Science of the Total Environment 650 (2019) 970–977

Contents lists available at ScienceDirect

Science of the Total Environment

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

Explaining variation in elevated blood lead levels among children in Minnesota using neighborhood socioeconomic variables

David C. Wheeler [a,*], Resa M. Jones [b], Mario Schootman [c], Erik J. Nelson [d]

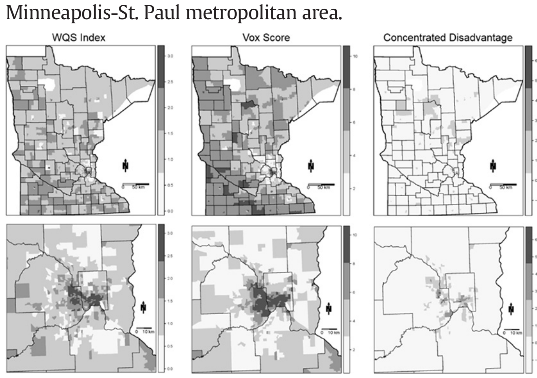
[a] Department of Biostatistics, Virginia Commonwealth University, United States of America
[b] Department of Epidemiology and Biostatistics, College of Public Health, Temple University and Fox Chase Cancer Center, Temple University Health, United States of America
[c] Department of Clinical Analytics and Insights, Center for Clinical Excellence, SSM Health, St. Louis, United States of America
[d] Department of Epidemiology and Biostatistics, Indiana University-Bloomington, United States of America

HIGHLIGHTS

- Elevated blood lead levels are significantly related to area socioeconomic status.
- Weighted quantile sum regression best explains elevated blood level risk.
- Percent of houses built prior to 1940 is most important for elevated blood levels.
- The conceptual model to identify areas of elevated blood lead risk is complex.
- Our complex model improves on existing approaches at explaining risk.

GRAPHICAL ABSTRACT

Estimated weighted quantile sum (WQS) index, Vox lead exposure score, and concentrated disadvantage index from principal components analysis to explain elevated blood lead level rates across census tracts in Minnesota with county lines drawn for reference. The top panel is the entire state and the bottom panel is focused on the Minneapolis-St. Paul metropolitan area.



ARTICLE INFO

Article history:
Received 2 June 2018
Received in revised form 7 September 2018
Accepted 7 September 2018
Available online 08 September 2018

Editor: D. Barcelo

Keywords:
Lead
Weighted quantile sum regression
Vox
Concentrated disadvantage
Socioeconomic status

ABSTRACT

Background: Childhood lead exposure is linked to numerous adverse health effects and exposure in the United States is highest among people living in substandard housing, which is disproportionately inhabited by socioeconomically disadvantaged individuals. In this paper, we compared the Vox lead exposure risk score and concentrated disadvantage based on principal component analysis (PCA) to weighted quantile sum (WQS) regression to determine which method was best able to explain variation in elevated blood lead levels (EBLLs).
Methods: We constructed indices for census tracts in Minnesota and used them in Poisson regression models to identify the best socioeconomic measure for explaining EBLL risk.
Results: All indices had a significant association with EBLL in separate models. The WQS index had the best goodness-of-fit, followed next by the Vox index, and then the concentrated disadvantage index. Among the most important variables in the WQS index were percent of houses built before 1940, percent renter occupied housing, percent unemployed, and percent African American population.
Conclusions: The WQS approach was best able to explain variation in EBLL risk and identify census tracts where targeted interventions should be focused to reduce lead exposure.

* Corresponding author.
E-mail address: dcwheeler@vcu.edu (D.C. Wheeler).

https://doi.org/10.1016/j.scitotenv.2018.09.088
0048-9697/© 2018 Published by Elsevier B.V.

1. Introduction

Lead is a widespread environmental contaminant and has been linked to numerous adverse health effects in children, particularly neurological and neurobehavioral deficits, lower IQ, slowed growth, and anemia (Agency for Toxic Substances and Disease Registry (ATSDR), 2007; Canfield et al., 2003; Chiodo et al., 2004; Grandjean and Landrigan, 2014; Lanphear et al., 2000; Lidsky and Schneider, 2003; Miranda et al., 2007; Nelson et al., 2015; Schnaas et al., 2000; Tellez-Rojo et al., 2006; Mielke et al., 1997; Mielke et al., 2017; Mielke et al., 2016). Lead can be ingested from a variety of sources including lead-based paint, household dust containing lead paint, soil, drinking water, and food (Mielke et al., 1997). Although there is no safe blood lead threshold in children, the U.S. Centers for Disease Control and Prevention recommends taking public health actions to reduce future lead exposure for children with blood lead levels (BLLs) at or above 5 μg/dL (Centers for Disease Control and Prevention (CDC), 2012; Wengrovitz et al., 2009). During 2007–2010, the percentage of children aged 1–5 years with BLLs at or above 5 μg/dL was 2.6%, or an estimated 535,000 children in the U.S. with elevated BLLs (EBLLs) (Centers for Disease Control and Prevention (CDC), 2013). Despite efforts by state and local health departments to reduce BLLs in children, the Healthy People 2020 objective to reduce BLLs to an average of 1.6 μg/dL is not likely to be achieved in the near future (Centers for Disease Control and Prevention (CDC), 2004; US Department of Health and Human Services, 2012). This may be due in part to the difficulty in identifying where to target remediation and prevention efforts because it is not feasible to obtain blood from children in a population-based manner. However, the best method to identify areas of EBLLs without obtaining blood is unknown.

Because risk of EBLL is highest among persons living in substandard housing, which are often inhabited by racial minorities and socioeconomically disadvantaged persons (Campanella and Mielke, 2008; Leech et al., 2016), socioeconomic measures of deprivation (e.g., Gini coefficient (Gini, 1997), population below the federal poverty level (U.S. Census Bureau, 2017), concentrated disadvantage (Sampson et al., 1997)) have been used to estimate risk of EBLLs for a variety of areal units (e.g., block groups, census tracts, ZIP Codes) (Boutwell et al., 2016; Hanna-Attisha et al., 2016; Krieger et al., 2003; Aelion et al., 2013; Carrel et al., 2017). In 2016, Vox published an online interactive map of lead exposure risk across census tracts in the U.S (Vox, 2016a). The Vox method calculated a score between 1 and 10 by weighting the joint effects of the proportion of the population living below the federal poverty level and the age of the housing stock based on household lead hazard estimates derived by Jacobs et al. (2002). While the Vox risk score has returned national attention to lead exposure, it uses only a small subset of the area-level covariates, ignoring the fact that many other area-level variables have been associated with EBLLs (Carrel et al., 2017; Jones et al., 2010; Moody and Grady, 2017). If these other variables are not included, then the score may not correctly identify all areas of high risk for EBLLs. This may mean that the Healthy People 2020 objective to reduce BLLs remains elusive.

In contrast to the variables included in the Vox score, area-level concentrated disadvantage uses more covariates to identify areas that are disadvantaged and potentially at increased risk for EBLL. The construction of this index is often based on principal components analysis (PCA). A potential limitation of PCA to construct the concentrated disadvantage index is that it is typically constructed based on the correlation or covariation pattern among the area-level variables without consideration of the relationship between these variables and the health outcome of interest. As a result, these indices may include variables that are not associated with the outcome and therefore may also not correctly identify all areas of high risk for EBLLs.

A recently developed method that may circumvent the limitations of both the Vox and PCA methods is weighted quantile sum (WQS) regression (Carrico et al., 2015). WQS regression is designed to accommodate correlated data when constructing an index. It can estimate both the effect of an index on a health outcome and the corresponding weights for each variable included in that index. Additionally, the estimated component weights in the index can be interpreted as measures of relative variable importance.

Given the possible limitations of the Vox and PCA methods, our objective was to construct an area-level socioeconomic status (SES) index using the newly developed Poisson WQS regression and compare its ability to explain the risk of EBLLs in Minnesota with the Vox risk score and concentrated disadvantage from PCA.

2. Methods

2.1. Study design

We assessed the association between various potential indicators of BLL risk and risk of EBLL across 1332 census tracts (out of 1338 tracts) in Minnesota from 2011 to 2015 using an ecological design. We selected Minnesota for this study because the statewide recommendation to perform childhood lead testing was similar to most states (Safer Chemicals Heathier Families, 2017), the Minnesota Department of Health has high lead surveillance reporting standards (State of Minnesota, 2017), and the data were publicly available.

2.2. Data

2.2.1. Blood lead levels

We obtained the counts and proportions of EBLL tests (≥5 μg/dL) among children <72 months in age who had BLL tests performed in Minnesota from 2011 to 2015 (Minnesota Department of Health, 2017). The test results included in our analyses are limited to one sample per child. It is important to note that lead testing is not universal in Minnesota and that children with risk factors for lead exposure (such as those living in older housing and families living in poverty) are targeted for testing. However, all blood lead tests performed in Minnesota are mandatorily reported to the Minnesota Department of Health Blood Lead Information System (State of Minnesota, 2017), thus the samples included in the analysis are representative of all blood lead testing conducted throughout the state. It should be noted that counts of fewer than four cases per census tract were suppressed, though the percentage of EBLLs were provided in these areas. For the suppressed tracts, we estimated the number of cases by multiplying the percentage of EBLLs by the total number of children tested and constrained these estimates to render between zero and four cases. Specifically, if a census tract reported between “1-4 EBLL cases” and an EBLL percent of 1.1% for a census tract with 115 children who were tested, the resulting product ($0.011 * 115 = 1.27$) was rounded to the nearest integer (e.g., 1 in the example provided) to estimate the number of EBLL cases in that census tract for our analyses. BLL tests were not available for six census tracts and these tracts were therefore excluded from the analysis. Further details regarding all aspects of data collection and reporting of lead exposure in Minnesota are available elsewhere (Minnesota Department of Health, 2018). A map of the crude rates of EBLLs among those tested is shown in Fig. 1. Some of the highest rates of EBLLs are in the Minneapolis-St. Paul metropolitan area.

2.2.2. Concentrated disadvantage

We constructed an index of concentrated disadvantage using PCA of eight variables from the 2011–2015 American Community Survey (U.S. Census Bureau, 2017) (ACS) using the R package psych (Revelle, 2017). Based on prior research (Sampson et al., 1997), we included the proportion of the population that was African-American, proportion of female-headed households, proportion of households receiving food stamps, proportion of uninsured individuals (neither private or public health insurance programs), percentage of households without employment during the past 12 months, proportion of households below the federal