.14)	0.98	0.93 (0.81, 1.07)	0.30
Cesium	1.23 (0.98, 1.50)	0.06	1.23 (1.00, 1.51)	0.05	1.22 (0.99, 1.50)	0.06
Copper	0.96 (0.82, 1.13)	0.65	0.96 (0.82, 1.14)	0.66	0.99 (0.84, 1.17)	0.89
Mercury	0.92 (0.82, 1.03)	0.12	0.92 (0.83, 1.03)	0.14	0.93 (0.83, 1.04)	0.18
Manganese	1.10 (0.90, 1.35)	0.33	1.10 (0.89, 1.35)	0.38	1.14 (0.92, 1.41)	0.23
Molybdenum	1.04 (0.90, 1.21)	0.58	1.05 (0.90, 1.22)	0.56	1.01 (0.87, 1.17)	0.91
Nickel	1.15 (0.98, 1.35)	0.08	1.15 (0.98, 1.35)	0.10	1.17 (0.99, 1.38)	0.06
Lead	1.20 (1.05, 1.37)	0.006	1.23 (1.08, 1.40)	0.002	1.24 (1.09, 1.41)	0.001
Antimony	1.07 (0.93, 1.22)	0.36	1.04 (0.91, 1.20)	0.55	1.09 (0.94, 1.26)	0.26
Tin	1.11 (1.01, 1.22)	0.04	1.11 (1.01, 1.22)	0.04	1.10 (1.00, 1.22)	0.05
Thallium	1.04 (0.93, 1.16)	0.52	1.04 (0.92, 1.16)	0.54	1.10 (0.99, 1.23)	0.09
Zinc	1.31 (1.11, 1.53)	0.001	1.31 (1.11, 1.55)	0.001	1.33 (1.13, 1.57)	0.0008

Note: all models were constructed by Cox proportional hazards model.

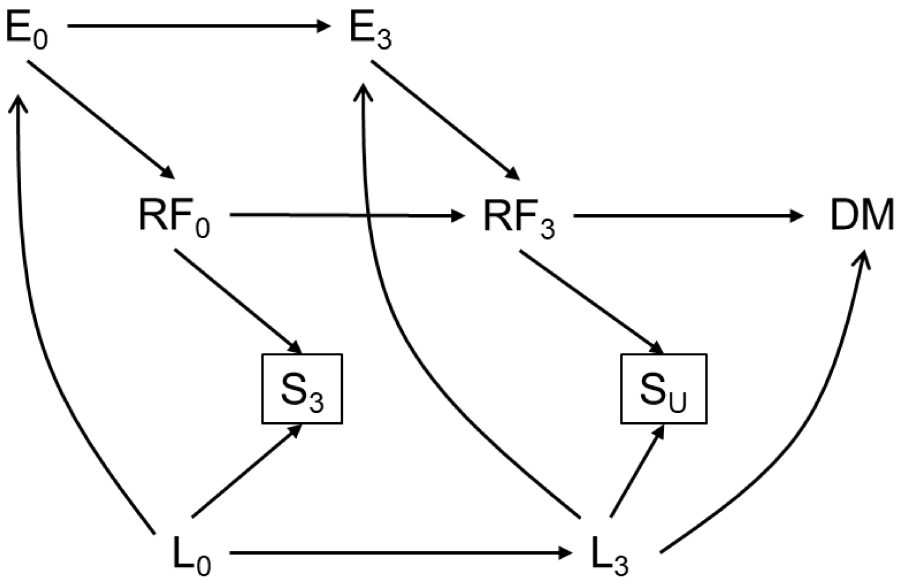
[a] Adjustment for education, household income, body mass index (baseline level), waist circumference (baseline level), smoking status, alcohol consumption, physical activity score, total energy intake, menopausal status, and use of hormone. For arsenic, cadmium and mercury, seafood and rice intake was additionally adjusted for; zinc intake from diets and supplements was additionally adjusted for zinc model.

[b] Additional adjustments for baseline levels of systolic blood pressure, total cholesterol level, high-density lipoprotein cholesterol level, triglyceride level, and family history of diabetes.

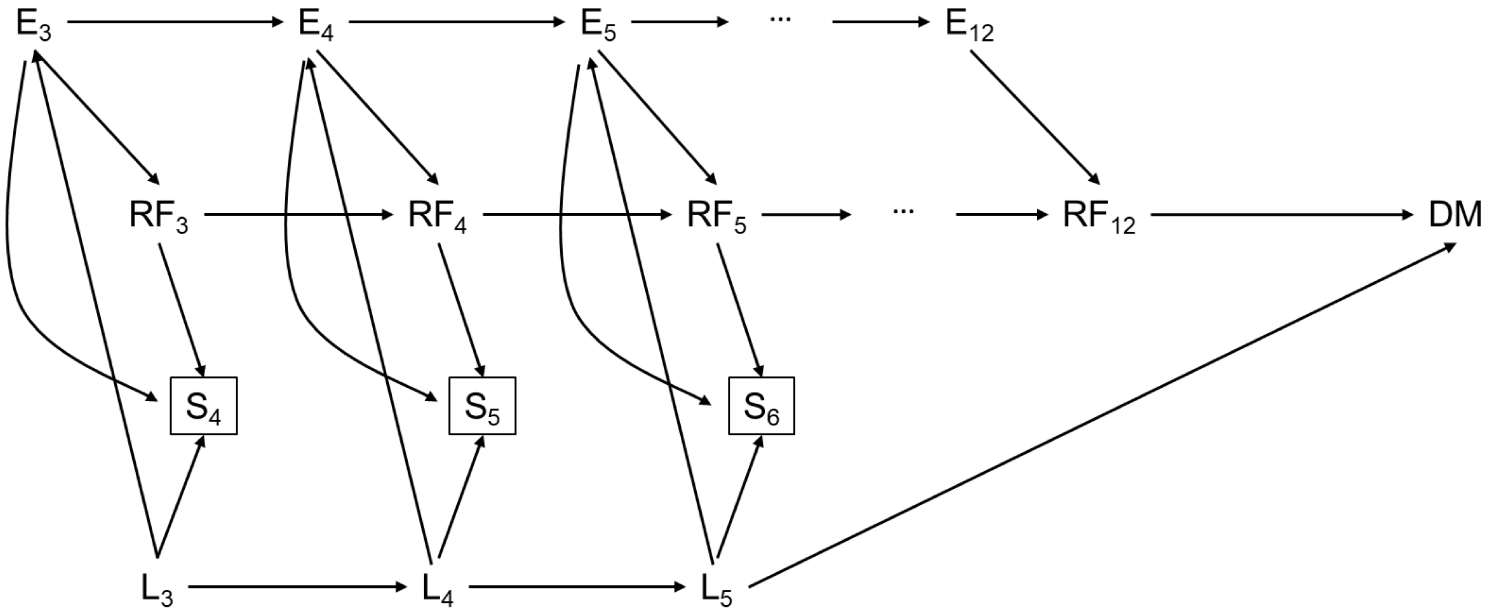
[c] Additional adjustments for baseline levels of systolic blood pressure, total cholesterol level, high-density lipoprotein cholesterol level, triglyceride level, and family history of diabetes, and fasting glucose.



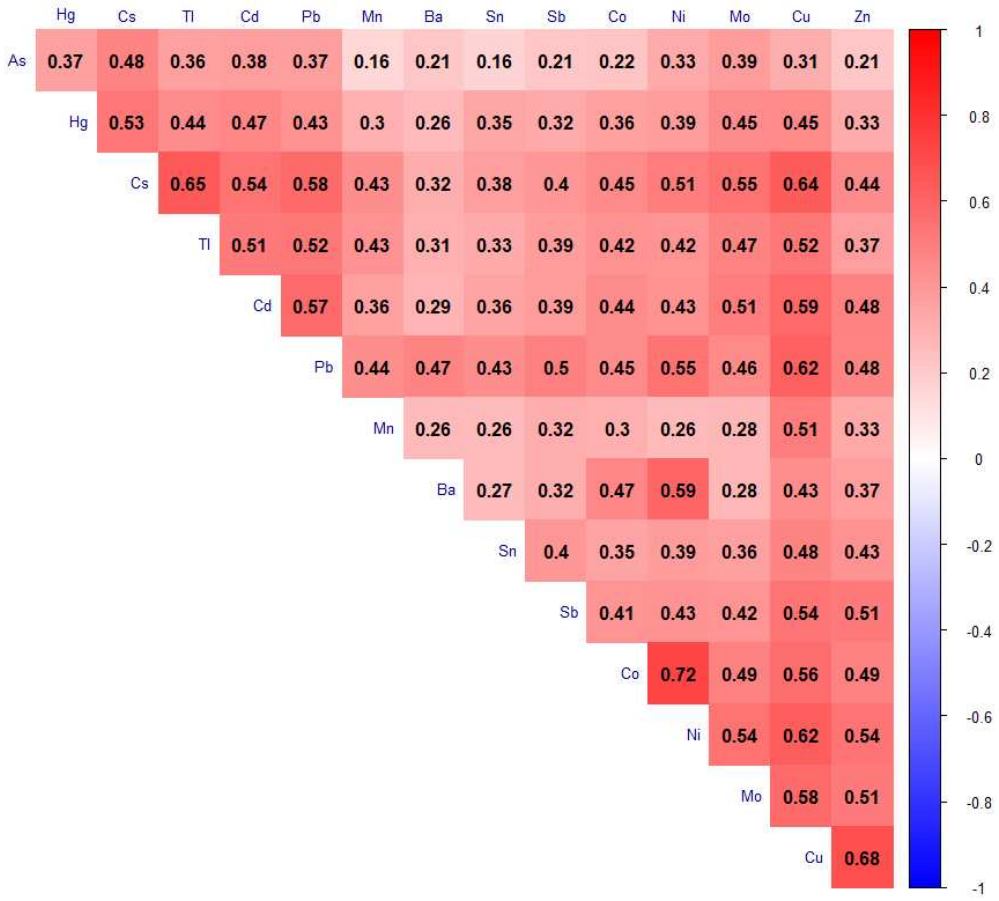
Supplementary Figure 1. Schematic diagram of analytic sample.



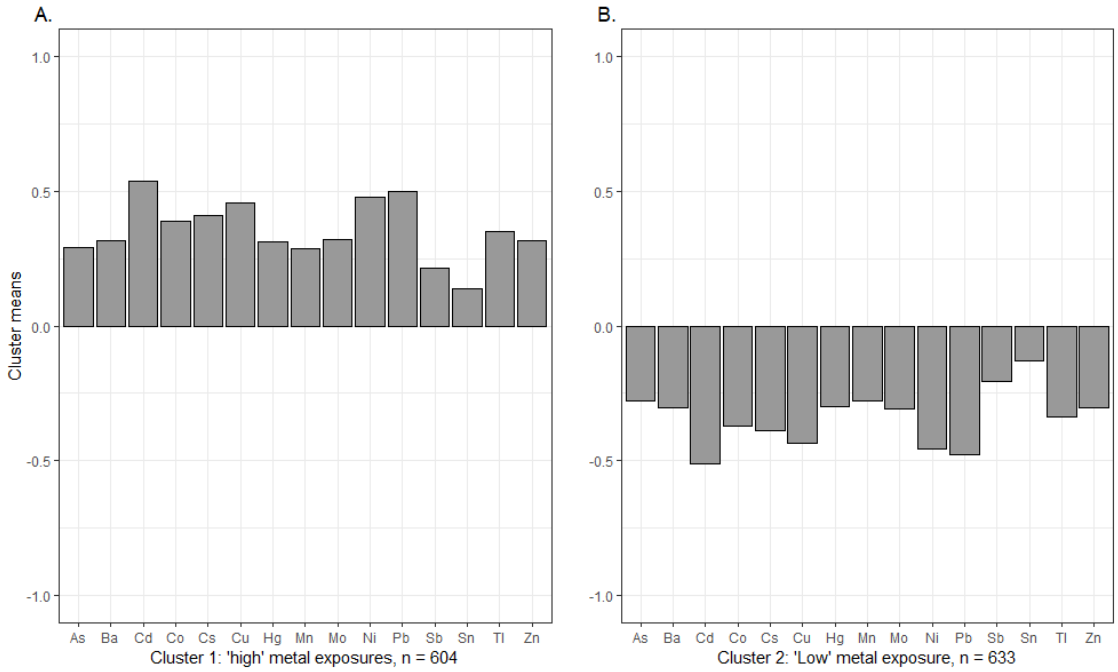
Supplementary Figure 2. Directed acyclic graphs illustrating selective participation in the Study of Women's Health Across the Nation Multi-Pollutant Substudy.



Supplementary Figure 3. Directed acyclic graphs illustrating selective attrition after baseline of the Study of Women's Health Across the Nation Multi-Pollutant Substudy.



Supplementary Figure 4. Spearman correlation matrix of metal concentrations. As: arsenic, Ba: barium, Cd: cadmium, Co: cobalt, Cs: cesium, Cu: copper, Hg: mercury, Mn: manganese, Mo: molybdenum, Ni: nickel, Pb: lead, Sb: antimony, Sn: tin, Tl: thallium, Zn: zinc.



Supplementary Figure 5. Cluster means of the 15 standardized log-transformed urinary metals using k-means clustering method. Y-axis (cluster means) represents the mean standardized log-transformed specific gravity adjusted metal concentrations. Cluster 1: “high” exposure pattern to metal mixtures; cluster 2: “low” exposure pattern to metal mixtures. As: arsenic, Ba: barium, Cd: cadmium, Co: cobalt, Cs: cesium, Cu: copper, Hg: mercury, Mn: manganese, Mo: molybdenum, Ni: nickel, Pb: lead, Sb: antimony, Sn: tin, Tl: thallium, Zn: zinc.

References

1 CDC. Laboratory Procedure Manual, Multi-Element in urine. NHANES 2011-2012. 2012.

2 Kim C. Does menopause increase diabetes risk? Strategies for diabetes prevention in midlife women. *Womens Health (Lond Engl)* 2012;**8**:155–67.

3 Hernán MA, Hernández-Díaz S, Robins JM. A Structural Approach to Selection Bias. *Epidemiology* 2004;**15**:615–25.

4 Weuve J, Tchetgen Tchetgen EJ, Glymour MM, *et al.* Accounting for Bias Due to Selective Attrition: the example of smoking and cognitive decline. *Epidemiology* 2012;**23**:119–28.

5 Weisskopf MG, Sparrow D, Hu H, *et al.* Biased Exposure–Health Effect Estimates from Selection in Cohort Studies: Are Environmental Studies at Particular Risk? *Environ Health Perspect* 2015;**123**:1113–22.