Johns Hopkins University, Dept. of Biostatistics Working Papers

1-8-2008

ON THE MERITS OF VOXEL-BASED MORPHOMETRIC PATH-ANALYSIS FOR INVESTIGATING VOLUMETRIC MEDIATION OF A TOXICANT'S INFLUENCE ON COGNITIVE FUNCTION

Shu-chih Su
Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health

Brian S. Caffo
Department of Biostatistics, The Johns Hopkins Bloomberg School of Public Health, bcaffo@jhsph.edu

Lynn E. Eberly
University of Minnesota, School of Public Health, Division of Biostatistics

Elizabeth Garrett-Mayer
Medical University of South Carolina, Department of Biostatistics, Bioinformatics and Epidemiology, Hollings Cancer Center

Walter F. Stewart
Johns Hopkins Bloomberg School of Public Health, Department of Epidemiology and Geisinger Clinic, Geisinger Center for Health Research

See next page for additional authors

Authors
Shu-chih Su, Brian S. Caffo, Lynn E. Eberly, Elizabeth Garrett-Mayer, Walter F. Stewart, Sining Chen, David Yousem, Christos Davatzikos, and Brian Schwartz

On the merits of voxel-based morphometric path-analysis for investigating volumetric mediation of a toxicant's influence on cognitive function

Shu-chih Su[1], Brian S. Caffo[1], Lynn E. Eberly[6], Elizabeth Garrett-Mayer[7], Walter F Stewart[3,8], Sining Chen [2,] David Yousem[5], Christos, Davatzikos[9] and Brian Schwartz[2,3,4]

Affiliations

Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

[1] Department of Biostatistics

[2] Department of Environmental Health Sciences

[3] Department of Epidemiology

Johns Hopkins School of Medicine, Baltimore, MD

[4] Department of Medicine

[5] The Russell H Morgan Department of Radiology and Radiological Sciences

University of Minnesota, School of Public Health, Minneapolis, MN

[6] Division of Biostatistics

Medical University of South Carolina, Charleston, SC

[7] Department of Biostatistics, Bioinformatics, and Epidemiology, Hollings Cancer Center

Geisinger Clinic, Danville, PA

[8] Geisinger Center for Health Research

University of Pennsylvania School of Medicine, Philadelphia, PA

[9] Department of Radiology

Correspondence

Address to Brian Caffo, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe St., Baltimore, MD 21205, Voice (410) 955-3504, Fax (410) 955-0958, email bcaffo@jhsph.edu

Abstract

We previously showed that lifetime cumulative lead dose, measured as lead concentration in the tibia bone by X-ray fluorescence, was associated with persistent and progressive declines in cognitive function and with decreases in MRI-based brain volumes in former lead workers. Moreover, larger region-specific brain volumes were associated with better cognitive function. These findings motivated us to explore a novel application of path analysis to evaluate effect mediation, specifically, whether the association of lead dose with cognitive function is mediated through brain volumes, on a voxel-wise basis. Voxel-wise path analysis, at face value, represents the natural evolution of voxel-based morphometry methods to answer questions of mediation. However, application of these methods to the former lead worker data demonstrated potential limitations in this approach. In particular, there was a tendency for results to be strongly biased towards the null hypothesis (lack of mediation). Moreover, a complimentary analysis using anatomically-derived regions of interest (ROI) volumes yielded opposing results, suggesting evidence of mediation. Specifically, in the ROI-based approach, there was evidence that the association of tibia lead with function in three cognitive domains (e.g., visuo-construction, executive functioning, eye-hand coordination) was mediated through the volumes of total brain, frontal gray matter, and/or possibly cingulate. A simulation study was conducted to investigate whether the voxel-wise results arose from an absence of localized mediation, or more subtle defects in the methodology. The simulation results showed the same null bias evidenced as seen in the lead workers data. Both the lead worker data results and the simulation study suggest that a null-bias in voxel-wise path analysis limits its inferential utility for producing confirmatory results.

Key words and phrases: brain volumes, cognitive function, environmental toxicology, path analysis, structural equation models, tibia lead

INTRODUCTION

In a longitudinal study of former organolead workers, we observed significant associations among lead dose, cognitive function and brain structure (Schwartz, Stewart et al. 2000; Links, Schwartz et al. 2001; Stewart, Schwartz et al. 2006). Specific findings include: (1) peak tibia lead (PTL), a measure of lifetime cumulative lead dose at the termination of employment, was found to be associated with smaller regional brain volumes (Stewart, Schwartz et al. 2006), (2) higher levels of PTL were associated with worse cognitive function at cross-section (Stewart, Schwartz et al. 1999) and decline in cognitive function over time (Schwartz, Stewart et al. 2000); and (3) larger brain volumes were associated with better cognitive function at cross-section (Schwartz, Chen et al. 2007). These results suggest the possibility that lead may exert its adverse effects on cognitive function through neurobiological and structural changes that are at least partially reflected in decreases in magnetic resonance imaging (MRI) measures of brain volume. That is, our hypothesis is that regional brain volumes mediate the relationship between cumulative lead dose and cognitive function.

Path analysis models, a subset of structural equation models (see Bollen 1989), can provide information on mediation by decomposing a total effect into direct and indirect components (Wright 1921). This approach is widely used, including using by neuroscientists to quantify functional relationships between multiple brain regions (see Buchel and Friston 1997; Buchel and Friston 1997). An early application of structural equation models to neuroimaging can be found in the work by McIntosh and Gonzalez-Lima (1991; 1994). In addition, structural equation models were used to evaluate potential mediation from data-derived volumetric regions of interest (ROIs) in the study reported on herein (Caffo, Chen et al. 2007). None of these studies have evaluated path analysis models at the voxel level.

The main purpose of this manuscript was to evaluate the utility of path analysis models for investigating questions of mediation, as applied to both anatomically-derived ROIs and to voxel-wise data from stereotactically normalized and smoothed structural MRI images. The latter builds on so-called voxel-based-morphometry (VBM), a whole-brain technique for characterizing localized structural differences in magnetic resonance images (Wright, McGuire et al. 1995; Ashburner and Friston 2000; Ashburner and Friston 2001; Mechelli, Price et al. 2005). To our knowledge, this manuscript is the first attempt to investigate voxel-wise path analysis models for mediation. This manuscript particularly focuses on the weaknesses of this voxel-wise approach, and uses the ROI analysis as motivation. Such weaknesses are unfortunate, since all other uses of path analysis in this area have applied the models only on univariate regional summaries. Such approaches fail in detecting localized effects existing across regional boundaries.

For our approach, we assume that a localized form of mediation would require a localized effect of PTL on volume, a localized effect of volume on function (as established in Schwartz, Chen et al. 2007) and a relevant overlap of these regions. Voxel-wise path analysis, used in conjunction with smoothing of the volumetric images, attempts to find this region of overlap, independently of regional anatomical boundaries.

We first describe the observational source data, followed by a description of the methodological approach. Subsequently, the ROI-based and voxel-wise path analysis results are presented. A simulation study is conducted to validate the concerns raised by our proposed approach. The limitations of this approach are discussed at the end.

MATERIALS AND METHODS

Motivating data

Data for the current analysis are from an ongoing prospective study of lead's impact on the

central nervous system structure and function. The study included male former organolead manufacturing workers (Stewart, Schwartz et al. 1999; Schwartz, Stewart et al. 2000; Stewart, Schwartz et al. 2006; Schwartz, Chen et al. 2007). The relevant subset of former lead manufacturing workers possessing cognitive testing, an acceptable MRI and key covariate data totaled 512 out of 628 subjects in the original data. The mean age of the subjects was 60.39 years with a standard deviation of 7.93 years (range 34.7 to 78.3 years). The average time since last occupational exposure to lead was 8.6 years with a standard deviation of 9.8 years (range 0 to 52 years). Subjects were recruited at one of several tours with tibia lead being measured at the time of the first visit. Structural MRIs were collected in a second phase of the study while cognitive function measurements were collected serially at each tour from the first visit. In this manuscript the cognitive function measurement obtained nearest in time to MRI acquisition. Further background information on the study design and demographic information about the sample can be found in the work by Stewart, Schwartz et al. (1999), Schwartz, Stewart et al. (2000), and Stewart, Schwartz et al. (2006) and Schwartz, Chen et al. (2007).

Data Acquisition

Lead dose was measured on study participants by X-ray fluorescence of the tibia. An extrapolation to estimate cumulative dose at the termination of employment, PTL, was obtained via kinetic models. This quantity has been shown to be a principal measure of the deleterious effects of lead (Links, Schwartz et al. 2001) and is used herein.

Patients were scanned on a General Electric 1.5 Tesla Signa scanner. Quantitative analyses were performed on the T1-weighted volume acquisitions with voxel sizes of $2.34 \times 0.98 \times 0.98$ mm^3 (Magnetom) and $1.20 \times 0.93 \times 0.93$ mm^3 (Signa)**.** The field of view was 24 cm and the matrix size was 256 × 256. After the acquisition of images, the images were preprocessed, including spatial normalization to common stereotactic coordinates, and then segmented into gray matter, white

matter, and cerebrospinal fluid voxels. A 10 mm full width at half maximum Gaussian smoother was applied using the SPM package for Matlab version 6.5 (Mathworks).

The cognitive function measures include a wide range of standard tests. The battery of neuropsychological tests was collapsed into six cognitive domain scores. In this study we only performed voxel-wise path analysis on the three of six domains that were previously found to be associated with lead dose among the 514 persons in this analysis (Caffo, Chen et al. 2007; Schwartz, Chen et al. 2007), namely visuo-construction, executive functioning, and eye-hand coordination. The tests that comprised the visuo-construction domain were the Rey complex figure, copy task, and block design from the Wechsler Adult Intelligence Scale. The tests that comprised the executive functioning domain were Purdue pegboard (assembly task minus both hands task), the Stroop test (C form time minus A form time), and the Trails-making test (B form time minus A form time). The tests that comprised the eye-hand coordination domain were Purdue Pegboard dominant hand, non-dominant hand, and both hands, and the Trails-making test A exam (Shih, Glass et al. 2006).

Data Analysis

In this manuscript, we investigated volumetric mediation of peak tibial lead's effect on cognitive function. Our goals were both to consider the scientific question of mediation as well as the methodological question of the efficacy of evaluating mediation on a voxel-wise scale. To this end, mediation was first considered using anatomically-based ROIs and later explored using voxel level data. A total of 20 ROIs were selected and defined for study using computerized template matching techniques (Stewart, Schwartz et al. 2006). For the voxel-wise approach, the regional analysis of volumes in normalized space (RAVENS) algorithm was used to map volumes into a common stereotactic space (Davatzikos 1996). Specifically, each transformed image is mapped onto a template brain, separately for gray and white matter, retaining volumetric information in the

process. That is, the intensity values for voxels in a RAVENS map represents a volume from the original image space. The original T1 intensities are not represented in the RAVENS images.

Path analysis models are used to analyze systems of hypothesized structural equations (Wright 1921; Wright 1934; Wright 1960). These models arise from a path diagram that describes hypothesized structural relationships between the study variables. We assume that these structural relationships are linear with Gaussian errors. Under these assumptions the sample covariance is sufficient for estimating path analysis parameters. As such, path analysis models provide information on mediation by decomposing the covariance matrix to conditional effects represented by the path diagrams.

The idea of mediation can be traced back to the early work of Wright (1921) in genetics and the work of Woodworth (1926) in psychology. Mediation implies a causal hypothesis where an independent (X) variable affects a mediator (Z) which in turn affects a dependent variable (Y) (Holland 1988; Sobel 1990). The evaluation of a mediating effect has been widely conducted in medical research for treatment and prevention (Judd and Kenny 1981; Donaldson 2001; Kraemer, Wilson et al. 2002).

To illustrate, consider a simple path diagram with three variables X, Y and Z *(*shown in Figure 1). Here, X denotes the independent variable, Z the mediator and Y the dependent variable. The parameter γ_1 is the path coefficient from X to Y; γ_2 is the path coefficient from X to Z. The structural equations can be written as follows:

$$Z_i = \alpha_1 + \gamma_2 X_i + \varepsilon_{2i}, \quad (1)$$

$$Y_i = \alpha_2 + \gamma_1 X_i + \gamma_3 Z_i + \varepsilon_{3i}, \quad (2)$$

where i indicates subject. Typically, one assumes that $\varepsilon_{2i} \sim N(0, \sigma_2^2)$ and $\varepsilon_{3i} \sim N(0, \sigma_3^2)$ are independent of each other.

The "total effect" of X on Y is then equal to $\gamma_1 + \gamma_2\gamma_3$. Here we say that γ_1 is the direct effect

of X on Y, as it represents a direct path from X to Y via equation (2). Then, $\gamma_2\gamma_3$ represents the indirect, or mediating, effect through Z. That is, a one unit change in X causes a γ_2 unit change in Z [from equation (1)], which, in turn, causes a γ_3 change in Y [from equation (2)]. Hence the term "indirect" (Bollen 1987) is used, by virtue of the effect having to pass through Z to get to Y. We use maximum likelihood (Joreskog 1967) for model estimation.

The main purpose of this paper is to examine to what extent brain volume mediates the association between peak tibia lead and measures of cognitive function. For this, we have developed a two-stage path model relating PTL (*X*) to volume (*Z*), and PTL and volume to cognitive function (*Y*), at each stage including relevant covariates (see Figure 2). In addition to the three key variables (PTL, volumes, domain scores), the following potential confounding variables: age (years), an indicator of the visit number for the cognitive function testing are fitted to account for a learning effect, and height (in inches), which was a surrogate for intrinsic brain size (e.g., intracranial volume, which was not measured).

The key structural equations for the path analysis model (Figure 2) are as follows:

$$\text{Domain} = \alpha_0 + \gamma_1 * \text{Lead} + \gamma_3 * \text{Brain Structure} + \gamma_4 * \text{Age} + \gamma_9 * \text{Visit Number} + \varepsilon_0,$$

$$\text{Volume} = \alpha_1 + \gamma_2 * \text{Lead} + \gamma_5 * \text{Age} + \gamma_7 * \text{Height} + \varepsilon_1.$$

The path model parameters were estimated using maximum likelihood assuming normality of the errors using the sem package in the R computing environment (http://www.r-project.org/).

We use a formal test statistic to assess the mediation by taking the indirect effect estimates, $\hat{\gamma}_2\hat{\gamma}_3$, and dividing by its standard error, and comparing the resulting statistic to standard normal quantiles. We use the multivariate delta-method (Bishop, Fienberg et al. 1975) to derive the variance estimate, yielding:

$$Var(\hat{\gamma}_2^2\hat{\gamma}_3^2) \approx \hat{\gamma}_2^2\hat{\sigma}_{\gamma_3}^2 + \hat{\gamma}_3^2\hat{\sigma}_{\gamma_2}^2 + 2\hat{\gamma}_2\hat{\gamma}_3\,\hat{Cov}(\hat{\gamma}_2,\hat{\gamma}_3),$$

where the estimated variance and covariance terms are given by the inverse observed information matrix.

RESULTS

Path Analysis Based on Regional Summaries

We examined extent that the association of PTL to the three cognitive domain cores is mediated through the lower volume in the 20 ROIs. The mediation proportion (i.e. the estimated proportion of the total effect of lead that is mediated through volume) was listed in Table 1 for the three cognitive domain scores and 20 ROI volumes. The estimated proportion of the total effect of PTL on the visuo-construction domain score that was attributable to changes in regional volumes, ranged from 2% to 24%, with the largest proportion for frontal gray matter. In contrast, this percentage ranged from 1% to 14% for executive functioning and from 2% to 12% for the eye-hand coordination domain.

The formal test for mediation of PTL's effect on visuo-construction and executive functioning was statistically significant (for a 5% type I error rate) when considering total brain and frontal gray matter regional volumes as mediators, respectively. In addition, the formal test of mediation of PTL's effect on executive functioning and eye-hand coordination was also statistically significant when the cingulate gyrus region was considered a mediator. We note that none of these effects would satisfy a strict Bonferroni-style multiplicity adjustment. However, clearly the regional volumes are quite correlated, so that such an analysis would be severely conservative.

Voxel-Wise Path Analysis

The voxel-wise mediation p-value maps are shown in Figures 3 (a)-(b), where the threshold was set at a 5% type I error rate. The results suggested that the mediation effect was most apparent for the visuo-construction domain [Figure 3 (a)], consistent with the ROI results. There were relatively fewer voxels with statistically significant mediation effects identified in the executive function and eye-hand coordination domains.

The liberal 5% threshold was initially used, but to address multiplicity issues, we also employed the Benjamini-Hochberg false discovery rate (FDR) procedure for multiple testing (1995; 2000). After such adjustment, few voxels and no relevant contiguous regions survived the threshold. It is noteworthy that the p-value distributions of the test for mediation were markedly non-uniform in the right tails [Figure 4 (d)]. In fact, the distribution suggested a conservative bias towards null results. In sequel, we conduct a simulation study to investigate this behavior more completely.

A Simulation Study

To evaluate the root causes of the non-uniformity of the distribution of the null p-values, we conducted a simulation study to investigate properties of tests for mediation. Thus we created a null distribution based on the empirical data in a setting where mediation is known to have occurred. For the simulation, voxels were represented by a 10 by 10 grid and a sum of these voxel values (representing the total volume in that area) then were associated with a cognitive function outcome (see below). The purpose of the simulation was to examine whether it was possible to identify a localized area of mediation via voxel-by-voxel path analysis.

To simulate realistic data we built on the three variable models considered in Figure 1, disregarding other covariates. We posited the existence of two important regions. We first let $\boldsymbol{A}$ be

the region where X had an impact on Z. That is, $\boldsymbol{A}$ was conceptually a region where PTL impacted brain volume. Secondly, consider another region, say $\boldsymbol{B}$, where Z had an impact on Y. The principal area of mediation, that is, the region where both lead impacts volume and volume impacts cognitive function, was then the area of overlap between these regions, say $\boldsymbol{C}$.

We first generated X measurements from a zero mean normal distribution. Subsequently, measurements for all voxels, regardless of their location, were then simulated from a zero mean normal distribution. The intensity of voxels within region $\boldsymbol{A}$, were shifted by an amount γ_2 times the X measurement, conceptually defining the area where lead impacts volume. Next, the total volume in area $\boldsymbol{B}$ was calculated by adding the voxel values in that region. The response Y was then simulated as a normal random variable with mean equal to γ_1 times the X measurement plus γ_3 times the volume in area $\boldsymbol{B}$. As such, this process mimicked settings where lead has an aggregate effect on area $\boldsymbol{A}$ while volume, through area $\boldsymbol{B}$, has an aggregate effect on cognitive function. Area $\boldsymbol{C}$, by virtue of being the area where both effects are present, represents the area of mediation. A more specific outline of the simulation process is given in Appendix A.

Two different scenarios were considered: the first where regions $\boldsymbol{A}$ and $\boldsymbol{B}$ completely overlapped and a second where regions $\boldsymbol{A}$ and $\boldsymbol{B}$ partially overlapped. That is, we considered the circumstance where the localized effect of lead on volume exactly overlapped the localized area where volume impacts cognitive function. For the first scenario, regions $\boldsymbol{A}$ and $\boldsymbol{B}$ were both in the square defined by pixels [2:5, 3:6] in the ten by ten dimensional grid defining the imaging space. In the second, regions $\boldsymbol{A}$ and $\boldsymbol{B}$ were in the square defined by the pixels [2:6, 2:6] and [3:7, 3:8] respectively.

The path coefficients γ_1 and γ_3 were set to be 0.8 and 1 respectively. The magnitude of γ_2 was varied at two different levels, moderate (γ_2=2) and strong (γ_2=5). The three measurement errors (Appendix A) were set at σ_1=0.05, σ_2=0.02 and σ_3=0.01, values consistent with the original

data.

The simulation results confirm that detecting localization of the mediation area was possible (Figure 5). There, the red highlighted area shows associations that were statistically significant via the test for mediation at *uncorrected* voxel-wise type I error rate of 0.05, and it clearly overlapped with the true mediating area used for simulation, ***C***. However, consider the distribution of the p-values obtained from the mediation test statistics (Figure 6), which ideally should be uniform with a spike near zero corresponding to the significant values. Instead, the distribution is right-skewed, with an excessive proportion of the values failing to reject the null hypothesis, mirroring the pattern observed in the voxel-wise results for the former lead workers. Furthermore, little of the mediation area survived strict, multiplicity adjusted, thresholds. Therefore, even though the mediation statistic appears to order the pixel-level results correctly, choosing an appropriate threshold was difficult due to the bias.

DISCUSSION

We considered decomposition of the total effect of cumulative lead dose on cognitive function using voxel-by-voxel path analysis as well as a parallel anatomically-derived ROI-based analysis. The ROI analysis suggested some degree of volumetric mediation in the three cognitive domains for which a statistically significant total effect was observed. It is worth noting that this ROI-based analysis presented herein differed from complimentary work found in (Caffo, Chen et al. 2007) by considering anatomically derived regions. However, the conclusions were similar. The most apparent mediators in the ROI-based analysis were total brain and frontal gray matter volumes for the association of PTl with performance in the visuo-construction domain. Other than for the cingulate gyrus in the executive functioning and eye-hand coordination domains, evidence for mediation was confined to larger areas, suggesting a more diffuse, non-localized pathway for lead to affect structure, and subsequently volume.

While the voxel-wise path analysis data results were in broad agreement with the ROI-based results, adequate thresholding remained an issue. In particular, multiplicity concerns are more of a factor for the voxel-wise approach, especially given the null bias of the p-values. This bias was confirmed by a simulation study and appeared to exist independent of the methods for calculating the standard error for the indirect effect, for which the multiple existing formulae differ from our own derivation; see (Sobel 1982). In addition, a computationally laborious non-parametric bootstrapping procedure was also performed to validate standard error calculations.

Explanations for this bias rely on the following violations of the standard assumptions for obtaining uniform p-value distributions in this setting: (i) there was statistical dependence arising from the model being applied across correlated voxel-level volumetric measurements; (ii) considering single voxels at a time may have resulted in an incorrect model, utilizing only the potentially minor mediation contribution of a single voxel and (iii) the multiplicative nature of the effect estimate, perhaps resulting in incorrect distributional assumptions for the test statistic, which in turn led to incorrectly calculated p-values. None of the three explanations appeared to dominate the process. For example, considering point (i) the p-value distributions for the path coefficients had marked uniform appearance, despite having the same statistical dependence issues. Point (ii) is somewhat addressed by the smoothing performed a-priori. However, as in point (i) we also note that this point appears to be much more problematic for investigating more subtle mediation hypotheses than considering direct effects associated with each path coefficient. In addition, the bandwidth of the smoother will certainly impact results, with too little smoothing making point (ii) more relevant and too much eliminating the possibility of localization. Point (iii) was considered by revisiting both the simulation study and the data analysis using the non-parametric bootstrap estimates of standard errors, yielding identical results. Moreover, another simulation setting (not shown), where the voxel-specific responses were independent, eliminated this bias despite the

product-type estimate and distributions used for simulation. Therefore it appeared to be an interaction of these departures that was the source of the bias.

In summary, we do not recommend the use of voxel-wise path analysis for the investigation of mediation despite being a clear generalization of voxel-based morphometry methods. Even in the highly idealized simulation setting, the distributional issues suggested that appropriate thresholding of p-value maps for mediation tests remains an unresolved question. Certainly, any multiplicity corrected threshold would be unduly pessimistic. In addition, the persistent null bias of the results casts doubts on the theoretical validity of the procedure.

Aside from the technical issues associated with performing voxel-wise path analyses, there remains the potential that the mediating effect of volume is diffuse or too variable across subjects to admit localized commonality across the population. As the simulation setting illustrated, it was possible for lead to have effects on cognition and a localized effect on volume, and for volume to have a localized effect on cognition, but no mediation exist. That is, if the regions do not overlap, no mediation will be present.

Table 1. Path analysis based on regional summaries (N=512). The fitted path model was adjusted for age, visit number and height. Three cognitive domain scores were considered, visuo-construction, executive functioning and eye-hand coordination. The mediation proportion represents the proportion of the total effect of lead on cognitive function that is indirect through brain volume. The indirect effects of total brain and frontal gray volumes (through a formal test of mediation) were statistically significant ($p < 0.05$) for the visuo-construction and executive functioning domains. In addition, the indirect effect was also statistically significance in the cingulate gyrus for the executive functioning and eye-hand coordination domains, and marginally so for visuo construction.

	Visuo-construction		Executive Functioning		Eye-hand Coordination	
	Mediation Prop. (%)	P-value	Mediation Prop. (%)	P-value	Mediation Prop. (%)	P-value
Total Brain	19	0.04	13	0.04	10	0.08
Total Gray	17	0.06	12	0.06	9	0.10
Total White	13	0.08	9	0.10	7	0.18
Frontal Gray	24	0.04	14	0.05	7	0.16
Occipital Gray	7	0.26	4	0.28	3	0.37
Parietal Gray	13	0.09	11	0.07	6	0.18
Temporal Gray	13	0.12	8	0.12	6	0.18
Frontal White	8	0.19	4	0.24	3	0.34
Occipital White	4	0.52	2	0.50	2	0.54
Parietal White	14	0.07	9	0.09	9	0.11
Temporal White	16	0.07	10	0.08	7	0.16
Cerebellum	2	0.49	4	0.30	5	0.26
Medial Temporal Lobe	12	0.16	8	0.16	5	0.23
Cingulate	14	0.06	13	0.05	12	0.05
Insula	13	0.08	7	0.16	8	0.13
Corpus Callosum	10	0.13	4	0.30	6	0.18
Internal Capsule	3	0.44	4	0.36	2	0.50
Hippocampus	4	0.34	2	0.40	4	0.31
Entorhinal Cortex	2	0.52	1	0.70	2	0.51
Amygdala	4	0.34	3	0.37	4	0.29

Figure 1 Simplified path model with a single mediator. *X* denotes the independent variable, *Z* the mediator and *Y* the dependent variable.

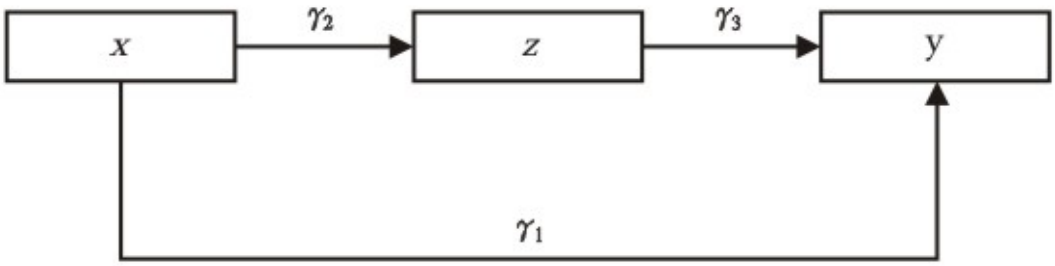


Figure 2 Path diagram used for the analysis of the former lead worker data presented herein. To examine to what extent brain volume mediates the association between peak tibia lead and measures of cognitive function, we have developed a two-stage path model relating PTL (*X*) to volume (*Z*) ,and PT (*X*) and volume (*Z*) to cognitive function (*Y*), at each stage including relevant covariates(age, the visit number and height.)

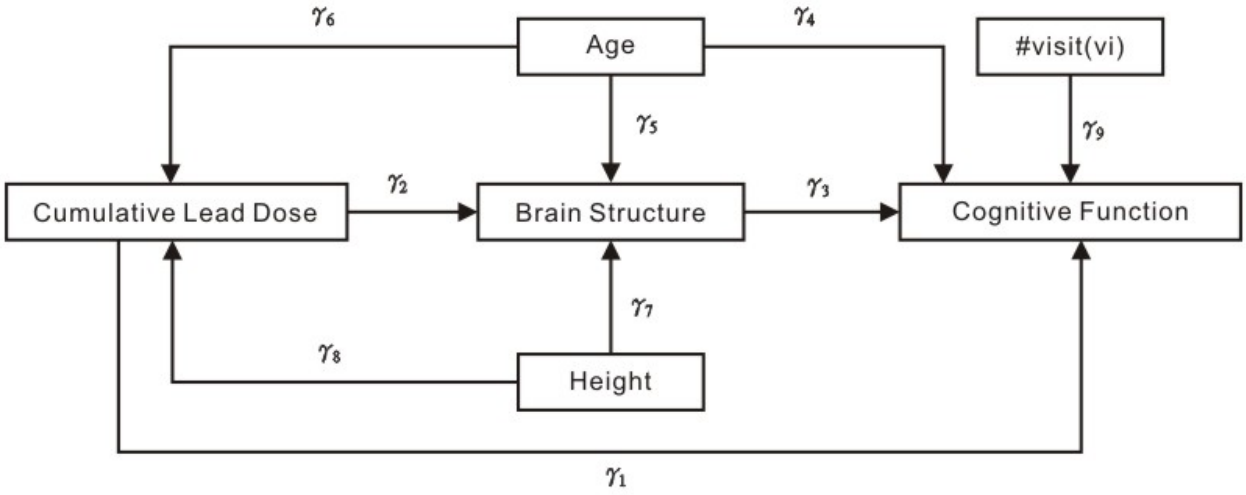
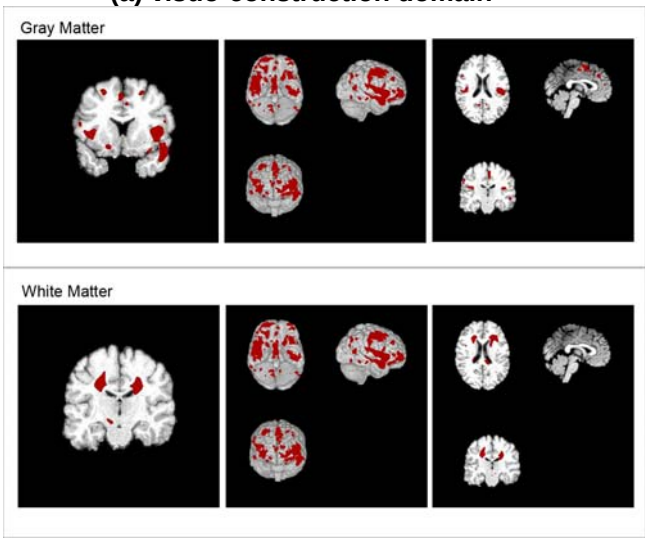
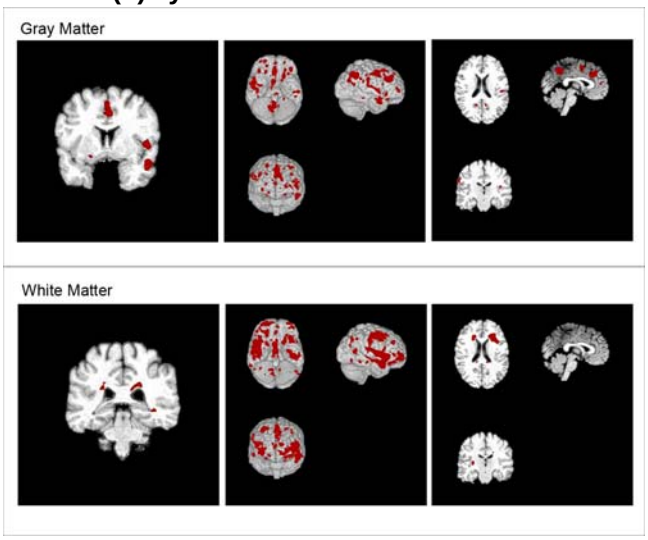


Figure 3 Spatial distribution of the mediation effect ((p-value threshold of 0.05)) for the former lead worker data(N=512) for three cognitive domains are shown for gray matter (A) and white matter (B). The results suggested that the mediation effect was most apparent for the visuoconstruction domain, consistent with the ROI results. There were relatively fewer voxels with statistically significant mediation effects identified in the executive function and eye-hand coordination domains.

(a) visuo-construction domain



(b) eye-hand coordination



(c) executive functioning

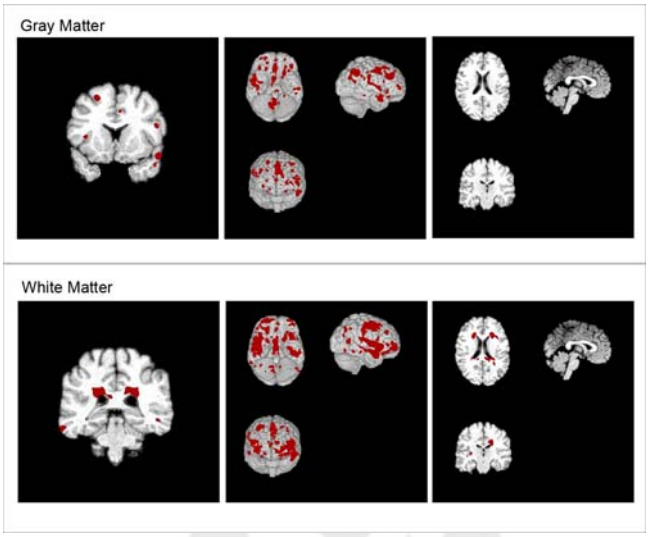


Figure 4 P-value distribution: white matter on eye-hand coordination : (a) p-value for γ_1; (b) p-value for γ_2; (c) p-value for γ_3; (d) p-value for testing $\gamma_2 * \gamma_3$. The distribution suggested that test for mediation was markedly non-uniform in the right tails while the non-uniformity is not present in tests for path coefficients.

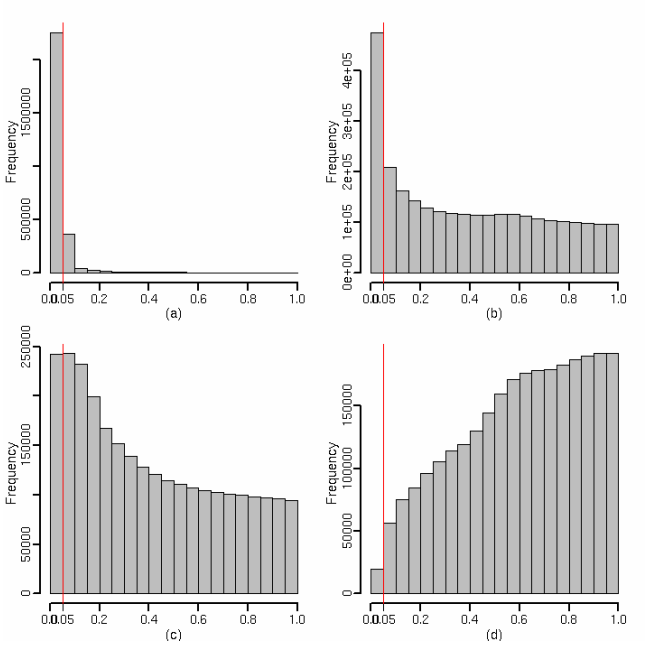


Figure 5 Spatial distribution of mediation effect from a representative simulation study. The four figures include the settings where: (a) regions ***A*** and ***B*** completely overlapped with a large γ_2; (b) ***A*** and ***B*** completely overlapped with a moderate γ_2; (c) ***A*** and ***B*** partially overlapped with a large γ_2; and (d) ***A*** and ***B*** partially overlapped with a moderate γ_2. The red highlighted area shows associations that were statistically significant via the test for mediation at *uncorrected* voxel-wise type I error rate of 0.05, and it clearly overlapped with the true mediating area used for simulation.

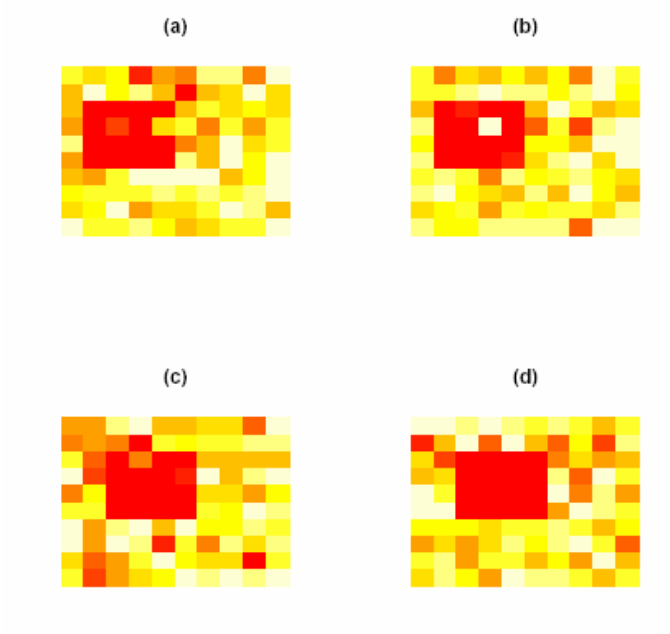
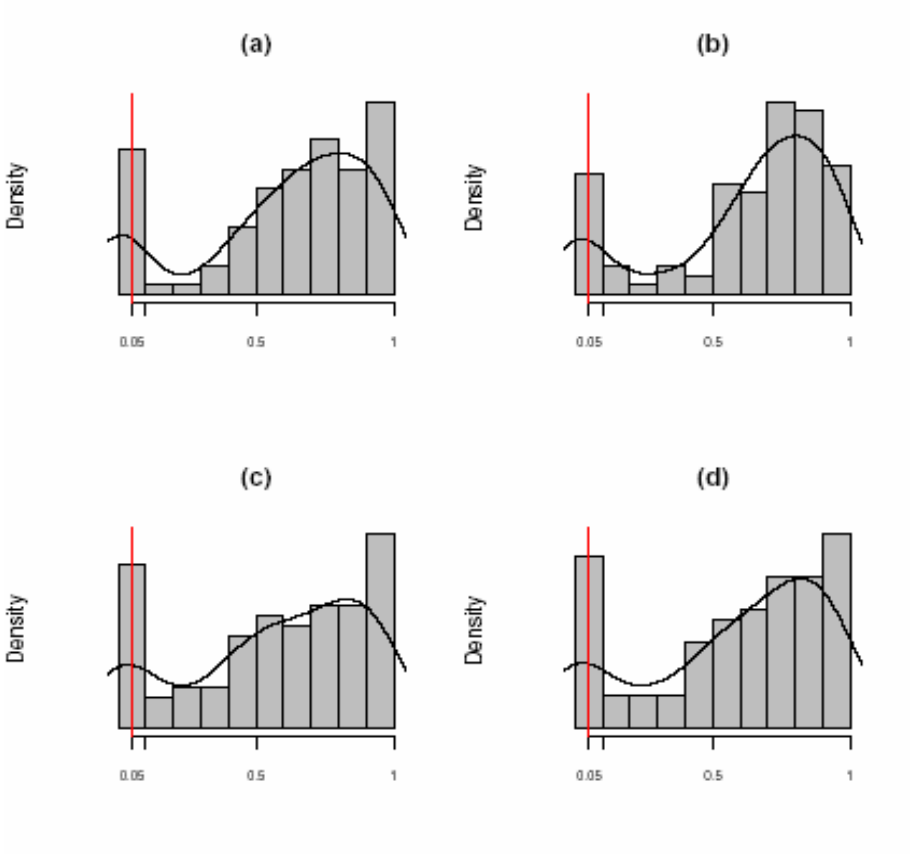


Figure 6 p-value distribution of mediation effect test from a representative simulation study. The four figures include the settings where: (a) regions ***A*** and ***B*** completely overlapped with a large γ_2; (b) ***A*** and ***B*** completely overlapped with a moderate γ_2; (c) ***A*** and ***B*** partially overlapped with a large γ_2; and (d) ***A*** and ***B*** partially overlapped with a moderate γ_2. the distribution is right-skewed, with an excessive proportion of the values failing to reject the null hypothesis, mirroring the pattern observed in the voxel-wise results for the former lead workers.



Appendix A

Outline of the Simulation Process

Consider a subject *i;* the simulation procedure follows as

1. Generate $X_i \sim N(0, \sigma_1^2)$

2. Generate $Z_i(v)$ data .

2.1 Generate a random sample, $\varepsilon_{2i}(v) \sim N(0, \sigma_2^2)$ at each pixel *v*

2.2 Define $Z_i(v) = I(v \in \mathbf{A}) * \gamma_2 * X_i + \varepsilon_{2i}(v)$ where I is an indicator of whether pixel *v* lies in region **A**.

3. Generate Y_i

3.1 Randomly sample $\varepsilon_{3i} \sim N(0, \sigma_3^2)$.

3.2 Let $W_i = \sum_{v \in C} Z_i(v)$ which is independent of pixel *v*.

3.3 $Y_i = \gamma_1 * X_i + \gamma_3 * W_i + \varepsilon_{3i}$.

REFERENCES

Ashburner, J. and K. Friston (2000). "Voxel-based morphometry - the methods." NeuroImage **11**: 805-821.

Ashburner, J. and K. Friston (2001). "Why voxel-based morphometry should be used." NeuroImage **14**(6): 1238–1243.

Benjamini, Y. and Y. Hochberg (1995). " Controlling the false discovery rate: a practical and powerful approach to multiple testing." Journal of the Royal Statistical Society **Ser. B,57**: 289–300.

Benjamini, Y. and Y. Hochberg (2000). "On the adaptive control of the false discovery rate in multiple testing with independent statistics." Journal of Educational and Behavioral Statistics **25**: 60–83.

Bishop, Y., S. Fienberg, et al. (1975). Discrete multivariate analysis: Theory and practice, Cambridge, MA: MIT Press.

Bollen, K. (1989). Structural equations with latent variables. New York, Wiley-Interscience.

Bollen, K. A. (1987). "Total, direct, and indirect effects in structural equation models." Sociological Methodology **17**: 37-69.

Buchel, C. and K. Friston (1997). Characterizing functional integration. Human Brain Function. R. S. J. Frackowiak, K. J. Friston, C. D. Frith, R. J. Dolan and J. C. Mazziotta, Academic Press USA**:** 127-140.

Buchel, C. and K. J. Friston (1997). "Modulation of connectivity in visual pathways by attention: Cortical interactions evaluated with structural equation modelling and fMRI." Cerebral Cortex **7**: 768-778.

Caffo, B., S. Chen, et al. (2007). Is mri based structure a mediator for cumulative environmental lead exposure's effect on behavioral function.

Davatzikos, C. (1996). "Spatial normalization of 3D brain images using deformable models." Journal of Computer Assisted Tomography **20**: 656-665.

Donaldson, S. I. (2001). Mediator and moderator analysis in program development. CA, Sage.

Holland, P. W. (1988). "Causal inference, path analysis, and recursive structural equation models." Sociological Methodology **18**: 449–484.

Joreskog, K. G. (1967). "Some contributions to maximum likelihood factor analysis." Psychometrika **32**(4): 443–482.

Judd, C. M. and D. A. Kenny (1981). "Process analysis: Estimating mediation in treatment evaluations." Evaluation Review **5**: 602–619.

Kraemer, H., T. Wilson, et al. (2002). "Mediators and moderators of treatment effects in randomized clinical trials." Archives of General Psychiatry **59**: 877-883.

Links, J. M., B. S. Schwartz, et al. (2001). "Characterization of toxicokinetics and toxicodynamics with linear systems theory: application to lead-associated cognitive decline." Environmental Health Perspectives **109**: 361-368.

McIntosh, A. and F. Gonzalez-Lima (1991). "Structural modeling of functional neural pathways mapped with 2-deoxyglucose: Effects of acoustic startle habituationthe auditory system." Brain Research **547**: 295-302.

McIntosh, A. and F. Gonzalez-Lima (1994). "Structural equation modeling and its application to network analysis in functional brain imaging." Human Brain Mapping **2**(2-22).

Mechelli, A., C. Price, et al. (2005). "Voxel-based morphometry of the human brain: Methods and applications." Current Medical Imaging Reviews: 105–113.

Schwartz, B. S., S. Chen, et al. (2007). "Relations of brain volumes with cognitive function in males 45 years and older with past lead exposure." Neuroimage **37**(2): 633-641.

Schwartz, B. S., W. F. Stewart, et al. (2000). "Past adult lead exposure is associated with longitudinal decline in cognitive function." Neurology **55**: 1144–1150.

Shih, R. A., T. A. Glass, et al. (2006). "Environmental lead exposure and cognitive function in community-dwelling older adults." Neurology **67**(9): 1556-1562.

Sobel, M. E. (1982). "Asymptotic confidence intervals for indirect effects in structural equation models." Sociological methodology **13**: 290-312.

Sobel, M. E. (1990). "Effect analysis and causation in linear structural equation models." Psychometrika **55**: 495–515.

Stewart, W., B. Schwartz, et al. (2006). "Past adult lead exposure is linked to neurodegeneration measured by brain mri." Neurology **66**: 1476–1484.

Stewart, W. F., B. S. Schwartz, et al. (1999). "Neurobehavioral function and tibial and chelatable lead levels in 543 former organolead workers." Neurology **52**(8): 1610-1617.

Woodworth, R. S. (1926). Dynamic Psychology, Worcester, MA: Clark University Press.

Wright, I., P. P. McGuire, J., et al. (1995). "A voxel-based method for the statistical analysis of gray and white matter density applied to schizophrenia." NeuroImage **2**: 244-252.

Wright, S. (1921). "Correlation and causation." Journal of Agricultural Research **20**: 557-585.

Wright, S. (1934). "The method of path coefficients." Annals of Mathematical Statistics **5**: 161-215.

Wright, S. (1960). "Path coefficients and path regressions: Alternative or complementary concepts?" Biometrics **16**: 189-202.