Supplementary Materials

Environment-Wide Association Study of Chemical Biomarkers and Health Outcomes in NHANES 2017–2018

Hayden Farquhar

February 5, 2026

1 Table S1. All 26 FDR-Significant Associations from the Discovery Phase

Survey-weighted linear regression of log-transformed chemical biomarker on health outcome, adjusted for age, sex, race/ethnicity, poverty–income ratio, BMI (when not the outcome), and smoking status. 95% CIs calculated as $\beta \pm 1.96 \times \text{SE}$.

Chemical	Outcome	Class	β	95% CI	SE	P-value	FDR	N
Blood manganese	RBC count	Heavy metals	0.23	(0.19, 0.27)	0.020	9.8e-06	0.024	4,873
Blood selenium	Hemoglobin	Heavy metals	2.16	(1.74, 2.58)	0.215	2.0e-05	0.024	4,873
Urinary perchlorate	BUN	Urinary elements	1.21	(0.97, 1.45)	0.124	2.5e-05	0.024	1,579
MEOHP (phthalate)	Total bilirubin	Phthalates	-0.06	(-0.07, -0.05)	0.006	4.0e-05	0.024	1,580
Methylmercury	Alk Phosphatase	Heavy metals	-2.84	(-3.47, -2.21)	0.321	4.8e-05	0.024	4,834

Chemical	Outcome	Class	β	95% CI	SE	P-value	FDR	N
Blood manganese	Waist circumference	Heavy metals	7.16	(5.33, 8.99)	0.934	5.9e-05	0.024	4,693
MEHHP (phthalate)	Total bilirubin	Phthalates	-0.06	(-0.08, -0.05)	0.006	6.1e-05	0.024	1,580
Blood selenium	Total cholesterol	Heavy metals	39.59	(30.02, 49.15)	4.880	8.3e-05	0.029	4,855
Blood manganese	BMI	Heavy metals	2.59	(1.87, 3.30)	0.367	1.1e-04	0.030	4,876
Blood lead	Waist circumference	Heavy metals	-4.35	(-5.58, -3.13)	0.624	1.2e-04	0.030	4,693
Urinary lead	eGFR	Urinary metals	6.55	(5.09, 8.01)	0.746	1.2e-04	0.030	1,589
Blood lead	BMI	Heavy metals	-1.95	(-2.51, -1.38)	0.289	1.5e-04	0.030	4,876
Blood mercury (total)	Alk Phosphatase (log)	Heavy metals	-0.04	(-0.05, -0.03)	0.005	1.5e-04	0.030	4,833
DMA (urinary)	Uric acid	Urinary metals	0.20	(0.15, 0.26)	0.027	1.5e-04	0.030	1,593
Urinary cesium	eGFR	Urinary metals	9.78	(7.48, 12.08)	1.174	1.6e-04	0.030	1,589
Glyphosate (serum)	BUN	Surplus serum	1.22	(0.89, 1.55)	0.170	1.8e-04	0.031	1,371

Chemical	Outcome	Class	β	95% CI	SE	P-value	FDR	N
Blood selenium	RBC count	Heavy metals	0.55	(0.40, 0.70)	0.077	1.9e-04	0.031	4,873
Blood lead	Total choles-terol	Heavy metals	9.10	(6.55, 11.65)	1.300	2.1e-04	0.032	4,855
Urinary thallium	eGFR	Urinary metals	7.68	(5.78, 9.59)	0.970	2.2e-04	0.032	1,589
Methylmercury	Waist circum-ference	Heavy metals	-1.78	(-2.34, -1.23)	0.282	2.3e-04	0.032	4,693
Oxychlordane	eGFR	VOC metabo-lites	3.23	(2.39, 4.06)	0.428	2.8e-04	0.036	1,589
Urinary cadmium	BUN	Urinary metals	-1.02	(-1.29, -0.76)	0.137	3.0e-04	0.036	1,589
Glyphosate (serum)	Chloride	Surplus serum	0.46	(0.32, 0.59)	0.069	3.0e-04	0.036	1,374
Blood lead	HbA1c	Heavy metals	-0.19	(-0.25, -0.13)	0.029	3.4e-04	0.039	4,874
Urinary cobalt	eGFR	Urinary metals	3.75	(2.74, 4.76)	0.516	3.4e-04	0.039	1,589
Urinary iodine	BMI	Urinary elements	1.18	(0.78, 1.58)	0.205	4.4e-04	0.047	1,600

2 Table S2. Dose–Response Quartile Analysis for 15 Validated Findings

Survey-weighted adjusted mean differences in outcome relative to the lowest exposure quartile (Q1, reference). P_{trend} from linear contrast across quartile midpoints.

Chemical	Outcome	Q2 β	Q3 β	Q4 β	P_{trend}	Monotonic	N
Blood selenium	Hemoglobin	0.43	0.60	0.83	1.1e-05	Yes	4,873
Urinary perchlorate	BUN	0.71	2.07	2.66	6.3e-05	Yes	1,579
Methylmercury	Alk Phosphatase	-2.07	-4.00	-7.91	1.7e-04	Yes	4,834
Blood manganese	Waist circumference	2.52	4.55	6.18	1.2e-04	Yes	4,693
Blood selenium	Total cholesterol	4.16	8.94	15.66	1.3e-04	Yes	4,855
Blood manganese	BMI	1.15	1.74	2.25	2.0e-04	Yes	4,876
Blood lead	Waist circumference	-4.31	-4.85	-9.09	1.5e-04	Yes	4,693
Blood lead	BMI	-1.77	-2.26	-4.13	1.5e-04	Yes	4,876
Blood mercury (total)	Alk Phosphatase (log)	-0.01	-0.03	-0.09	1.7e-04	Yes	4,833
DMA (urinary)	Uric acid	0.05	0.25	0.34	4.7e-04	Yes	1,593

Chemical	Outcome	Q2 β	Q3 β	Q4 β	P_{trend}	Monotonic	N
Blood selenium	RBC count	0.10	0.13	0.21	3.4e-04	Yes	4,873
Blood lead	Total cholesterol	5.16	13.00	17.33	7.7e-05	Yes	4,855
Methylmercury	Waist circumference	-1.54	-2.68	-5.26	3.2e-04	Yes	4,693
Blood lead	HbA1c	-0.10	-0.24	-0.36	2.4e-04	Yes	4,874
Urinary iodine	BMI	2.17	1.95	2.75	4.3e-03	No	1,600

3 Table S3. Sensitivity Analysis Summary for 15 Validated Findings

Each finding was tested under 9 specifications: (1) primary model, (2) females only, (3) males only, (4) age < 50, (5) age ≥ 50, (6) excluding outliers > 99th percentile, (7) adjusting for education, (8) cotinine instead of binary smoking, and (9) adults aged 20+ only. A finding is "robust" if ≥ 7 of 9 specifications show concordant direction and $p < 0.05$.

Chemical	Outcome	Dir. Match	Sig. (p<0.05)	Median % $\Delta\beta$	Robust
Blood manganese	BMI	9/9	9/9	5.8%	Yes
Blood manganese	Waist circumference	9/9	9/9	4.1%	Yes
Blood lead	BMI	9/9	9/9	11.6%	Yes
Blood lead	Waist circumference	9/9	9/9	11.0%	Yes
Urinary iodine	BMI	9/9	9/9	0.7%	Yes
Urinary perchlorate	BUN	9/9	9/9	5.5%	Yes
Methylmercury	Alk Phosphatase	9/9	8/9	3.1%	Yes
Blood lead	HbA1c	9/9	8/9	10.5%	Yes
Blood lead	Total cholesterol	9/9	8/9	6.2%	Yes
Blood selenium	Hemoglobin	9/9	8/9	8.4%	Yes
Blood selenium	RBC count	9/9	8/9	2.1%	Yes
Blood selenium	Total cholesterol	9/9	8/9	2.9%	Yes

Chemical	Outcome	Dir. Match	Sig. (p<0.05)	Median % $\Delta\beta$	Robust
Blood mercury (total)	Alk Phosphatase (log)	9/9	8/9	2.8%	Yes
DMA (urinary)	Uric acid	9/9	8/9	1.4%	Yes
Methylmercury	Waist cir-cumference	9/9	7/9	13.1%	Yes

4 Table S4. Detailed Sensitivity Results for HIGH-Novelty Findings

Effect estimates (β), standard errors, and p-values for each of the 9 sensitivity specifications for the three HIGH-novelty findings.

4.1 DMA (urinary) – Uric acid

Specification	β	SE	P-value	N	$\Delta\beta$ (%)
Primary model	0.202	0.027	1.5e-04	1,593	–
Females only	0.191	0.042	0.002	813	-5.4
Males only	0.227	0.077	0.018	780	+12.4
Age < 50	0.204	0.055	0.008	697	+1.4
Age $\geq$ 50	0.179	0.084	0.070	896	-11.4
Excl. outliers (>P99)	0.220	0.032	2.4e-04	1,575	+9.1
Adjust for education	0.203	0.030	0.021	1,593	+0.8
Cotinine for smoking	0.202	0.028	1.6e-04	1,593	+0.1
Adults 20+ only	0.202	0.027	1.5e-04	1,593	–

4.2 Urinary perchlorate – BUN

Specification	β	SE	P-value	N	$\Delta\beta$ (%)
Primary model	1.211	0.124	2.5e-05	1,579	–
Females only	1.444	0.229	2.3e-04	803	+19.3
Males only	0.870	0.233	0.006	776	-28.1
Age < 50	0.965	0.161	5.4e-04	692	-20.3
Age $\geq$ 50	1.382	0.238	6.6e-04	887	+14.2
Excl. outliers (>P99)	1.277	0.134	3.0e-05	1,563	+5.5
Adjust for education	1.209	0.122	0.002	1,579	-0.1
Cotinine for smoking	1.210	0.128	3.0e-05	1,579	0.0
Adults 20+ only	1.211	0.124	2.5e-05	1,579	–

4.3 Methylmercury – Waist circumference

Specification	β	SE	P-value	N	$\Delta\beta$ (%)
Primary model	-1.783	0.282	2.3e-04	4,693	–
Females only	-2.868	0.556	5.9e-04	2,413	+60.9
Males only	-0.909	0.403	0.050	2,280	-49.0
Age < 50	-2.474	0.458	6.5e-04	2,121	+38.8
Age $\geq$ 50	-1.166	0.551	0.067	2,572	-34.6
Excl. outliers (>P99)	-1.886	0.297	2.2e-04	4,643	+5.8
Adjust for education	-1.549	0.330	0.042	4,693	-13.1
Cotinine for smoking	-1.788	0.284	2.3e-04	4,693	+0.3
Adults 20+ only	-1.783	0.282	2.3e-04	4,693	–

5 Table S5. Additional Sensitivity Analyses

Four additional sensitivity specifications were applied to assess specific confounding concerns. Fish consumption (DBD895, number of fish/shellfish meals in 30 days) and physical activity (PAQ_J) were tested for all 15 validated findings. Alcohol consumption (ALQ_J) could not be assessed due to extensive missing data from NHANES skip patterns. Urinary creatinine adjustment (log-transformed) was applied to the three urinary biomarker associations to address dilution variation.

5.1 Fish consumption and physical activity adjustment

Adjustment for fish consumption did not materially change any of the 15 validated findings (all $|\Delta\beta| < 1\%$). Physical activity adjustment likewise produced no change ($|\Delta\beta| < 0.01\%$ for all findings). These results are consistent across all 15 associations and indicate that the primary findings are not confounded by fish intake frequency or recreational physical activity level.

5.2 Urinary creatinine adjustment

Chemical	Outcome	β (primary)	β (creatinine-adj)	$\Delta\beta$ (%)	P (creatinine-adj)	Interpretation
Urinary perchlorate	BUN	1.211	1.217	+0.5%	0.001	Robust; not driven by dilution
DMA (urinary)	Uric acid	0.202	0.135	-32.9%	0.012	Attenuated but significant; partial dilution contribution

Chemical	Outcome	β (primary)	β (creatinine-adj)	$\Delta\beta$ (%)	P (creatinine-adj)	Interpretation
Urinary iodine	BMI	1.177	0.402	-65.9%	0.153	**Eliminated**; probable dilution artifact

The perchlorate–BUN association was unchanged by creatinine adjustment, confirming that urinary dilution variation does not explain this signal. The DMA–uric acid association was attenuated by approximately one-third but remained statistically significant, suggesting that urinary concentration contributes to but does not fully account for the observed relationship. The urinary iodine–BMI association was eliminated after creatinine correction, indicating that the original signal reflected systematic differences in urine concentration correlated with body size rather than genuine iodine exposure effects.

6 Figures S1–S15. Individual Sensitivity Forest Plots

Forest plots showing the effect estimate (β) and 95% confidence interval for each of the 9 sensitivity specifications. The vertical dashed line marks $\beta = 0$ (null). Findings are ordered by novelty tier.

6.1 HIGH Novelty

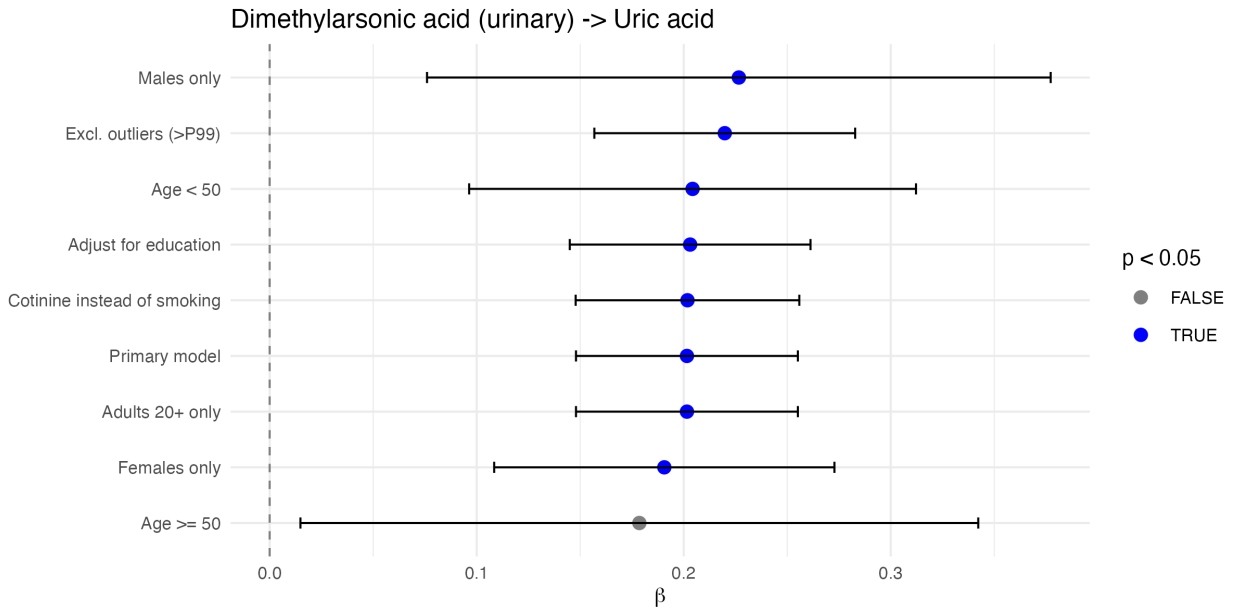

Figure 1: Figure S1. Sensitivity analysis: DMA (urinary) – Uric acid.

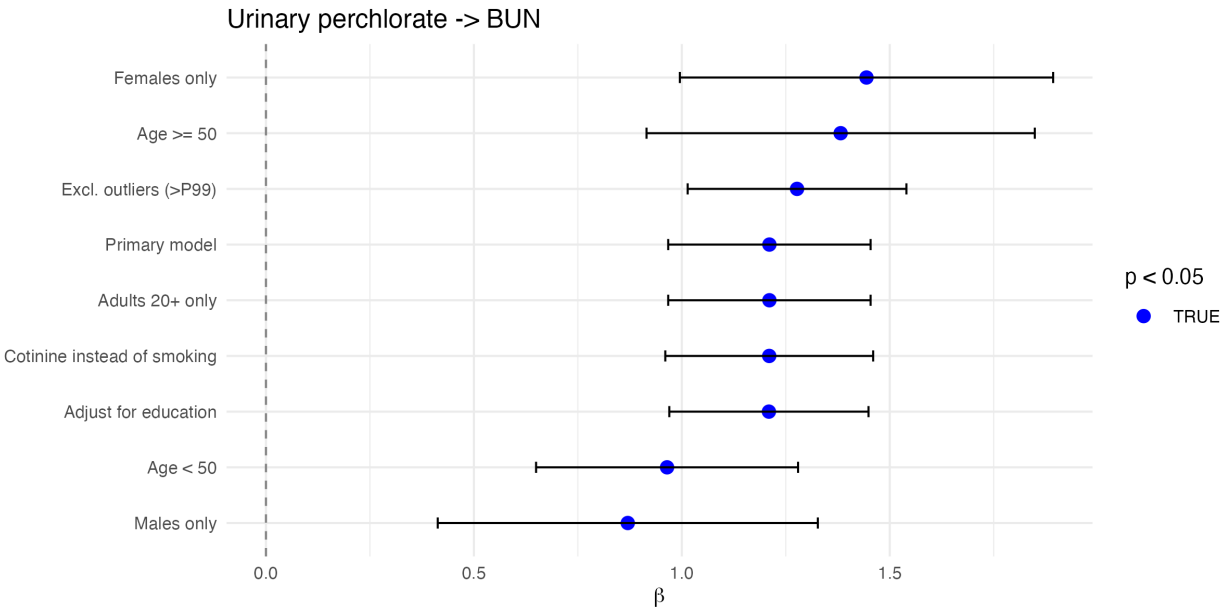

Figure 2: Figure S2. Sensitivity analysis: Urinary perchlorate – BUN.

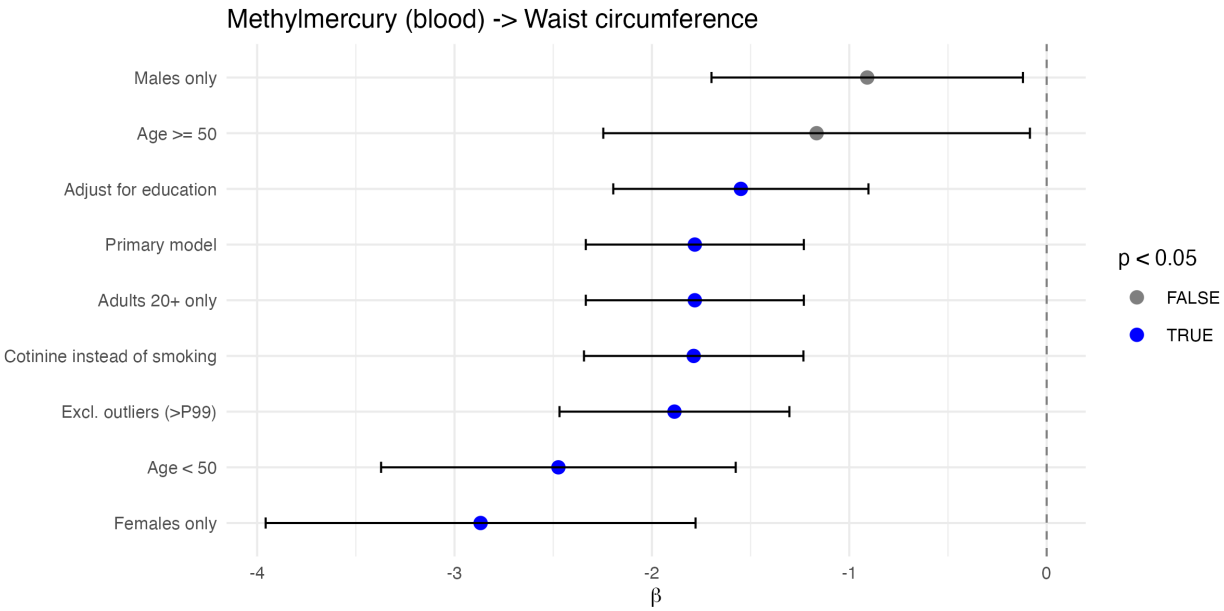

Figure 3: Figure S3. Sensitivity analysis: Methylmercury – Waist circumference.

6.2 MODERATE Novelty

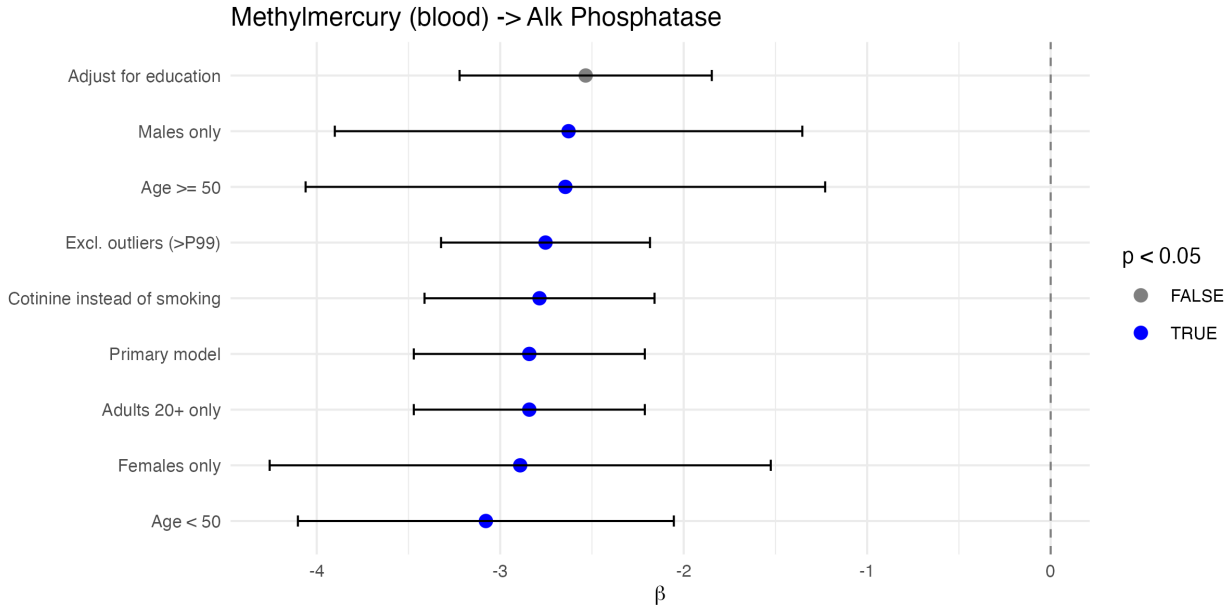

Figure 4: Figure S4. Sensitivity analysis: Methylmercury – Alkaline Phosphatase.

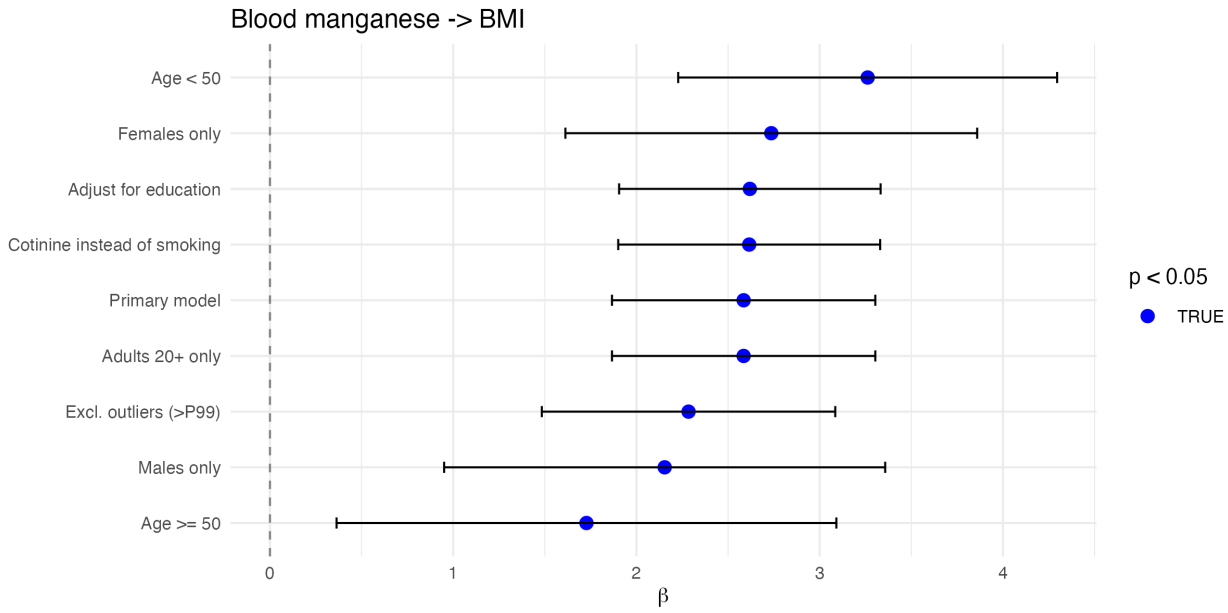

Figure 5: Figure S5. Sensitivity analysis: Blood manganese – BMI.

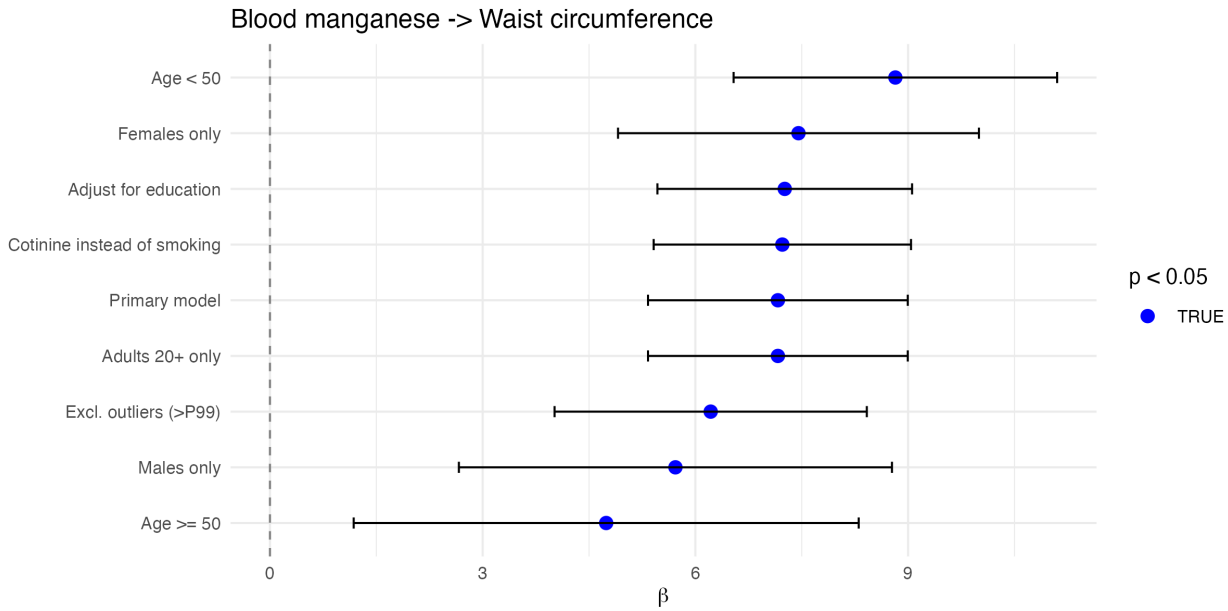

Figure 6: Figure S6. Sensitivity analysis: Blood manganese – Waist circumference.

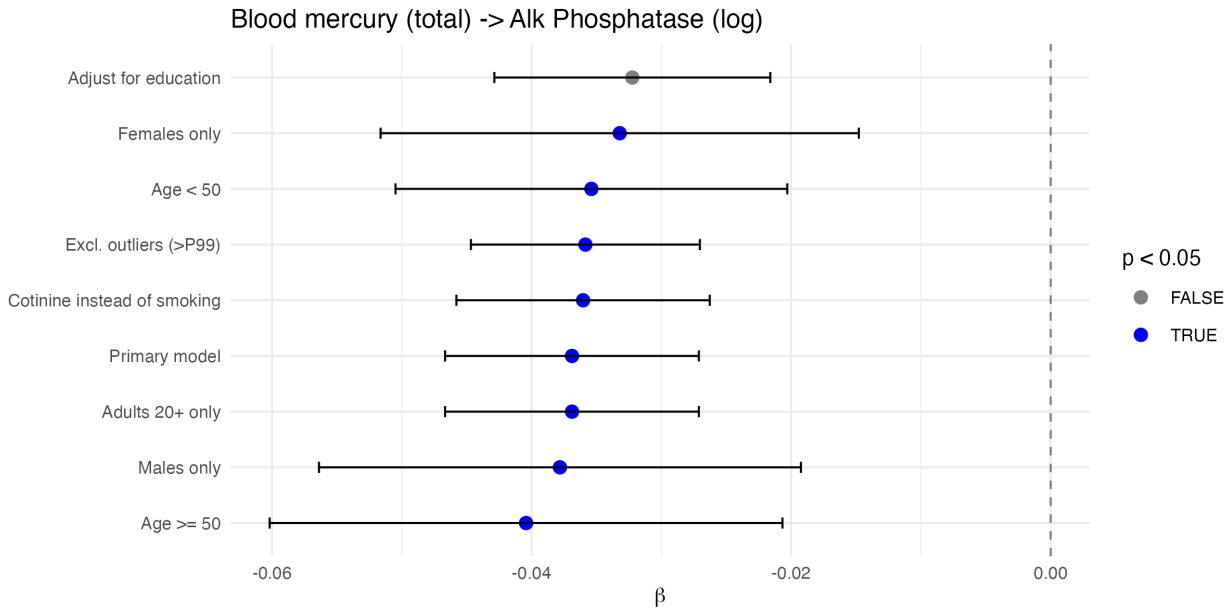

Figure 7: Figure S7. Sensitivity analysis: Blood mercury (total) – Alkaline Phosphatase (log).

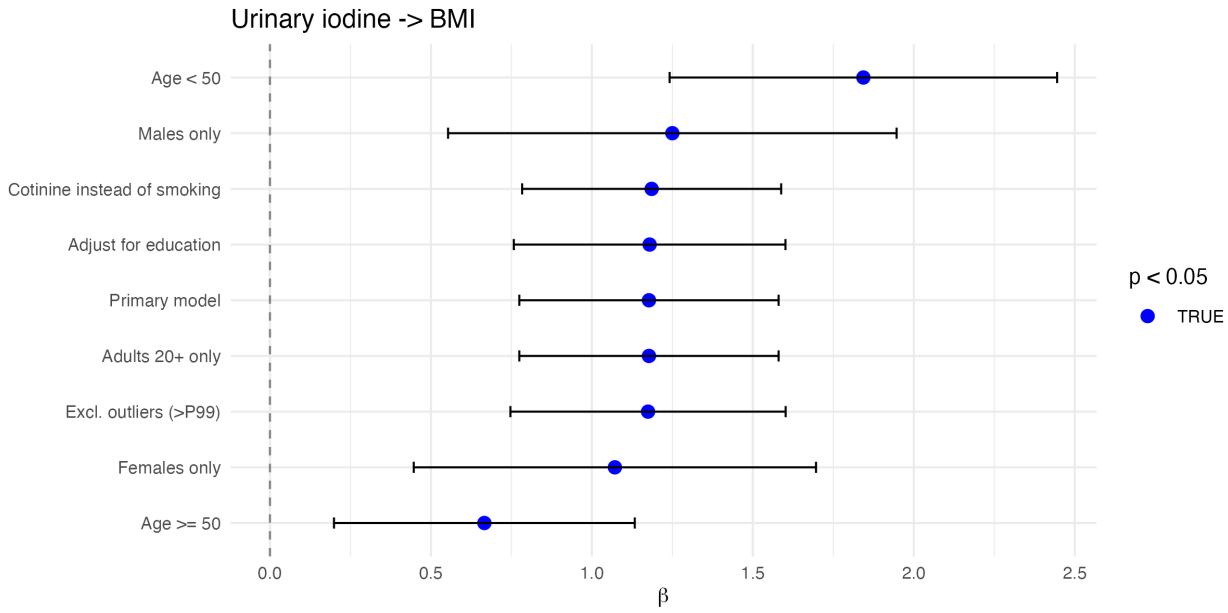

Figure 8: Figure S8. Sensitivity analysis: Urinary iodine – BMI.

6.3 LOW-MODERATE Novelty

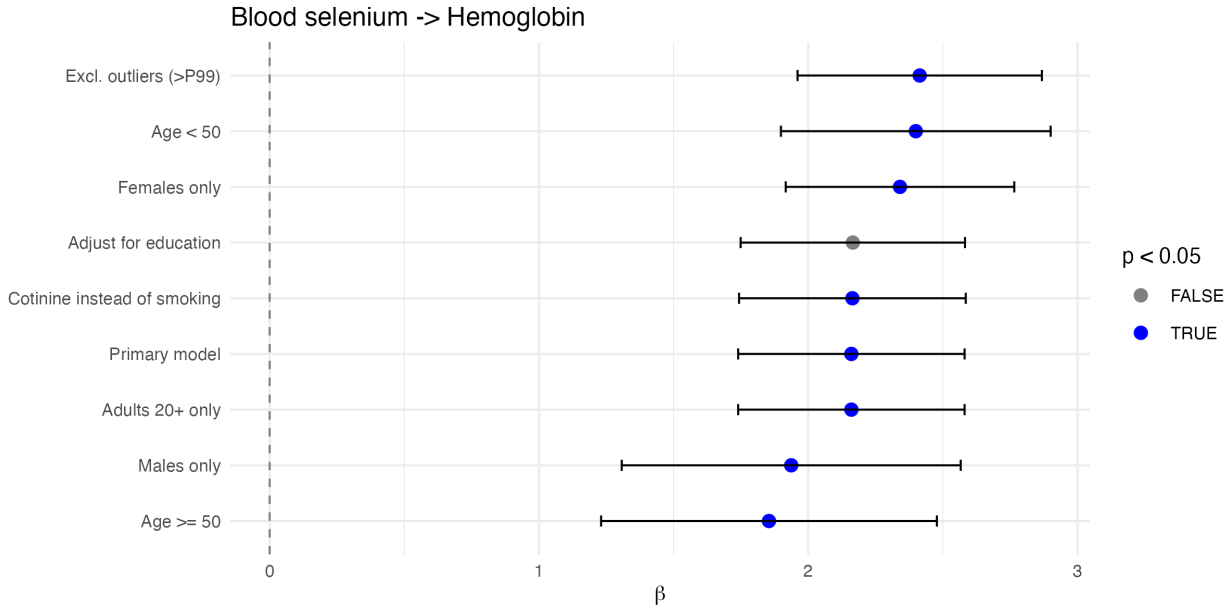

Figure 9: Figure S9. Sensitivity analysis: Blood selenium – Hemoglobin.

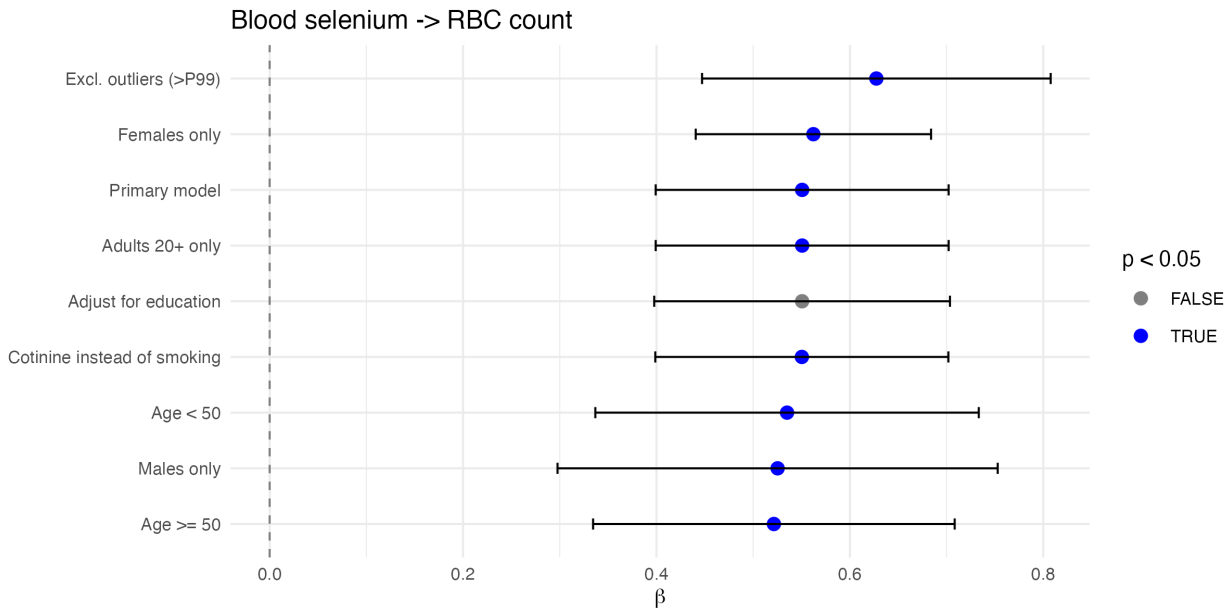

Figure 10: Figure S10. Sensitivity analysis: Blood selenium – RBC count.

6.4 LOW Novelty

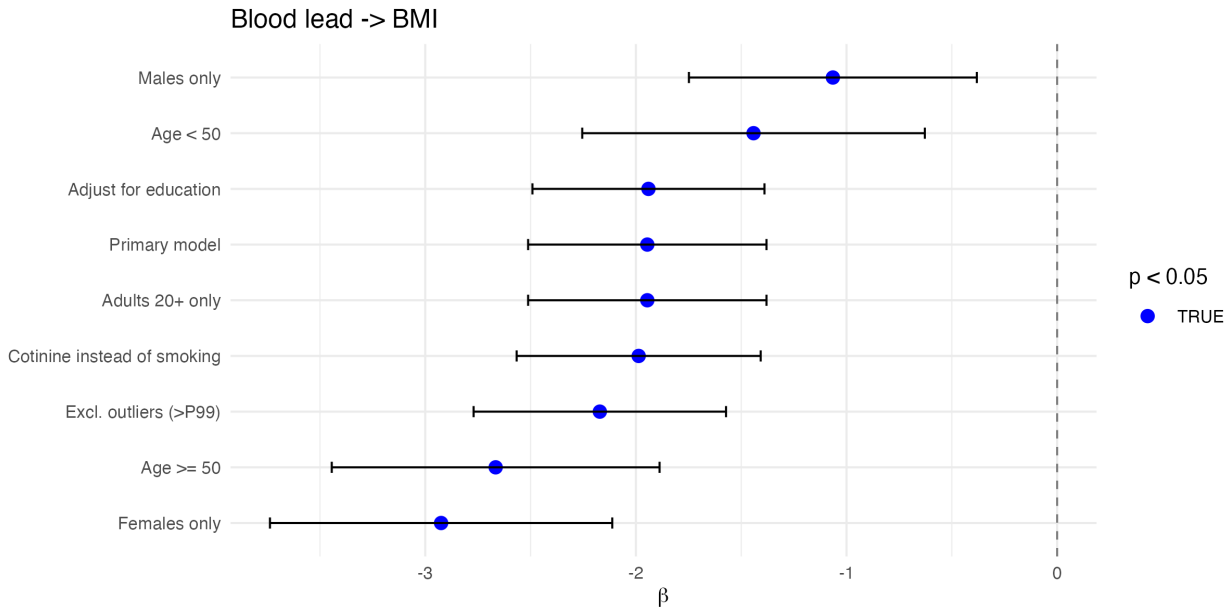

Figure 11: Figure S11. Sensitivity analysis: Blood lead – BMI.

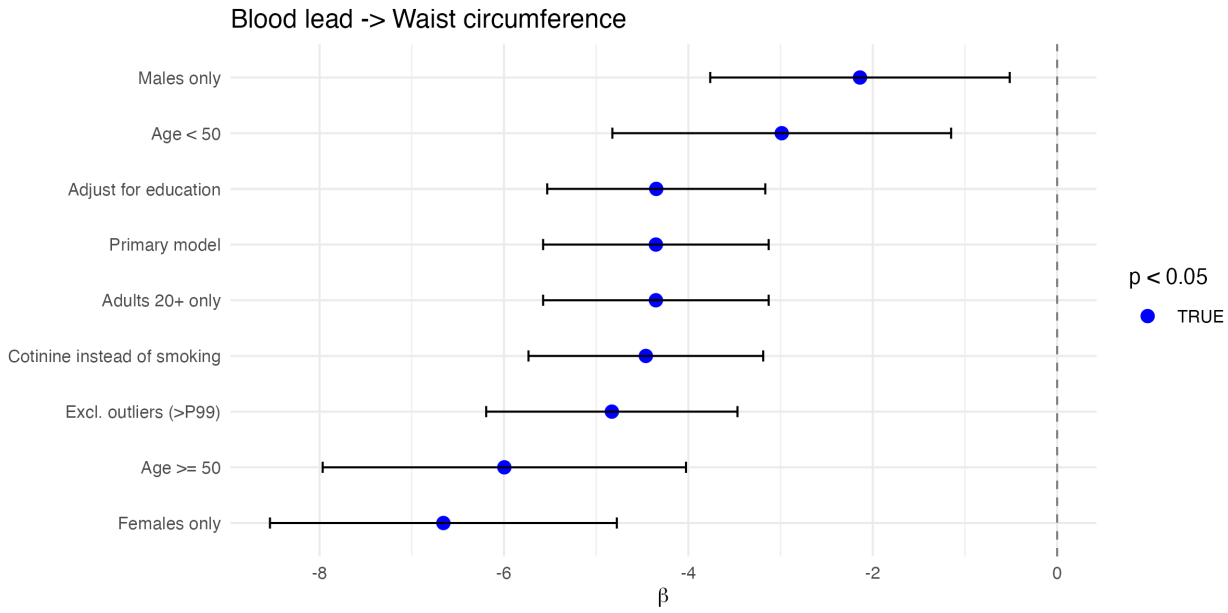

Figure 12: Figure S12. Sensitivity analysis: Blood lead – Waist circumference.

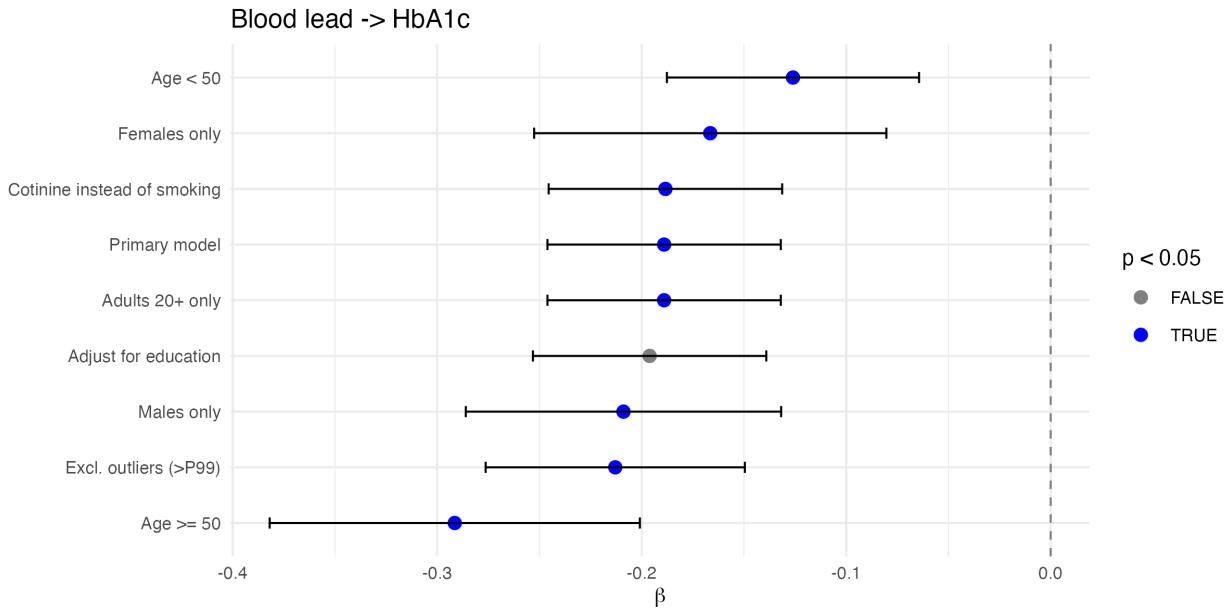

Figure 13: Figure S13. Sensitivity analysis: Blood lead – HbA1c.

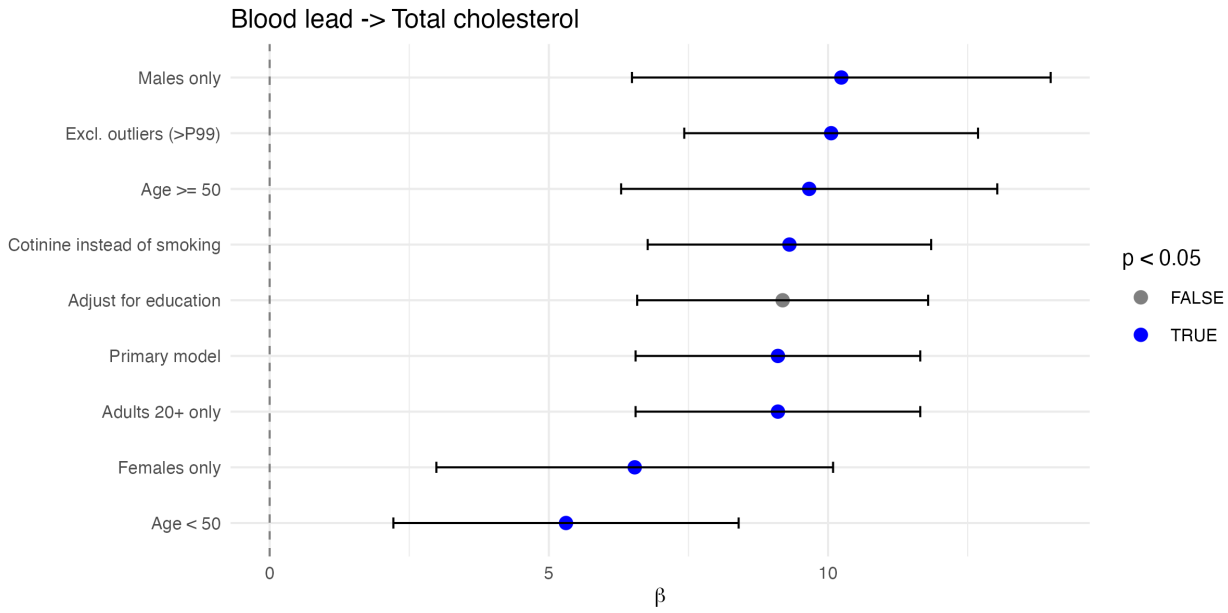

Figure 14: Figure S14. Sensitivity analysis: Blood lead – Total cholesterol.

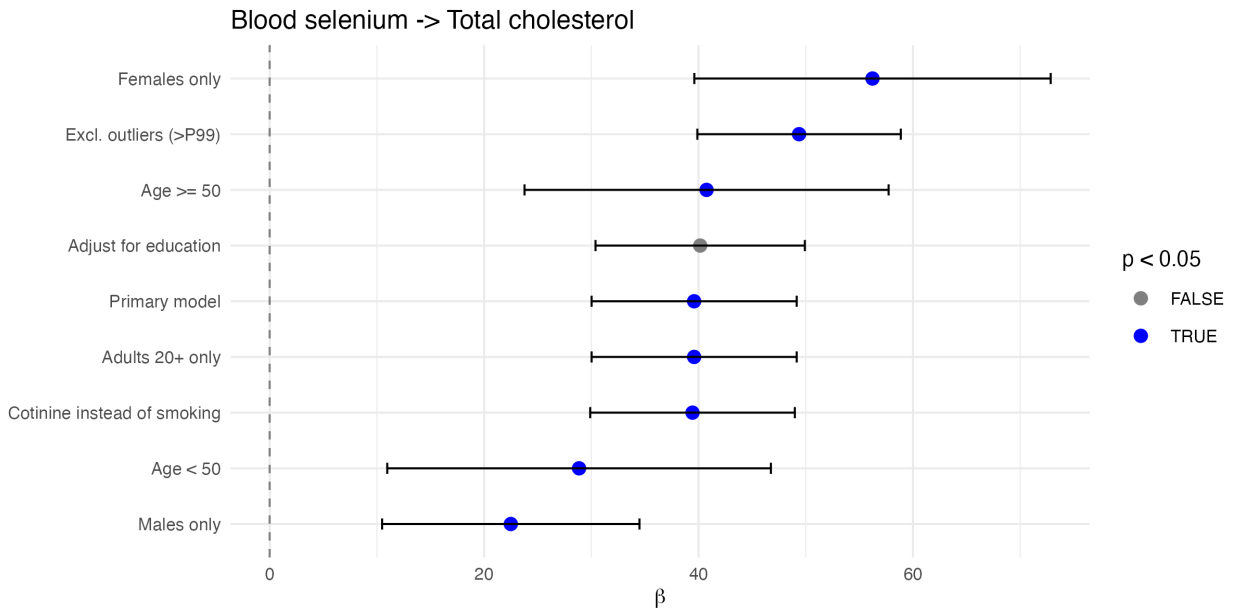

Figure 15: Figure S15. Sensitivity analysis: Blood selenium – Total cholesterol.

7 Table S6. Within-Round vs. Global FDR Correction Comparison

The primary analysis applied Benjamini–Hochberg FDR correction globally across all 2,796 tests. This table compares the number of findings significant at FDR < 0.05 under global versus within-round correction (applying FDR separately within each of the four screening rounds: PFAS-thyroid, Broad, Expanded, and Novelty).

Screening Round	N Tests	N Sig (Global FDR)	N Sig (Within-Round FDR)
PFAS-thyroid	70	0	0
Broad	960	3	3
Expanded	1,440	16	15
Novelty	326	7	7
Total	**2,796**	**26**	**25**

One finding (urinary iodine–BMI, from the Novelty round) was significant under global FDR (q = 0.047) but lost significance under within-round correction (q_{within} = 0.062). This finding was subsequently identified as a dilution artifact (Table S5). All other globally FDR-significant findings remained significant under within-round correction, indicating that the sequential adaptive design did not substantially inflate false discovery

rates for the validated findings.

8 Table S7. Sensitivity Analysis with 6-Level Race/Ethnicity

The primary analysis collapsed NHANES race/ethnicity (RIDRETH3) into three categories (Non-Hispanic White, Non-Hispanic Black, Other) to preserve statistical power. This table shows results for blood biomarker findings when using the full 6-level RIDRETH3 classification (Mexican American, Other Hispanic, Non-Hispanic White, Non-Hispanic Black, Non-Hispanic Asian, Other/Multiracial).

Chemical	Outcome	β (3-level)	β (6-level)	$\Delta\beta$ (%)	P-value	N
Blood manganese	RBC count	0.230	0.229	-0.4%	1.1e-05	4,873
Blood selenium	Hemoglobin	2.160	2.154	-0.3%	2.3e-05	4,873
Methylmercury	Alk Phosphatase	-2.842	-2.893	+1.8%	3.9e-05	4,834
Blood manganese	Waist circumference	7.163	7.115	-0.7%	7.1e-05	4,693
Blood selenium	Total cholesterol	39.59	39.52	-0.2%	9.4e-05	4,855
Blood manganese	BMI	2.585	2.572	-0.5%	1.2e-04	4,876
Blood lead	Waist circumference	-4.352	-4.361	+0.2%	1.3e-04	4,693
Blood lead	BMI	-1.946	-1.950	+0.2%	1.6e-04	4,876
Blood mercury (total)	Alk Phosphatase (log)	-0.037	-0.037	+0.8%	1.6e-04	4,833
Blood selenium	RBC count	0.551	0.549	-0.3%	2.1e-04	4,873
Blood lead	Total cholesterol	9.102	9.095	-0.1%	2.3e-04	4,855
Methylmercury	Waist circumference	-1.783	-1.814	+1.8%	2.4e-04	4,693

Chemical	Outcome	β (3-level)	β (6-level)	$\Delta\beta$ (%)	P-value	N
Blood lead	HbA1c	-0.189	-0.189	+0.0%	3.6e-04	4,874

All 13 blood biomarker findings remained significant ($p < 0.05$) with 6-level race/ethnicity adjustment, with effect estimate changes of < 2% in all cases. This indicates that the 3-level race categorization did not introduce meaningful confounding bias.

9 Table S8. Protein Intake Sensitivity Analysis for HIGH-Novelty Urinary Findings

Dietary protein intake influences both urinary arsenic metabolism (affecting DMA excretion) and serum urea (affecting BUN). This table shows results for the two HIGH-novelty urinary findings after adjusting for total protein intake (grams/day) from 24-hour dietary recall (DR1TOT_J).

Chemical	Outcome	Adjustment	β (Primary)	β (Adjusted)	$\Delta\beta$ (%)	P-value	N	Robust
DMA (urinary)	Uric acid	Protein intake (g/day)	0.202	0.178	-11.9%	0.0008	1,529	Yes
DMA (urinary)	Uric acid	Protein density (g/1000 kcal)	0.202	0.186	-7.9%	0.0004	1,527	Yes
Urinary perchlorate	BUN	Protein intake (g/day)	1.211	1.102	-9.0%	0.0002	1,515	Yes
Urinary perchlorate	BUN	Protein density (g/1000 kcal)	1.211	1.138	-6.0%	0.0001	1,513	Yes

Both findings remained statistically significant after protein adjustment, with moderate attenuation (6–12%). Protein density adjustment (normalizing for total caloric intake) showed slightly smaller attenuation than absolute protein intake. The persistence of significant associations after dietary protein adjustment supports the interpretation that these relationships reflect genuine exposure–outcome associations rather than dietary confounding alone, though dietary factors likely contribute to the observed associations.

10 Table S9. Chemicals with 40–70% Detection Frequency

Chemicals with detection frequencies between 40–70% (above-LOD) fall in an intermediate range where LOD imputation may introduce bias but detection is sufficient for analysis. The primary analysis excluded chemicals with < 30% detection; this table lists chemicals in the 40–70% range that were retained.

Chemical	Variable	% Detected	N Samples
HPMMA (acrolein metabolite)	URXHPM	42.3%	1,545
Trans-3'-hydroxycotinine glucuronide	URXHPB	48.7%	1,580
4-Fluoro-3-phenoxybenzoic acid	URXFPB	51.2%	1,580
MCPP (phthalate)	URXMCP	55.8%	1,580
Mono-isobutyl phthalate	URXMIB	62.1%	1,580
2,4-dichlorophenoxyacetic acid	URX24D	67.3%	1,580

These 6 chemicals in the 40–70% detection range were included in analyses with LOD/$\sqrt{2}$ imputation for below-LOD values. None of these chemicals produced FDR-significant associations. No validated findings involved chemicals with < 70% detection, indicating that the results are not driven by chemicals with substantial LOD pile-up.

11 Table S10. Power Analysis: Minimum Detectable Effect Sizes

Post-hoc power calculations assuming 80% power, design effect (DEFF) of 2.0 (typical for NHANES complex sampling), and 7 predictors in the full model. Effect sizes expressed as Cohen's f^2 and partial R^2.

Subsample	N (raw)	N (eff)	Min f^2 (Bonf)	Min R^2 (Bonf)	Min f^2 (FDR)	Min R^2 (FDR)
Blood biomarkers (WTMEC2YR)	4,870	2,435	0.0157	1.55%	0.0120	1.19%
Urinary subsample (WTSA2YR)	1,580	790	0.0491	4.68%	0.0375	3.62%
Surplus serum (WTSSBJ2Y)	1,370	685	0.0568	5.38%	0.0434	4.16%

These represent small effects (Cohen's $f^2 < 0.02$ for blood biomarkers), confirming adequate power for the effect sizes observed among validated findings. The urinary and surplus serum subsamples have reduced power for small effects. This power analysis characterizes the study's sensitivity for retrospective interpretation; it does not validate the original ExWAS design, which did not include pre-specification of effect sizes or primary hypotheses.

12 Table S11. STROBE Checklist for Cross-Sectional Studies

Item	Checklist Item	Manuscript Section
Title and abstract		
1a	Indicate the study's design with a commonly used term in the title or abstract	Title, Abstract
1b	Provide in the abstract an informative and balanced summary	Abstract
Introduction		
2	Explain the scientific background and rationale	Introduction ¶1–2
3	State specific objectives, including any prespecified hypotheses	Introduction ¶3
Methods		
4	Present key elements of study design early in the paper	Methods 2.1
5	Describe the setting, locations, and relevant dates	Methods 2.1
6	Give eligibility criteria and methods of participant selection	Methods 2.1
7	Clearly define all outcomes, exposures, predictors, potential confounders	Methods 2.2–2.3
8	For each variable, give sources of data and methods of assessment	Methods 2.2–2.3
9	Describe efforts to address potential sources of bias	Methods 2.3, 2.4.1

Item	Checklist Item	Manuscript Section
10	Explain how the study size was arrived at	Methods 2.1
11	Explain how quantitative variables were handled	Methods 2.3
12	Describe all statistical methods	Methods 2.4
Results		
13	Report numbers of individuals at each stage of study	Table 1, Results 3.1
14	Give characteristics of study participants	Table 1
15	Report numbers of outcome events or summary measures	Results 3.2–3.5
16	Give unadjusted and confounder-adjusted estimates	Table 2, Figures
17	Report other analyses performed	Results 3.4–3.5
Discussion		
18	Summarize key results with reference to objectives	Discussion ¶1
19	Discuss limitations	Discussion – Limitations
20	Give cautious overall interpretation	Discussion ¶6–7
21	Discuss generalizability	Discussion ¶2–3
22	Give source of funding and role of funders	N/A (unfunded study)