Toxicology Letters 228 (2014) 34–41

Contents lists available at ScienceDirect

Toxicology Letters

journal homepage: www.elsevier.com/locate/toxlet

Effects of environmental levels of cadmium, lead and mercury on human renal function evaluated by structural equation modeling

Jerome P. Trzeciakowski[a], Lesley Gardiner[a], Alan R. Parrish[b,*]

[a] Department of Medical Physiology, College of Medicine, Texas A&M Health Science Center, College Station, TX, USA
[b] Department of Medical Pharmacology and Physiology, University of Missouri School of Medicine, One Hospital Drive, Columbia, MO, USA

HIGHLIGHTS

- SEM is a novel method to model multi-variable, biological issues.
- The impact of cadmium on renal function was sigmoidal.
- The impact of lead on renal function was linear.

ARTICLE INFO

Article history:
Received 30 January 2014
Received in revised form 8 April 2014
Accepted 9 April 2014
Available online 21 April 2014

Keywords:
Cadmium
Kidney Function
Lead
Mercury
SEM
Structural Equation Model

ABSTRACT

A relationship between exposure to heavy metals, including lead and cadmium, and renal dysfunction has long been suggested. However, modeling of the potential additive, or synergistic, impact of metals on renal dysfunction has proven to be challenging. In these studies, we used structural equation modeling (SEM), to investigate the relationship between heavy metal burden (serum and urine levels of lead, cadmium and mercury) and renal function using data from the NHANES database. We were able to generate a model with goodness of fit indices consistent with a well-fitting model. This model demonstrated that lead and cadmium had a negative relationship with renal function, while mercury did not contribute to renal dysfunction. Interestingly, a linear relationship between lead and loss of renal function was observed, while the maximal impact of cadmium occurred at or above serum cadmium levels of 0.8 μg/L. The interaction of lead and cadmium in loss of renal function was also observed in the model. These data highlight the use of SEM to model interaction between environmental contaminants and pathophysiology, which has important implications in mechanistic and regulatory toxicology.

1. Introduction

Lead (Pb), cadmium (Cd) and mercury (Hg) are known to be nephrotoxic at high levels (Gonick, 2008; Sommar et al., 2013). Additive or synergistic interactions among heavy metals such that toxic effects may result even at low levels following environmental exposure (Fadrowski et al., 2010; Wallin et al., 2013; Weaver et al., 2011a). With few exceptions (Navas-Acien et al., 2009; Sanchez et al., 2001; Shelley et al., 2012), most studies have examined the effects of each metal in isolation. In this study, we have constructed a structural equation model (SEM) to identify and quantify the effects of lead, cadmium and mercury in a random subsample subjects participating in the National Health and Nutrition Examination Survey.

In structural equation modeling (SEM), a hypothesized model of relationships between variables is designed and then evaluated to determine if the experimental data supports that model. SEM has two key features: the measurement model, which defines the relationships between measurable variables and non-measurable latent factors, and the structural model, which delineates the path links and coefficients between and among the latent variables (Collin et al., 2009). As a modeling technique SEM has several advantages in that it allows for the modeling of complex, multivariate processes beyond simple correlations among single sets of

Abbreviations: ACR, urine albumin–urine creatinine ratio; BUN, blood urea nitrogen; CD, coefficient of determination; Cd, cadmium; CFI, comparative fit index; CrCl, creatinine clearance; eGFR, estimated glomerular filtration rate; Hg, mercury; H_0, null hypothesis; NHANES, National Health and Nutrition Examination Survey; Pb, lead; P_{close}, p-value for null hypothesis that RMSEA is ≤0.05; PSU, primary sampling unit; SCr, serum creatinine concentration; SD, standard deviation; SE, standard error; SEM, structural equation model; SHg, (total) serum mercury concentration; SPb, serum lead concentration; UCr, urine creatinine concentration; UHg, urine mercury concentration; UPb, urine lead concentration; VIF, variance inflation factor.

* Corresponding author at: Medical Pharmacology and Physiology, University of Missouri School of Medicine, MA 415 Medical Sciences Building, One Hospital Drive, Columbia, MO 65212, USA. Tel.: +1 573 884 4391.
E-mail address: parrishar@health.missouri.edu (A.R. Parrish).

http://dx.doi.org/10.1016/j.toxlet.2014.04.006

variables; it is not limited by measurable variables, but it allows for the inclusion of latent factors, i.e., factors that cannot be measured or observed on their own, but that can be expressed by measurable variables (Kline, 1991). SEM is also able to accurately measure unreliable events because it can quantify an error measurement that is indicative of errors including as biological variance. Most data sets are imperfect and more common modeling techniques, such as multiple regression and observed variable path analyses, cannot account for these flaws; however SEM compensates for these issues (Kline, 1991).

Although SEM has been utilized in the fields of sociology and psychology for many years, it is underutilized in the biological sciences. However, our group (Gardiner et al., 2012) and others (Fisher et al., 2011) have used SEM to model chronic kidney disease (CKD), which has complicated pathophysiology involving a number of factors. Given the complexity of assessing the relationship between environmental exposures and CKD, SEM is a valuable tool to begin to assess the relationship between heavy metals and renal dysfunction.

2. Methods

2.1. Study population

Demographic, laboratory and examination variables were obtained from 5 consecutive 2-year cycles of the National Health and Nutrition Examination Survey (continuous NHANES), which are made available online for public use by the Centers for Disease Control (Centers for Disease Control and Prevention 1999–2008). Because subjects are identified by sequence number only, no special permissions are required to access the data. Of the 51,653 subjects that were both interviewed and examined from 1999–2008, 30,257 had simultaneous entries for serum lead, serum cadmium and serum (total) mercury; 8847 had entries for detectable levels of urine lead, urine cadmium and urine mercury. Blood metal measures were available on all but 551 of the subjects with urine metal values. Individuals missing one or more measures of kidney function were excluded from this set, which left $n = 7236$ subjects for analysis. Demographic and other relevant characteristics of the study population are listed in Table 1.

Table 1
Demographics and related characteristics of the subjects.

Age	
Mean (yrs)	40.99 ± 0.40[a]
Distribution	
0–5 yrs	0%
6–11 yrs	0%
12–19 yrs	12.3%
20–39 yrs	37.0%
40–59 yrs	34.0%
60+	16.6%
Race/ethnicity	
Mexican American	8.6%
Other hispanic	4.9%
NonHispanic white	70.3%
NonHispanic black	11.5%
Other races	4.7%
Sex	
Male	41.0%
Female	59.0%
Kidney status	
GFR ≥ 60 ml/min/1.73 m^2	94.6%
GFR < 59 ml/min/1.73 m^2	5.4%
Macroalbuminuria	
Absent (ACR < 300 mg/g)	99.0%
Present (ACR ≥ 300 mg/g)	1.0%
Body measures	
BMI (kg/m^2)	27.71 ± 0.14[a]
Waist circumference (cm)	94.60 ± 0.39[a]
Weight (kg)	78.36 ± 0.46[a]

[a] Mean ± linearized SE.

2.2. Data preparation

Because of changes in assay methods, serum creatinine values for the 1999–2000 and 2005–2006 data sets had to be adjusted to ensure comparability with standard creatinine (Selvin et al., 2007). Creatinine clearance was calculated from the corrected serum creatinine values using the Cockcroft–Gault formula (Cockcroft and Gault, 1976). Albuminuria was calculated as the ratio of urine albumin to urine creatinine (ACR) expressed in units of mg/g. Limits of detection for blood and urine metals varied slightly across the survey cycles. In those subjects where the result was below the limit of detection, a concentration equal to the limit of detection divided by the square root of two was used (Centers for Disease Control and Prevention 2007–2008; Centers for Disease Control and Prevention, 2009, 2013). Metal concentration data that contain values below a lower detection limit are referred to as left-censored or censored from below. Excluding metal concentrations below the limit of detection (LOD) is not recommended, as it not only reduces the sample size but also yields upwardly biased results (Hornung and Reed, 1990). A number of methods have been proposed for handling values falling below the LOD (Helsel, 2010). A fraction of the LOD (e.g. LOD/2 or LOD/$\sqrt{2}$) is often substituted for the problem values in regression modeling. Metal concentrations falling below the LOD in NHANES surveys are pre-transformed by substitution with LOD/$\sqrt{2}$ prior to publishing, and have been used in this format by investigators working with NHANES data (Navas-Acien et al., 2009; Shelley et al., 2012). The bias is small if the percentage of data below the LOD is small and the data are not highly skewed (Baccarelli et al., 2005).

Given recent concerns over the use of data substitution, we investigated an alternate method for handling the problem: multiple imputation. Model-based multiple imputation is an alternative to substitution for left-censored data (Baccarelli et al., 2005; He et al., 2010). To examine the effect of multiple imputation in this study, each metal concentration below the LOD was first replaced by a missing value code. Then, for each missing value, 20 new values were generated using Markov Chain Monte Carlo (MCMC) simulations, to create 20 complete data sets containing no missing values (Rubin, 1987; Schafer, 1997). These data sets were then used as the basis for imputation by Bayesian estimation of the SEM model in MPlus. Briefly, the SEM model was run for each of the 20 complete data sets, and combined by the MPlus program into a single set of results that incorporated uncertainty due to the missing data. An assumption of multiple imputation is that the data are missing at random. To the extent that metal values falling below the LOD may not comply with this assumption, some bias is expected. There was no substantial difference in structure, coefficients or fit indices between the model derived from multiple imputation and that derived using LOD/$\sqrt{2}$ substitution (data not shown). This suggests that for this data set, the SEM model is robust to changes in the method of handling metal concentrations falling below the limit of detection.

All continuous variables were tested for normality prior to analysis. The Jarque–Berra statistic (Jarque and Bera, 1980) provides a sensitive index of both skewness and kurtosis and was used to evaluate the need for transformation. Based on this metric, all 10 observed variables used in the model were found to require log-transformation prior to analysis. We tested for possible multicollinearity among the transformed measures by computing variance inflation factors (VIF): these were found to be negligible (VIF range: 1.06–2.23, mean = 1.77). When two or more predictor variables in a statistical model are highly correlated, it becomes difficult to statistically determine which variable has the most impact on the predicted result. The variables are collinear, and the results show what is termed multicollinearity. Multicollinearity increases the variance (standard errors) of the model coefficients and can cause what should be significant predictors to be considered non-significant. Variance inflation factors measure how much the variances of the estimated coefficients are increased over what they would be in the absence of correlations among the predictor variables. There is no formal cutoff for the upper limit of acceptable VIF values; however values above 5 are a usually a cause of concern and values of 10 are a definite indicator of extreme multicollinearity (Kutner et al., 2004).

NHANES guidelines recommend using weights corresponding to the smallest subpopulation containing any of the variables of interest. Metals in urine were generally measured in a random 1/3rd subsample of the participants; therefore these subsample weights were used for constructing combined 10-year sample weights across survey cycles.

2.3. Data analysis

Development and testing of the structural equation model was performed with MPlus software (version 6.11, Muthén and Muthén; www.StatModel.com). Data were imported from an ASCII file in free format, with each row representing a subject and each column a variable. In addition to the 10 observed variables, the data contained stratification, cluster (PSU), and sample weight variables as required for analysis of complex survey designs. Sample correlation and covariance matrices for the data are provided in Table 2. The default estimation method in MPlus for survey data is a maximum likelihood estimator (MLR) that results in parameter estimates and standard errors that are robust to non-normality and non-independence of observations when used with complex data types. Stata for Windows (version 13, StataCorp LP; www.stata.com) was used for constructing the diagram of the model solution. Statistical analysis of path effects and predicted latent factor scores