Environmental Research 151 (2016) 505–512

Contents lists available at ScienceDirect

Environmental Research

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

Impacts of geocoding uncertainty on reconstructed PFOA exposures and their epidemiological association with preeclampsia ☆

Raghavendhran Avanasi [a,b,*], Hyeong-Moo Shin [c], Veronica M. Vieira [a,d], Scott M. Bartell [a,d,e]

[a] Environmental Health Sciences Graduate Program, 2032, AIRB, University of California, Irvine, CA 92697-3957, USA
[b] ICF International. Fairfax, Virginia, USA
[c] Department of Public Health Sciences, One Shields Avenue, MS1-C, Davis, CA 95616-8638, USA
[d] Program in Public Health, AIRB, University of California, Irvine, CA 92697-3957, USA
[e] Department of Statistics and Department of Epidemiology, University of California, Irvine, CA, USA

ARTICLE INFO

Article history:
Received 10 May 2016
Received in revised form
16 August 2016
Accepted 17 August 2016

Keywords:
C8 Health Project
Geocoding uncertainty
Differential exposure mischaracterization
Perfluorooctanoate
Preeclampsia

ABSTRACT

Many epidemiology studies have investigated associations of perfluorooctanoate (PFOA) exposures with a variety of adverse health outcomes for participants in the C8 Health Project. The exposure concentrations (i.e., air and groundwater) used in these studies were determined primarily based on participant's residential locations. However, for residential addresses that could not be geocoded to the street level, the exposure concentrations were assigned based on population-weighted ZIP code centroid, which may result in exposure mischaracterization. The aim of this current study is to evaluate the potential impact of mischaracterized exposure concentrations due to geocoding uncertainty on the predicted serum PFOA concentrations and the epidemiological association between PFOA exposure and preeclampsia. For both workplace addresses and incompletely geocoded residential addresses, we used Monte Carlo (MC) simulation to assign alternate geographic locations within the reported ZIP code (instead of population-weighted ZIP code centroids) and the corresponding exposure concentrations. We found that mischaracterization of residential exposure due to population-weighted ZIP code centroid assignment had no significant impact on the serum PFOA concentration predictions and the epidemiological association of PFOA exposure with preeclampsia. In contrast, the uncertainty in workplace exposure moderately impacted the rank exposure among the participants. We observed a 41% increase in the average adjusted odds ratio of preeclampsia occurrence that may be due to differing proportions of cases (64.3%) and controls (54.5%) with workplace address geocodes during pregnancy. This finding suggests that differential exposure mischaracterization can be reduced by obtaining accurate exposure information such as street addresses and tap water consumption, for both workplaces and residences. The analysis we present is one approach for estimating the potential impacts of positional errors in a geocoding-based exposure assessment on exposure estimates and epidemiological study results.

1. Introduction

Geographic Information Systems (GIS) have been used in numerous environmental health studies for assessing the exposure of participants to contaminants of interest via proximity analysis, integration of environmental monitoring data, individual-level exposure estimation, design of exposure metrics, and reconstructing exposure through activity patterns (Ali et al., 2002; Bell et al., 2015; Bellander et al., 2001; Beyea and Hatch, 1999; Elgethun et al., 2003; Floret et al., 2003; Nuckols et al., 2004; Reynolds et al., 2003; Rull and Ritz, 2003; Shin et al., 2011a; Vieira et al., 2010, 2013). The use of GIS in environmental exposure assessment can improve our understanding of the associations between environmental exposures and adverse health outcomes (Beyea and Hatch, 1999; Nuckols et al., 2004).

Geocoding, the process of matching addresses to geographic locations (latitude and longitude), is an important step in using GIS for exposure assessment (Bonner et al., 2003). One primary application of geocoding is to assign individual-level environmental exposures based on their location in an exposed

Abbreviations: GIS, Geographic Information Systems; C8, PFOA Perfluorooctanoate; PWD, Public Water District; MC, Monte Carlo; AOR, Adjusted Odds Ratio; IQR, Inter Quartile Range; CI, Confidence Interval; PI, Probability Interval

☆We obtained approval (HS#2013-9421) from the Institutional Review Board at the University of California, Irvine, to work with the human subject data in this current study.

* Corresponding author.
E-mail address: ragaavanasi@gmail.com (R. Avanasi).

http://dx.doi.org/10.1016/j.envres.2016.08.019
0013-9351/

geographic area (Elgethun et al., 2003; Shin et al., 2011a; Vieira et al., 2013; Ward et al., 2005). Partial matching of addresses, such as a street name without the house number or a ZIP code without a specific street, or errors in geocoding can lead to positional errors in the exposure assessment, potentially leading to exposure mischaracterization. This can impact the validity of the epidemiological studies that use the resulting exposure estimates (Bonner et al., 2003; Elgethun et al., 2003; Vieira et al., 2010, 2013). Researchers and the National Institutes of Health have called for more investigation into the potential impacts of geocoding uncertainty on the results of epidemiological studies (Henry and Boscoe, 2008; US Department of Health and Human Services, 2014; Zandbergen, 2009). A recent report from a Health and Environmental Sciences Institute (HESI) workshop also recommended the characterization and evaluation of uncertainty in environmental epidemiology studies to better understand the potential sources of bias and to utilize results from epidemiological analyses for risk assessment (Burns et al., 2014).

The C8 Science Panel studies investigated associations of perfluorooctanoate (PFOA) serum concentrations predicted by a GIS-based exposure assessment (Shin et al., 2011a, 2011b) with a variety of adverse health outcomes such as ulcerative colitis, kidney and testicular cancer, pregnancy outcomes, abnormal thyroid function, and abnormal kidney function (Barry et al., 2013; C8 Science Panel, 2011; Lopez-Espinosa et al., 2012; Savitz et al., 2012a, 2012b; Steenland et al., 2013; Watkins et al., 2013). Predicted serum PFOA concentrations for 2005–2006 were well correlated (r_s=0.68) with measured serum PFOA concentrations in the same year. Geocoding was used to locate participant residential addresses geographically to assign air and water PFOA concentrations for each year, over 58 years-1951 to 2008. This was done by spatially joining the addresses with the pipe distribution networks of the six participating public water districts (PWDs) to which all the consented participants of the C8 Health Project belonged (Shin et al., 2011b; Vieira et al., 2013). About 12% of the addresses (mostly rural addresses) with ZIP codes within the six PWDs could not be geocoded and thus population weighted ZIP code centroids were used to assign PWDs and the corresponding PFOA water concentrations. The assignment of population weighted ZIP code centroids for addresses that could not be geocoded to the street level can be considered as a single geographic imputation method (analogous to a mean imputation method). Such imputation or geocoding at a coarse spatial resolution can introduce geographic bias/positional errors in the exposure classification (Henry and Boscoe, 2008; Zandbergen, 2009). Also, it has been noted that there is greater potential for positional errors when geocoding rural addresses compared to geocoding urban addresses (Vieira et al., 2010; Ward et al., 2005).

The aim of this study is to evaluate the potential impacts of geocoding uncertainty on the estimated serum PFOA concentrations of participants in the C8 Health Project. Specifically, we examine the impacts of single geographic imputation, which may have resulted in mischaracterized water PFOA concentrations for those participants geocoded to population-weighted ZIP code centroids. We also examine the corresponding impact on the association between the estimated serum PFOA concentrations and the occurrence of preeclampsia (Savitz et al., 2012), an epidemiological analysis that has been discounted for the use of modeled rather than directly measured serum PFOA concentrations (Johnson and Sutton, 2014; Koustas et al., 2014). We use Monte Carlo (MC) simulation to assign alternate geographic locations within the reported ZIP code for all residential addresses that were geocoded to a population-weighted ZIP code centroid and the reported work addresses, and recalculate the prediction of serum PFOA concentrations and the epidemiological association with preeclampsia for each set of alternate geographic locations.

2. Materials and methods

2.1. PFOA exposure assessment

The PFOA exposure assessment by Shin et al. (2011a, 2011b) had two distinct modeling components. The first part of the PFOA exposure assessment used a suite of environmental fate and transport models to predict yearly PFOA outdoor air and groundwater concentrations for 1951–2008 in the region surrounding the Washington Works facility and the six impacted PWDs. Detailed explanation of the PFOA fate and transport modeling can be found in Shin et al. (2011a). Briefly, the modeling system utilized yearly PFOA release rates from the Washington Works facility, along with PFOA physicochemical properties, local meteorology, and hydrogeology to predict the yearly air and water concentrations of PFOA for the area serviced by the six PWDs: the City of Belpre, Little Hocking Water Association, Tuppers Plains Chester Water District, the Village of Pomeroy Water District, Lubeck Public Service District, and Mason County Public Service District. The model also estimated PFOA exposure for shallower private drinking water wells located in the study area.

Next, an integrated exposure and pharmacokinetic model system was used to predict the yearly serum PFOA concentrations for all consented participants in the C8 Health Project study. This model system utilized the predicted yearly PFOA air and water concentrations (Shin et al., 2011a), standard inhalation and standard/self-reported tap water ingestion rates (U.S. EPA, 2009), PWD pipe distribution networks, along with self-reported participant information collected through a questionnaire as part of the C8 Health Project (Frisbee et al., 2009). These included detailed participant residential/work histories and participant demographics such as age, gender, and body weight. Based on the self-reported information, the drinking water source at each residential history was categorized as public, private, bottled water, or mixed. GIS was then used to link participant residential addresses with modeled air and water PFOA concentrations and predict yearly combined inhalation and ingestion (total) exposure doses for all the participants. A one-compartment pharmacokinetic model was then applied to estimate the yearly serum PFOA concentrations based on a single elimination half-life. More details on the exposure reconstruction/pharmacokinetic modeling are described by Shin et al. (2011b).

2.2. GIS

GIS methods were used to assign historical outdoor air and groundwater concentrations for each participant. With respect to ingestion exposure, GIS was used first to map the pipe distribution systems of the six PWDs included in the exposure modeling system (Shin et al., 2011b). Next, the participant residential addresses were geocoded using TeleAtlas and ArcView/NAVTEQ (Vieira et al., 2010). Among the residential addresses with ZIP codes in any of the six PWDs, approximately 12% of the addresses could not be geocoded to the street level (Shin et al., 2011b; Vieira et al., 2013) and hence, a population-weighted ZIP code centroid was used to assign environmental concentrations instead of the street level geocode. Later, within the GIS, the geocoded addresses were spatially joined with the PWD pipe distribution system to assign PWD-specific annual average PFOA water concentrations to those addresses that were serviced by any specific PWD. As described in the text and Shin et al. (2011b, Fig. 1) study, based on the participant's geocoded residential address, the PFOA water concentrations were assigned for each reported residence for each participant. Any discrepancies between the self-reported water sources and the geocoded water sources (~9% of the addresses) were reviewed manually to determine the most likely source. For the