Exploring Heavy Metal Correlation with Gender, Race and Other Attributes

Vincent Preikstas; Dr. Isabelle Bichindartiz
Biomedical and Health Informatics Department

Introduction

Heavy metal toxicity is a well studied topic within medicine and academia

Research in Syracuse measured metabolites both targeted and untargeted, blood levels of mercury, cadmium and lead, demographic information and various health markers from children at Upstate Medical Center from 2013 to 2017. The set consisted of about 298 patients

Using binary logistic regression and linear regression, various attributes were assessed for correlation with heavy metals

The end goal of the project was to test correlation between gender, race and other attributes with presence of Lead, Cadmium or Mercury in blood

Background

Wealth of knowledge and research already exists looking at heavy metal correlation and linking disease to heavy metal exposure

Literature was reviewed to see if there was anything novel our analysis could tell us

Using correlation to identify at risk populations was considered from an article on the CDC website

High Blood pressure in post-menopausal women with low exposure to lead was considered in conjunction with an article outlining the effects of heart failure related to hypertension.

This very brief review identified at least one possibly higher risk population (postmenopausal women) and gave some precedent that going over similar markers for different populations could yield different correlations and outcomes with heavy metals.

In our project we want to see that for these children if there is any correlation with heavy metals on measures that might be overlooked

Project Design

Initial dataset included 298 rows of data with over 600 attributes per subject

Attributes were excluded down to 41 key attributes

Main variables analyzed included lead, mercury and cadmium levels, gender, race, and several others. The full data dictionary is available in the project final report

IBM's SPSS analysis tool was used to do binary logistic regression and linear regression. Both used a significance value of p = 0.05

A binary regression was done using gender as the dependent variable with two versions run, one with variables being considered independent of each other and one with variables being considered together.

Race was then used as the dependent variable in the same analysis that was done using gender

Ibuprophen, asperin and vitamins were also used as the dependent variable in a binary logistic regression, but only measured against the three heavy metals

Cadmium, lead and mercury each were used as dependent variables in a linear regression analysis

Results

Please see the output pictures from SPSS to the right for a sample of the analysis results. All tables of analysis results are available in the final report document. Some were omitted here to save space

Gender was not correlated with any of the heavy metals

Race was correlated with lead when considering attributes separately and was correlated with cadmium when considering attributes together

Aspirin, ibuprofen and vitamins were not found to have any significant correlations with heavy metal levels in blood

The model for socioeconomic score was found to be predictive of lead levels although the model did not have high enough significance to be considered although it was close (p = 0.089)

Binary Logistic Regression, Gender as Dependent Variable, Attributes Not Together

Variables not in the Equation[a]

			Score	df	Sig.
Step 0	Variables	yr	1.963	1	.161
		race	.516	1	.472
		childage	2.462	1	.117
		sesscore	.080	1	.777
		bmipct	3.184	1	.074
		TCmgdL	.863	1	.353
		HDLmgdL	5.059	1	.024
		nonHDLmgdL	.028	1	.866
		GlumgdL	.823	1	.364
		Cd	.014	1	.905
		pb	1.987	1	.159
		hg	.007	1	.935
		SystolicBP	.150	1	.699
		DiastolicBP	1.501	1	.220
		HeartRate	4.411	1	.036
		LVCardout	.028	1	.868
		LVCardindex	1.601	1	.206
		LVMass	.127	1	.722
		LVMassHt	.014	1	.904
		LVMassindex	3.598	1	.058
		Height	2.439	1	.118
		Weight	3.393	1	.065
		BSA	3.624	1	.057
		aspirin	2.278	1	.131
		Ibuprofen	8.115	1	.004
		vitamins	.119	1	.730
		Latexpaint	1.241	1	.265
		oilpaint	4.401	1	.036
		leadpaint	.026	1	.871
		Alanine	11.851	1	.001
		Glycine	.177	1	.674
		Valine	3.712	1	.054
		Leucine	1.149	1	.284
		Threonine	4.782	1	.029
		Serine	2.001	1	.157
		Proline	2.782	1	.095
		Methionine	1.030	1	.310
		Phenylalanine	.000	1	.991
		Glutamine	2.466	1	.116

Questions

Please feel free to email at vpreikst@oswego.edu with any further questions or to request a copy of the final report.

Binary Logistic Regression, Gender as Dependent Variable, Attributes Considered Together

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	Upper
Step 1[a]	yr	-.015	.300	.003	1	.960	.985	.547	1.772
	race	.219	.652	.113	1	.737	1.245	.347	4.469
	childage	-1.355	.391	12.003	1	.001	.258	.120	.555
	sesscore	-.265	.302	.771	1	.380	.767	.424	1.387
	bmipct	.017	.015	1.287	1	.257	1.018	.987	1.048
	TCmgdL	.390	.481	.658	1	.417	1.478	.575	3.796
	HDLmgdL	-.433	.482	.809	1	.368	.648	.252	1.667
	nonHDLmgdL	-.406	.482	.711	1	.399	.666	.259	1.713
	GlumgdL	-.021	.028	.562	1	.453	.980	.928	1.034
	Cd	6.575	5.722	1.320	1	.251	716.808	.010	53257084.90
	pb	-.696	.461	2.278	1	.131	.499	.202	1.231
	hg	-.124	.198	.392	1	.531	.884	.600	1.301
	SystolicBP	.015	.030	.270	1	.603	1.015	.958	1.076
	DiastolicBP	-.041	.041	1.037	1	.308	.960	.886	1.039
	HeartRate	.089	.033	7.260	1	.007	1.093	1.025	1.166
	LVCardout	-2.619	2.020	1.681	1	.195	.073	.001	3.818
	LVCardindex	1.673	2.650	.399	1	.528	5.330	.030	960.449
	LVMass	-.012	.187	.004	1	.949	.988	.685	1.426
	LVMassHt	.116	.363	.102	1	.749	1.123	.551	2.289
	LVMassindex	-.091	.198	.209	1	.648	.913	.619	1.347
	Height	.259	.256	1.025	1	.311	1.296	.784	2.142
	Weight	.155	.304	.259	1	.611	1.167	.643	2.117
	BSA	-9.545	27.679	.119	1	.730	.000	.000	2.601E+19
	aspirin	2.844	1.546	3.385	1	.066	17.184	.830	355.578
	Ibuprofen	-1.366	.531	6.616	1	.010	.255	.090	.722
	vitamins	-.178	.504	.124	1	.724	.837	.312	2.248
	Latexpaint	.682	.598	1.297	1	.255	1.977	.612	6.389
	oilpaint	1.116	.593	3.545	1	.060	3.053	.955	9.757
	leadpaint	.280	.258	1.175	1	.278	1.323	.797	2.196
	Alanine	.006	.003	2.999	1	.083	1.006	.999	1.012
	Glycine	.003	.004	.425	1	.515	1.003	.994	1.012
	Valine	-.005	.011	.224	1	.636	.995	.974	1.016
	Leucine	.013	.032	.175	1	.676	1.013	.952	1.078
	Threonine	.003	.008	.168	1	.682	1.003	.987	1.020
	Serine	-.003	.006	.261	1	.609	.997	.984	1.009
	Proline	-.002	.004	.195	1	.658	.998	.991	1.006
	Methionine	-.013	.068	.036	1	.849	.987	.864	1.128
	Phenylalanine	-.035	.042	.690	1	.406	.966	.889	1.049
	Glutamine	.001	.001	1.003	1	.317	1.001	.999	1.003
	Constant	12.207	604.868	.000	1	.984	200188.505		

Linear Regression with Lead as the Dependent Variable

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	17.723	39	.454	1.384	.089[b]
	Residual	44.978	137	.328		
	Total	62.701	176			

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	163.123	133.979		1.218	.225
	yr	-.080	.067	-.137	-1.209	.229
	gender	-.142	.110	-.119	-1.291	.199
	race	-.060	.137	-.050	-.437	.663
	childage	-.123	.076	-.185	-1.626	.106
	sesscore	-.173	.060	-.252	-2.873	.005
	bmipct	.000	.003	-.018	-.109	.913
	TCmgdL	.145	.101	6.677	1.439	.153
	HDLmgdL	-.155	.101	-3.481	-1.536	.127
	nonHDLmgdL	-.144	.101	-6.141	-1.429	.155
	GlumgdL	-.002	.002	-.100	-1.213	.227
	Cd	1.915	.975	.162	1.963	.052
	hg	.025	.050	.039	.499	.618
	SystolicBP	.008	.006	.113	1.208	.229
	DiastolicBP	.001	.009	.006	.064	.949
	HeartRate	-.008	.006	-.176	-1.272	.206
	LVCardout	-.245	.387	-.480	-.634	.527
	LVCardindex	.392	.501	.522	.781	.436
	LVMass	.010	.010	.391	.975	.331
	LVMassHt	-.006	.008	-.172	-.720	.472
	LVMassindex	-.009	.007	-.264	-1.334	.184
	Height	.040	.042	.595	.955	.341
	Weight	.073	.063	1.734	1.168	.245
	BSA	-5.872	5.868	-2.134	-1.001	.319
	aspirin	.154	.264	.047	.584	.560
	Ibuprofen	-.020	.110	-.016	-.179	.858
	vitamins	-.070	.106	-.057	-.662	.509

Future Plans

Further analysis of targeted and untargeted metabolites could bear fruit

Looking at certain groups of attributes together could be interesting. Considering metabolites, heavy metals, and health markers together to see where there is any limiting going on would be of special interest. Finding evidence of various metabolites regulating or influencing toxicity could prove to be novel information

References

At-Risk Populations. (2019, July 30). Retrieved from https://www.cdc.gov/nceh/lead/prevention/populations.htm

Singh, S., & Sharma, S. (2019, November 14). High-Output Cardiac Failure. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK513337/

Lead Levels Linked to Blood Pressure. (2003, March 25). Retrieved from https://www.webmd.com/hypertension-high-blood-pressure/news/20030325/lead-levels-linked-to-blood-pressure#