Variability and epistemic uncertainty in water ingestion rates and pharmacokinetic parameters, and impact on the association between perfluorooctanoate and preeclampsia in the C8 Health Project Population

Raghavendhran Avanasi[1], **Hyeong-Moo Shin**[2], **Veronica M. Vieira**[1,3], and **Scott M. Bartell**[1,3,4]

[1]Environmental Health Sciences Graduate Program, 2032, AIRB, University of California, Irvine, CA 92697-3957

[2]Department of Public Health Sciences, One Shields Avenue, MS1-C, Davis, CA 95616-8638

[3]Program in Public Health, AIRB, University of California, Irvine, CA 92697-3957

[4]Department of Statistics and Department of Epidemiology, University of California, Irvine, California, U.S.A

Abstract

We recently utilized a suite of environmental fate and transport models and an integrated exposure and pharmacokinetic model to estimate individual perfluorooctanoate (PFOA) serum concentrations, and also assessed the association of those concentrations with preeclampsia for participants in the C8 Health Project (a cross-sectional study of over 69,000 people who were environmentally exposed to PFOA near a major U.S. fluoropolymer production facility located in West Virginia). However, the exposure estimates from this integrated model relied on default values for key independent exposure parameters including water ingestion rates, the serum PFOA half-life, and the volume of distribution for PFOA. The aim of the present study is to assess the impact of inter-individual variability and epistemic uncertainty in these parameters on the exposure estimates and subsequently, the epidemiological association between PFOA exposure and preeclampsia. We used Monte Carlo simulation to propagate inter-individual variability/ epistemic uncertainty in the exposure assessment and reanalyzed the epidemiological association. Inter-individual variability in these parameters mildly impacted the serum PFOA concentration predictions (the lowest mean rank correlation between the estimated serum concentrations in our study and the original predicted serum concentrations was 0.95) and there was a negligible impact on the epidemiological association with preeclampsia (no change in the mean adjusted odds ratio (AOR) and the contribution of exposure uncertainty to the total uncertainty including sampling variability was 7%). However, when epistemic uncertainty was added along with the inter-

Corresponding Author: Scott Bartell, PhD, Associate Professor, 2032 AIRB, Program in Public Health, University of California, Irvine, Irvine, CA 92697-3957, Phone: (949) 824-5919, Fax: (949) 824-0529, sbartell@uci.edu.

individual variability, serum PFOA concentration predictions and their association with preeclampsia were moderately impacted (the mean AOR of preeclampsia occurrence was reduced from 1.12 to 1.09, and the contribution of exposure uncertainty to the total uncertainty was increased up to 33%). In conclusion, our study shows that the change of the rank exposure among the study participants due to variability and epistemic uncertainty in the independent exposure parameters was large enough to cause a 25% bias towards the null. This suggests that the true AOR of the association between PFOA and preeclampsia in this population might be higher than the originally reported AOR and has more uncertainty than indicated by the originally reported confidence interval.

Keywords

C8 Science Panel; measurement error; individual-level exposure uncertainty; perfluorooctanoate; preeclampsia

1.1 Introduction

Many recent environmental epidemiology studies evaluating associations between perfluorooctanoate (PFOA) and various health effects including ulcerative colitis, kidney and testicular cancer, pregnancy outcomes, abnormal thyroid function, abnormal liver function, and abnormal kidney function have been based on participants from the C8 Health Project/C8 Science Panel Studies (Barry et al., 2013; C8 Science Panel, 2011; Gallo et al., 2012; Lopez-Espinosa et al., 2012; Savitz et al., 2012a; Savitz et al., 2012b; Steenland et al., 2013; Watkins et al., 2013). These individuals were environmentally exposed to PFOA from decades of emissions from DuPont's Washington Works facility in the Mid-Ohio valley (Frisbee et al., 2009). PFOA (in the form of ammonium perfluorooctanoate-APFO) was released by the facility from 1950s to early 2000s, and exposure to PFOA for participants in the surrounding communities occurred primarily via inhalation of contaminated air and ingestion of ground water (Paustenbach et al., 2007; Shin et al., 2011a, b). For many of the epidemiological studies, individual exposure assignments were based on model predictions of yearly PFOA serum concentrations for consented C8 Health Project participants. The retrospective serum PFOA concentrations were reconstructed since 1951 using individual residential and water use histories, historical exposure concentrations, and a default elimination half-life. The predicted PFOA serum concentrations were well correlated with measured 2005–2006 serum concentrations ($r_s = 0.68$), supporting the validity of the retrospective exposure estimates (Shin et al., 2011a, b). The measured serum levels in the C8 Health population ranged over several orders of magnitude. The population mean (among the 48,998 consented participants) for serum PFOA concentration was 82.9 (with standard deviation of 240.8) ng/mL; the 2.5th to 97.5th percentile ranging from 4.3 to 530.4 ng/mL (Frisbee et al., 2009).

One of the C8 Science Panel studies analyzed the association between predicted PFOA serum concentration at the year of pregnancy and preeclampsia among the participants and found a moderate association, with an adjusted odds ratio (AOR) of 1.13 per interquartile range increase in ln serum PFOA (Savitz et al., 2012a). The interquartile range was 2.19 units on the log scale; thus there was a 13% increase in the odds of preeclampsia per

exp(2.19) ≈ 9-fold increase in serum PFOA. The C8 Science Panel concluded that a probable link exists between PFOA exposure and the occurrence of pregnancy-induced hypertension/preeclampsia (C8 Science Panel, 2011). However, the validity of this study has been questioned by one group of researchers who excluded it from a meta-analysis of PFOA exposure and fetal growth due to the retrospectively modeled exposure assignments with limited validation by measured biomarkers (Johnson and Sutton, 2014; Koustas et al., 2014). Previous studies have shown that the use of modeled pollutant concentrations and self-reported activity patterns can introduce exposure measurement error (Sarnat et al., 2010; Wu et al., 2013), as can studies that rely only on a single biomarker measurement to characterize each individual's exposure (Bartell et al., 2004; Bradman et al., 2013; Prentice et al., 2013; Tsuchiya et al., 2012) For the Savitz et al. (2012a) study, uncertainties in spatiotemporal predictions of PFOA water/air concentrations and in individual-level variables (e.g., water ingestion rates, PFOA half-life, PFOA volume of distribution) used in the dose-reconstruction and pharmacokinetic models likely resulted in some exposure measurement error, potentially affecting the validity of the epidemiological findings.

In a recent uncertainty analysis (Avanasi et al., *in press*), we used a Monte Carlo (MC) simulation methodology to characterize uncertainty in PFOA groundwater concentrations predicted from environmental fate and transport models, and determined its potential impacts on serum PFOA concentration predictions and the association between PFOA and preeclampsia (Avanasi et al., *in press*). We found that shared water PFOA concentration uncertainty, which is correlated within individuals over time and between individuals with shared water sources, substantially impacts the PFOA serum concentration predictions but only mildly impacts the epidemiological association between PFOA and preeclampsia. This appears to be due to the fact that shared uncertainty, even at a high magnitude, does not perturb the rank order of exposure among the preeclampsia cases and controls of the study. However, preliminary analyses in that study suggested that uncertainty in independent exposure parameters such as the tap water ingestion rates might have a larger impact on epidemiological associations than that in shared PFOA water concentrations (Avanasi et al., *in press*).

The objective of the present study is to determine the potential impacts of other input parameter uncertainties on the PFOA serum concentration predictions and the association between PFOA and preeclampsia. The input parameter uncertainties included in this study are realistic inter-individual variability and more subjective epistemic uncertainty in independent (non-shared) exposure factors such as water ingestion rates assigned using either self-reported (Frisbee et al., 2009) or population-level default values (U.S.EPA, 2011), PFOA elimination half-life, and PFOA volume of distribution. It has been previously identified that distinguishing these two types of uncertainty is important and commonly not addressed by researchers. Variability differs from epistemic uncertainty in a way that it represents heterogeneity in a parameter of interest, while epistemic uncertainty arises out of our lack of knowledge/understanding of the value of a parameter or its variability (Burmaster and Anderson, 1994; Cullen and Frey, 1999; Finley and Paustenbach, 1994; Morgan and Henrion, 1990). In this manuscript we obtained realistic variability distributions on these parameters from literature wherever possible. We then used Monte Carlo

simulation to determine impacts on the predicted serum concentrations and the association between PFOA and preeclampsia.

2.0 Materials and Methods

2.1. Environmental fate and transport modeling

The historical PFOA air and groundwater concentrations used in the exposure assessment were predicted by an integrated fate and transport model system (Shin et al., 2011a). This modeling system included a series of linked environmental fate and transport models to predict the yearly PFOA air and groundwater concentrations for the years 1951 – 2008, for the area that covers the communities that consented to the C8 Health Project (includes participants from the six public water districts- the City of Belpre, Little Hocking Water Association, Tuppers Plains Chester Water District, the Village of Pomeroy Water District, Lubeck Public Services District, and Mason County Public Service District). These models utilized information on historical release rates of PFOA, local meteorological and hydrogeological data, and PFOA physiochemical properties. More details on the modeling and the calibration methodology are described by Shin et al. (2011a).

2.2. Exposure-reconstruction and pharmacokinetic modeling

The predicted PFOA air and groundwater concentrations were then used in an exposure model to predict yearly PFOA exposure doses through inhalation and ingestion for each of the participants (Shin et al., 2011b). To predict yearly total exposure doses (combined inhalation and ingestion doses) for each of the participants, this exposure model utilized information on: self-reported participant demographics such as age, gender, body weight; residential/work histories; standard (recommended mean) inhalation rates (U.S. EPA, 2009); standard (self-reported, if available) water ingestion rates; and information on the historical pipe distribution systems of each of the six public water districts. Self-reported water ingestion rates (number of cups per day) were available for approximately 50% of the C8 Health Project participants (Shin et al., 2011b) and were used when available. This is because the original C8 Health Project questionnaire did not include questions about participant water ingestion rate. The participants were originally asked to assess water intake by filling a 'water usage at residence' questionnaire for each address the participant resided in. The questions asked for the main source of water (public/private/bottled water/other) used for drinking, cooking, showering/bathing in each residential address by each participant. Tap water ingestion rates were later obtained between 2010 and 2011 when participants were re-contacted by telephone for follow-up studies conducted by the C8 Science Panel (Winquist et al., 2013). Tap water ingestion rates were self-reported using the following ranges: 0–1, 1–2, 3–4, 5–6, 7–8, > 8 cups per day. The midpoint of each range was used in the original exposure model, yielding a mean of 5.44 cups per day. For the self-reported participants the model also assumed a constant water consumption rate for each individual over the entire exposure period.

A one-compartment pharmacokinetic model (PK) was used to predict yearly PFOA serum concentrations in the study participants, based on their individual yearly total PFOA dose estimates and assuming 100% absorption by both exposure routes (Shin et al., 2011b). The

total PFOA serum concentration was computed as the sum of PFOA serum concentration due to emissions from the Washington Works plant and PFOA serum concentration due to background exposure from other exposure sources (Shin et al., 2011b). By default, the model used age- and gender-specific PFOA volume of distribution based on previously recommended volume of distribution per unit body weight derived from a monkey model by Butenhoff et al. (2004). The PFOA elimination half-life in the PK model was fixed at 3.5 years for all participants based on follow-up of retired workers by Olsen et al. (2007).

2.3. Epidemiological analysis

The Savitz et al. (2012a) epidemiological analysis of self-reported preeclampsia in association with PFOA exposure utilized the predicted PFOA serum concentrations at the year of birth for 10,189 pregnancies occurring between 1990 and 2006 in the C8 Health Project population. Among these pregnancies, 730 occurrences of preeclampsia were reported and the analysis adjusted for confounding by parity, maternal age, education level and smoking status. The original reported adjusted odds ratio (AOR) for the occurrence of preeclampsia was 1.13 (1.00–1.28) per interquartile range (IQR) of PFOA serum concentration in log (natural logarithm) scale (nanograms per milliliter-ng/mL). When we restrict (modify) the analysis to 10,149 pregnancies among 6134 mothers (725 preeclampsia cases) by removing those mothers with previous work histories in the Washington Works facility, thereby focusing only on the contribution from non-occupational sources of exposure, the modified analysis yielded an AOR of 1.12 (1.00–1.26). We received approval from the Institutional Review Board at the University of California, Irvine (HS#2013-9421) to conduct our analyses.

2.4 Methodology: Monte Carlo simulation I

In the first MC simulation, we study the impact of variability in the individual-level input parameters including the standard tap water ingestion rates or the self-reported water ingestion rates, the PFOA half-life, and the PFOA volume of distribution. We used the outputs from the fate and transport models (the PFOA air and water concentrations) and ran multiple iterations of the exposure and the pharmacokinetic models. For each iteration, the individual-level independent exposure parameters were varied by randomly sampling from variability distributions developed based on our literature survey described below.

2.4.1. Variability in standard/self-reported water ingestion rates—In the original exposure model (Shin et al., 2011b), for the participants who did not provide a self-reported tap water ingestion rate (~30 % of the Savitz et al., 2012a study participants), default water ingestion rates based on the EPA Exposure Factors Handbook (U.S. EPA, 2011) were used. These are mean water ingestion rates for different age categories (for the U.S. population) and include estimates of water ingested directly as a beverage and water used in preparing various food/beverages. Percentile estimates of these water ingestion rate distributions are also available for each age category in the Exposure Factors Handbook. In order to account for inter-individual variability in the MC analysis, we assumed the water ingestion rate to be log-normally (LN) distributed and calculated the log-normal parameters (the log mean, μ_a and the log standard deviation, σ_a) based on the mean and 95th percentile values for each age category (Table 1). Log-normal distributions are non-negative which is very useful for

describing variability/uncertainty in biological parameters; they have been previously used in a number of Monte Carlo applications for that reason (Limpert et al., 2001, Morgan et al., 1990). Further, Roseberry and Burmaster (1992) found that the use of log-normal distributions to describe water ingestion rates provided robust results for all age-groups and highly recommend their use.

In the MC analysis, for each iteration, a percentile of the water ingestion rate for any specific individual was randomly sampled from a uniform distribution and used throughout the individual's lifetime to compute the water ingestion rate from the age-specific log-normal variability distributions as shown. We chose to hold the percentile of water ingestion rate constant instead of the absolute water ingestion rate in order to accommodate the fact that an individual's water consumption changes during the aging process. For example, using the rates based on the EPA Exposure Factors Handbook, the mean water ingestion rate for an individual whose age is 21 years is 32 % more than that if the age is 18, 19 or 20 years.

$$p\tilde{\ }U\left(0,1\right)$$
$$IR_i = F^{-1}\left(p\right) \quad [1]$$

Where,

p is the sampled percentile for individual i,

F^{-1} is the inverse distribution function (or "quantile function") for LN (μ_a, σ_a),

IR_i is the computed water ingestion rate for individual i,

μ_a is the log mean specific to individual i's age category and,

σ_a is the log standard deviation specific to individual i's age category.

Holding the percentile of the water ingestion rate constant throughout the lifetime for any iteration simulated the effect of consistently assigned high or low water consumption on PFOA exposures. For example, an individual assigned to the 92nd percentile of water consumption would remain at that percentile within each age group, throughout his or her lifetime, for that MC iteration.

For the participants who had self-reported tap water ingestion rates (~ 70 % of the participants), they reported the number of cups of water consumed in a day in one of the six categories: 0–1, 1–2, 3–4, 5–6, 7–8, > 8 cups. The original exposure model had used the mean of each category as the water ingestion rate, except for the > 8 cups category, which used 9 cups. These rates were held constant throughout the individual's lifetime.

In the MC analysis, to account for the variability within each category, for any individual, we randomly sampled a value of a water ingestion rate based on a uniform distribution shown below, instead of the median value (as used in the original model) and used it throughout all years of the individual's lifetime.

$$m_i \sim U \ (L_i, U_i) \quad \text{[2]}$$

Where,

m_i is the sampled self-reported water ingestion rate (L/day) for individual i,

L_i and U_i are the lower and upper bounds of the uniform distribution.

For example, $m_i \sim U$ (1.183, 1.4196) for an individual who reported 5–6 cups per day of tap water ingestion, assuming the volume of one cup is 0.2367 L.

2.4.2. Variability in the PFOA elimination half-life—In the original model, the PFOA half-life used in the pharmacokinetic model was fixed at 3.5 years for all individuals based on a study by Olsen et al. (2007) which followed 26 retired fluorochemical production workers, over five years. In order to account for inter-individual variability in the PFOA half-life, we assumed it to be log-normally distributed and calculated the log-normal parameters (μ_h = 1.247 and σ_h = 0.399) based on the reported arithmetic mean and standard deviation of the PFOA half-life among the 26 subjects in the Olsen et al. (2007) study. Log-normal distributions have been previously used to describe inter-individual variability/uncertainty in pharmacokinetic parameters such as elimination half-life, because they are non-negative (Clewell et al., 1999; Limpert et al., 2001).

In the MC analysis, for each iteration, the PFOA half-life for any specific individual was randomly sampled from the log-normal variability distributions as shown below and used throughout each year of the individual's lifetime.

$$T_{\frac{1}{2}i} LN \ (\mu_h, \sigma_h) \quad \text{[3]}$$

Where,

$T_{\frac{1}{2}i}$ is the PFOA half-life for individual i,

μ_h is the log mean of the PFOA half-life and,

σ_h is the log standard deviation of the PFOA half-life.

2.4.3. Variability in the PFOA volume of distribution—The PFOA volume of distribution per weight used in the original model (female = 0.198 L/Kg) was based on a study in cynomolgus monkeys by Butenhoff et al. (2004). This volume of distribution was held constant for all the participants in the original model (Shin et al., 2011b). To account for inter-individual variability in the volume of distribution, we assumed it to be log-normally distributed and calculated the parameters (μ_{vd} = 5.2453 and σ_{vd} = 0.3540) based on the reported arithmetic mean and standard deviation of the PFOA volume of distribution among the 3 female monkey subjects in the study.

In the MC analysis, for each iteration, the PFOA volume of distribution for any specific individual was randomly sampled from the log-normal variability distributions shown below and used throughout the individual's entire lifetime.

$$VD_i \sim LN\left(\mu_{vd}, \sigma_{vd}\right) \quad [4]$$

Where,

VD_i is the PFOA volume of distribution for individual i,

μ_{vd} is the log mean of the PFOA volume of distribution and,

σ_{vd} is the log standard deviation of the PFOA volume of distribution.

2.4.4 Monte Carlo simulations II: Epistemic uncertainty in self-reported water ingestion rates, PFOA half-life and PFOA volume of distribution—In addition to studying the impact of variability in the self-reported water ingestion rates as described in section 2.4., in a separate MC simulation (II), we studied the impact of epistemic uncertainty in the three individual-level parameters. For uncertainty regarding the validity of each self-reported water ingestion rate, we used a log-normal distribution shown below.

$$IR_i \sim LN\left(\mu_i, \sigma_i\right) \quad [5]$$

$$\mu_i = \log\left(m_i\right) \quad [6]$$

Where,

IR_i is the sampled water ingestion rate for individual i,

μ_i is the individual specific log mean and,

σ_i is the log standard deviation reflecting potential inaccuracy in the self-reported water ingestion rates.

This was a two-level analysis evaluating both uncertainty in the self-reported water ingestion rate category and, as described in section 2.4.1, variability within each selected category. The median of the log-normal distribution was sampled randomly from a uniform distribution as described in equation 2. The standard deviation ($\sigma_i = 0.446$) was computed from the confidence interval of the mean difference between a questionnaire (self-reports) and 3-day diary based on water ingestion rates studied by Shimokura et al. (1998). This standard deviation represents the uncertainty in self-reported water ingestion rates as a surrogate for true rates.

We also included the epistemic uncertainty in PFOA half-life and PFOA volume of distribution in the MC simulation II. The PFOA half-life used in the study was based on a study (Olsen et al., 2007) of mainly older (ranged between 55–75 years) males (24 males

and 2 females) and may not be representative of the participants in the Savitz et al., 2012a study, which were mainly women of child-bearing age (ranged between 14–45 years). Therefore, we included epistemic uncertainty by increasing the log standard deviation of the log-normal distribution in equation 3 to 0.798 (2-times the original value in the half-life estimate used in the MC simulation I). Because only three monkeys were used in the Butenhoff et al. (2004) study to calculate the PFOA volume of distribution for females, and monkeys were used as a surrogate for humans in our analysis (Shin et al., 2011b), there could be additional uncertainty in the parameter. To address these epistemic sources of uncertainty in the volume of distribution, we multiplied the log standard deviation by a factor of 5 (i.e., from 0.35 to 1.77). Although this multiplier is subjective and somewhat arbitrary due to the absence of experimental data in humans, we compared the resulting serum PFOA concentration predictions with the original predictions and the 2005–2006 measured serum PFOA concentrations to judge whether the results were reasonable. In MC simulation II, the same variability distributions from MC simulation I and these three epistemic uncertainties (self-reported water ingestion rates, PFOA half-life and PFOA volume of distribution) were all included in the model.

The MC simulations (300 iterations each) were run using MATLAB (The Mathworks Inc., Natick, MA, 2000) and R (http://www.r-project.org/) was used to reanalyze the epidemiological association between the estimated serum PFOA concentrations and preeclampsia. For each MC iteration of the exposure model, the serum PFOA concentrations were estimated and multiple logistic regression was used to compute the AOR per IQR of the estimated serum concentrations. The rank correlation between the simulated and original serum PFOA concentration estimates for the 10,149 participants between the years 1990 and 2006 was calculated for both MC simulations. The plot of the mean and 95% probability interval over 300 iterations of the rank correlation compares the yearly rank exposure of the participants (among all 300 iterations) with the original model predicted rank exposure for each MC simulation. The contribution of inter-individual variability/epistemic uncertainty to the total uncertainty in the epidemiological association (including participant sampling uncertainty) was calculated for each MC simulation using the law of total variance as described in our previous study (Avanasi et al., *in press*). Briefly, if b is the log odds parameter estimate and X is a collection of individual exposure estimates, the total variance (as a measure of uncertainty) in b is given by $var(b) = E(var(b|X)) + var(E(b|X))$, where var is the variance and E corresponds to the expected value. The first and the second terms in the summation correspond to the participant sampling variability and the inter-individual variability/epistemic uncertainty, respectively. The formula $var(E(b|X))/var(b)$ gives the relative contribution of the inter-individual variability/epistemic uncertainty to the total uncertainty.

3.0 Results

3.1. Impact on predicted serum concentrations

We calculated the mean, median, 25th and 75th percentile predicted serum concentrations (ng/mL) of each iteration for 10,149 participants and their mean and 95% probability interval among 300 iterations for each MC simulation (Table 2), along with the

corresponding values for the modified original data (including only non-occupationally exposed participants as described in section 2.3.). The mean serum PFOA concentration among the 300 iterations of the MC simulation II was nearly 5-times that of the modified original serum PFOA concentrations, but the median was similar. Across the 300 iterations of the two MC simulations, the IQR of the log serum PFOA concentrations varied minimally with the 95% probability interval of the IQR ranging between 1.32 and 2.29 for MC simulation I and between 0.96 and 2.86 for MC simulation II. The mean and the 95% probability interval for the IQR for the two simulations (including the original IQR) are given in Table 3.

The rank correlation between the simulated and original serum PFOA concentration estimates for the 10,149 participants between the years 1990 and 2006 was calculated. The plot of the mean and 95% probability interval over 300 iterations of the rank correlation for the Monte Carlo simulation I is presented in Figure 1. A similar plot for MC simulation II is shown in Figure 2. For the MC simulation II, as seen in Figure II, we find that the estimated and the original predicted serum concentrations are well correlated, around 0.92 (in the year 1990), but over time, this reduces (due to the addition of variability/uncertainty) and the lowest mean rank correlation was 0.77 (for the years 2003for the years 2004 and 2005). This represents the simulation with the maximum impact of variability/epistemic uncertainty on the rank exposure, with the other simulation (the Monte Carlo simulation I) producing higher rank correlations (as shown in Figure 1) with the original serum PFOA concentrations.

3.2. Impact on the epidemiological association

The mean with 95% probability interval of AOR for each simulation was calculated and is presented in Table 4. The percent contribution of the variability/epistemic uncertainty to the overall uncertainty is also presented in Table 4. We find that the addition of epistemic uncertainty to variability increases the contribution of exposure uncertainty to the total uncertainty (including sampling variability) in the epidemiological association by 21% (comparing the contributions of exposure uncertainty between MC simulations I and II). The AOR of preeclampsia occurrence is also reduced from 1.12 to 1.09 (25 % reduction), indicating the impact of variability/epistemic uncertainty on the AOR. As previously reported in the literature (Armstrong, 1998), the increasing variability/epistemic uncertainty in independent exposure factors resulted in increasing bias of the AOR towards the null and increased the contribution of exposure uncertainty to overall uncertainty. Figure 3 is a boxplot of the point estimates of the log odds of preeclampsia occurrence from the 300 iterations for MC simulation I and MC simulation II. This plot shows the variation in preeclampsia risk attributed to PFOA across the iterations, compared with the log odds of preeclampsia occurrence in the original modified analysis (dotted horizontal line in Figure 3).

4.0 Discussion

From the results of MC simulation I, we found that realistic inter-individual variability (determined based on our literature review) in independent exposure parameters such as the water ingestion rates, the PFOA half-life, and the PFOA volume of distribution impacted the

absolute serum PFOA concentration predictions (Table 2). The rank order of estimated serum PFOA concentrations was only mildly impacted (the lowest mean rank correlation between the estimated serum concentrations in our study and the original predicted serum concentrations was 0.95 for the years 2003 the lowest mean rank correlation between the estimated serum concentrations in our study and the original predicted serum concentrations was 0.95 for the years 2004 and 2005 as seen in Figure 1), without causing any change to the mean AOR of preeclampsia occurrence among the participants (within rounding error). The reduction in rank exposure over the years can be explained by the fact that the participants are cumulatively exposed to PFOA over the years and that PFOA is partly retained in the serum over time. Hence, any impact of uncertainty on the serum PFOA predictions will also be cumulative over the years and thus magnified over multiple years of exposure. The overall contribution of exposure uncertainty to the total uncertainty (including participant sampling variance) for MC simulation I was low, around 7%.

In contrast, in MC simulation II, when epistemic uncertainty is added along with the inter-individual variability in the same exposure parameters, both the absolute serum PFOA concentration predictions (Table 2) and the rank order of estimated serum PFOA concentrations were moderately impacted. The mean serum PFOA concentrations increased nearly 5-fold compared to the modified original predictions, suggesting that the addition of epistemic uncertainty has a significant impact on the serum concentration predictions. However, the use of lognormal distributions with large standard deviations can result in some unrealistic values being sampled from the long tails of the distributions and inflation of the mean. The median is not impacted by the long tails, as seen in Table 2. The lowest mean rank correlation between the estimated and the original predicted serum concentrations was 0.77 for the years 2003, 2004 and 2005 as seen in Figure 2. The mean AOR was reduced by 25% (from 1.12 to 1.09). The total contribution of exposure uncertainty to the total uncertainty also increased to nearly 33% (compared to the original model). These results support previous literature that non-differential variability and epistemic uncertainty in independent parameters of individual-level exposure models and pharmacokinetic models can lead to bias in the effect estimates of environmental epidemiological studies towards the null and reduce the power as well as the precision of these studies (Thomas et al., 1993; Armstrong, 1998).

Our recent uncertainty analysis study in the PFOA exposure assessment modeling system on the C8 Heath Project/C8 Science panel study population (Avanasi et al., *in press*) focused on correlated exposure uncertainty in the predicted public water district PFOA water concentrations from environmental fate and transport models (correlated within each participant over the years and between participants with a shared drinking water source). Despite larger uncertainties in the predicted PFOA water concentrations, the impact of correlated exposure uncertainty on the epidemiological association is negligible compared to the impact of variability/epistemic uncertainty in independent exposure parameters seen here. Our previous study (Avanasi et al., *in press*) showed that shared uncertainty substantially impacted the absolute PFOA serum concentration predictions, but only mildly impacted the rank order of estimated serum PFOA concentrations (the lowest mean rank correlation between the estimated and the original predicted serum concentrations was 0.89)

and did not impact the mean AOR between PFOA serum concentration predictions and preeclampsia. These results seemed counter-intuitive, considering that large changes to exposure assignments might be expected to cause large changes in the epidemiological results. Together, these two studies suggest that in the PFOA exposure assessment modeling system, independent sources of error are more likely to change the rank order of exposure of participants and in turn impact the AOR of association than correlated uncertainty arising out of shared exposure sources. This may be due to relatively large differences (spanning several orders of magnitude) in PFOA concentrations between water districts for the C8 Science Panel study area (Shin et al., 2011a), so that the rank order of exposure among participants is mostly preserved despite large changes to the PFOA groundwater concentration for each water district.

Notably, a prospective analysis of the same study population reported an AOR of 1.27 (95% CI: 1.05, 1.55) for PFOA and pregnancy-induced hypertension using measured serum concentrations and 2005–2010 birth records instead of modeled exposure assignments and self-reported health outcomes (Darrow et al., 2013). Preeclampsia is a type of pregnancy-induced hypertension that also includes proteinuria; the presence or absence of preeclampsia was not recorded on the birth certificates and pregnancy-induced hypertension was not included on the C8 Health Project questionnaire (C8 Science Panel, 2011). Together, the retrospective and prospective studies show consistent associations between PFOA and pregnancy-induced hypertension (with or without proteinuria) in this study population, with low sensitivity to exposure uncertainty.

4.1. Limitations

Although the probability distributions for each of the exposure parameters in MC simulation I were derived from empirical evidence/self-reported values (percentile estimates of water ingestion rate distributions for standard water ingestion rates based on the self-reported ranges of the number of cups of water consumed in a day, the PFOA half-life based on the Olsen et al., 2007 study, and the PFOA volume of distribution based on the Butenhoff et al., 2004 study), the choice of probability distributions reflecting epistemic uncertainty in MC simulation II for the PFOA half-life (two times the log standard deviation of the variability distribution) and PFOA volume of distribution (five times the log standard deviation of the variability distribution) were based on subjective judgments, with less clear interpretability. For example, the 95% probability interval (PI) of the half-life epistemic uncertainty distribution is 1.59–7.61 years and that of the volume of distribution is 0.006–6.09 L/Kg. As a sensitivity analysis, we also looked at an uncertainty distribution for the volume of distribution with a log standard deviation which is 10 times that used in the variability analysis. The 95% probability interval for this analysis is between 0.0002–195.6 L/Kg, resulting in a 50% decrease (from 1.12 to 1.06) in the mean AOR, but the analysis produced implausible serum PFOA concentration predictions with a mean of 23604.6 (95% PI: 4643.9, 151368.5) ng/mL, orders of magnitude larger than 2005–2006 measured serum PFOA concentrations among consented C8 Health Project participants (n=48,998) for which the 95% PI was 4.3, 530.4 ng/mL. An epistemic uncertainty distribution for the volume of distribution with five times the standard deviation of the variability distribution is more

plausible considering the resulting mean (95% PI) PFOA serum concentrations among the 300 iterations was 267.6 (63.5, 578.8).

It is to be noted that there is a data gap with respect to PFOA volume of distribution in humans, with the Butenhoff et al. (2004) monkey study being the only source of this information. Other existing PK models (Andersen et al., 2006; Loccisano et al., 2012; Loccisano et al., 2013; Tan et al., 2008) of PFOA in rats, monkeys and humans have all used the volume of distribution results of the Butnhoff et al. (2004) study to examine the kinetics of PFOA. Also, it is to be noted that the sample size of 3 female monkeys adds to the limitations of this study. However, choosing the appropriate epistemic uncertainty distributions (for half-life and volume of distribution) for MC simulation II is inherently subjective, as epistemic uncertainty refers to unmeasured attributes.

Because our MC simulation evaluates the impact of adding parameter variability/uncertainty (adds new measurement error) to the exposure estimates, rather than "correcting for" or "removing" measurement error, the AORs reported in Tables 3 and 4 indicate the expected direction of bias in the epidemiologic results due to additional measurement error rather than the AORs that would result from removing the measurement error. So, had there been no measurement error in the original exposure estimates, the AOR for serum PFOA and preeclampsia in this study population would likely have been higher than 1.12, with a wider confidence interval than originally estimated, because the original AORs were likely biased towards the null. Nonetheless, our simulations indicate that the epidemiological association between PFOA serum concentration predictions and preeclampsia in this study population is not very sensitive to variability/uncertainty in the retrospective exposure assignments.

Another limitation of this analysis is that we did not consider additional confounders or investigate impacts of uncertainty in the confounders used in the Savitz et al., 2012 study. A recent analysis by Watkins et al. (2013) investigated the association between measured/ model-predicted PFOA serum levels and the estimated glomerular filtration rate (eGFR), which could be a potential confounder for epidemiological studies that rely on measured serum PFOA. However, they noted that eGFR was not used in the models predicting serum PFOA concentrations for the C8 Health Project population, and therefore cannot be a confounder of epidemiological studies that rely on modeled (rather than measured) serum PFOA. Measurement error in the confounders could also have an impact on the AOR. Although parity, maternal age, and education level are likely to have been reported accurately by most participants, smoking is often subject to under-reporting, particularly among pregnant women (Shipton et al., 2009; Ford et al., 1997).

4.2. Conclusions

In the uncertainty analysis presented here, we studied the impact of realistic inter-individual variability/epistemic uncertainty in independent exposure parameters including the standard and self-reported water ingestions rates, PFOA half-life, and PFOA volume of distribution on the predicted serum PFOA concentrations and the association between PFOA and preeclampsia in the participants of the C8 Health Project. Variability and epistemic uncertainty in these independent parameters change the rank exposure among the study participants enough to cause a 25% bias in the AOR towards the null. This result is in line

with the previous literature which suggested that independent exposure measurement error can bias the effect estimate of an epidemiology study and suggests that the true AOR of association between PFOA and preeclampsia in the C8 Health Project/C8 Science Panel study population might be higher than originally reported with a wider confidence interval considering the effects of exposure uncertainty. We found it useful to study the impacts of these two types of exposure uncertainty (independent vs. correlated) separately. We think that future epidemiology studies with complex exposure scenarios and multiple sources of variability/uncertainty can separate out the two kinds of uncertainty and study them separately to better understand their potential impacts and to what extent, if any, they threaten the validity of epidemiological studies.

Acknowledgments

Abbreviations

C8	PFOA Perfluorooctanoate
MC	Monte Carlo
AOR	Adjusted Odds Ratio
LN	Log-Normal
PK	Pharmacokinetic
IQR	Inter Quartile Range
PI	Probability Interval
eGFR	Estimated Glomerular Filtration Rate

References

Andersen ME, Clewell HJ, Tan YM, Butenhoff JL, Olsen GW. Pharmacokinetic modeling of saturable, renal resorption of perfluoroalkylacids in monkeys-Probing the determinants of long plasma half-lives. Toxicology. 2006; 227:156–164.10.1016/j.tox.2006.08.004 [PubMed: 16978759]

Armstrong BG. Effect of measurement error on epidemiological studies of environmental and occupational exposures. Occup Environ Med. 1998; 55:651–6. [PubMed: 9930084]

Avanasi R, Shin HM, Vieira VM, Savitz DA, Bartell SM. Impact of Exposure Uncertainty on the Association between Perfluorooctanoate and Preeclampsia in the C8 Health Project Population. Environ Health Perspect. in press. 10.1289/ehp.1409044

Barry V, Winquist A, Steenland K. Perfluorooctanoic acid (PFOA) exposures and incident cancers among adults living near a chemical plant. Environ Health Perspect. 2013; 121:1313–1318.10.1289/ehp.1306615 [PubMed: 24007715]

Bartell SM, Griffith WC, Faustman EM. Temporal error in biomarker-based mean exposure estimates for individuals. J Expo Anal Environ Epidemiol. 2004; 14:173–9.10.1038/sj.jea.7500311 [PubMed: 15014548]

Bradman A, Kogut K, Eisen EA, Jewell NP, Quirós-Alcalá L, Castorina R, et al. Variability of organophosphorous pesticide metabolite levels in spot and 24-hr urine samples collected from young children during 1 week. Environ Health Perspect. 2013; 121(1):118–124. [PubMed: 23052012]

Burmaster DE, Anderson PD. Principles of good practice for the use of Monte Carlo techniques in human health and ecological risk assessments. Risk Anal. 1994; 14:477–81. [PubMed: 7972955]

Butenhoff JL, Kennedy GL, Hinderliter PM, Lieder PH, Jung R, Hansen KJ, Gorma GS, Nokers PE, Thomford PJ. Pharmacokinetics of perfluorooctanoate in cynomolgus monkeys. Toxicol Sci. 2004; 82:394–406.10.1093/toxsci/kfh302 [PubMed: 15470233]

C8 Science Panel. Probable link evaluation of pregnancy induced hypertension and preeclampsia 1–6. 2011. Available: http://www.c8sciencepanel.org/prob_link.html. [Accessed 1 February]

Clewell HJ, Gearhart JJM, Gentry PR, Covington TR, Vanlandingham CB, Crump KS, Shipp AM. Evaluation of the Uncertainty in an Oral Reference Dose for Methylmercury Due to Interindividual Variability in Pharmacokinetics. Risk Anal. 1999; 19:547–58. [PubMed: 10765421]

Cullen, AC.; Frey, HC. Probabilistic techniques in exposure assessment: a handbook for dealing with variability and uncertainty in models and inputs. Springer Science and Business Media; 1999. p. 1-14.Chapter 1

Darrow LA, Stein CR, Steenland K. Serum Perfluorooctanoic Acid and Perfluorooctane Sulfonate Concentrations in Relation to Birth Outcomes in the Mid-Ohio Valley, 2005–2010. Environ Health Perspect. 2013:2005–2010.

Finley B, Paustenbach D. The benefits of probabilistic exposure assessment: three case studies involving contaminated air, water, and soil. Risk Anal. 1994; 14:53–73. [PubMed: 8146404]

Ford RP, Tappin DM, Schluter PJ, Wild CJ. Smoking during pregnancy: how reliable are maternal self reports in New Zealand? J Epidemiol Community Health. 1997; 51:1–2.

Frisbee SJ, Brooks AP, Maher A, Flensborg P, Arnold S, Fletcher T, Steenland K, Shankar A, Knox SS, Pollard C, Halverson JA, Vieira VM, Jin C, Leyden KM, Ducatman AM. The C8 Health Project: Design, Methods, and Participants. Environ Health Perspect. 2009; 117:1873–1882.10.1289/ehp.0800379 [PubMed: 20049206]

Gallo V, Leonardi G, Genser B, Lopez-Espinosa MJ, Frisbee SJ, Karlsson L, Ducatman AM, Fletcher T. Serum perfluorooctanoate (PFOA) and perfluorooctane sulfonate (PFOS) concentrations and liver function biomarkers in a population with elevated PFOA exposure. Environ Health Perspect. 2012; 120:655–660.10.1289/ehp.1104436 [PubMed: 22289616]

Johnson P, Sutton P. The Navigation Guide—Evidence-Based Medicine Meets Environmental Health: Systematic Review of Human Evidence for PFOA Effects on Fetal Growth. Env Heal. 2014; 122:1040–1051.10.1289/ehp.1307893

Koustas E, Lam J, Sutton P, Johnson PI, Atchley DS, Sen S, Robinson KA, Axelrad DA, Woodruff TJ. The Navigation Guide - evidence-based medicine meets environmental health: systematic review of nonhuman evidence for PFOA effects on fetal growth. Environ Health Perspect. 2014; 122:1015–27.10.1289/ehp.1307177 [PubMed: 24968374]

Limpert E, Stahel WA, Abbt M. Log-normal distributions across the sciences: keys and clues. Biosciences. 2001; 51(5):341.

Loccisano AE, Campbell JL, Butenhoff JL, Andersen ME, Clewell HJ. Evaluation of placental and lactational pharmacokinetics of PFOA and PFOS in the pregnant, lactating, fetal and neonatal rat using a physiologically based pharmacokinetic model. Reprod Toxicol. 2012; 33:468–90.10.1016/j.reprotox.2011.07.003 [PubMed: 21872655]

Loccisano AE, Longnecker MP, Campbell JL, Andersen ME, Clewell HJ. Development of PBPK models for PFOA and PFOS for human pregnancy and lactation life stages. J Toxicol Environ Health A. 2013; 76:25–57.10.1080/15287394.2012.722523 [PubMed: 23151209]

Lopez-Espinosa M, Mondal D, Armstrong B, Bloom MS, Fletcher T. Thyroid Function and Perfluoroalkyl Acids in Children Living Near a Chemical Plant. Environ Health Perspect. 2012; 120:1036–1041.10.1289/ehp.1104370 [PubMed: 22453676]

Morgan, G.; Henrion, M.; Small, M. Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. New York, NY: Cambridge University Press; 1990.

Olsen GW, Burris JM, Ehresman DJ, Froehlich JW, Seacat AM, Butenhoff JL, Zobel LR. Half-life of serum elimination of perfluorooctanesulfonate, perfluorohexanesulfonate, and perfluorooctanoate in retired fluorochemical production workers. Environ Health Perspect. 2007; 115:1298–305.10.1289/ehp.10009 [PubMed: 17805419]

Paustenbach DJ, Panko JM, Scott PK, Unice KM. A methodology for estimating human exposure to perfluorooctanoic acid (PFOA): a retrospective exposure assessment of a community (1951–2003). J Toxicol Environ Health A. 2007; 70:28–57.10.1080/15287390600748815 [PubMed: 17162497]

Prentice RL, Tinker LF, Huang Y, Neuhouser ML. Calibration of self-reported dietary measures using biomarkers: an approach to enhancing nutritional epidemiology reliability. Curr Atheroscler Rep. 2013; 15:353.10.1007/s11883-013-0353-5 [PubMed: 23881548]

Roseberry AM, Burmaster DE. Lognormal distributions for water intake by children and adults. Risk Anal. 1992; 12:99–104. [PubMed: 1574621]

Sarnat SE, Klein M, Sarnat JA, Flanders WD, Waller LA, Mulholland JA, Russell AG, Tolbert PE. An examination of exposure measurement error from air pollutant spatial variability in time-series studies. J Expo Sci Environ Epidemiol. 2010; 20:135–46.10.1038/jes.2009.10 [PubMed: 19277071]

Savitz DA, Stein CR, Bartell SM, Elston B, Gong J, Shin HM, Wellenius GA. Perfluorooctanoic acid exposure and pregnancy outcome in a highly exposed community. Epidemiology. 2012; 23:386–92.10.1097/EDE.0b013e31824cb93b [PubMed: 22370857]

Savitz DA, Stein CR, Elston B, Wellenius GA, Bartell SM, Shin H, Vieira VM, Fletcher T. Children's Health Relationship of Perfluorooctanoic Acid Exposure to Pregnancy Outcome Based on Birth Records in the Mid-Ohio Valley. Environ Health Perspect. 2012; 120:1201–1207. [PubMed: 22450153]

Shimokura GH, Savitz DA, Symanski E, Symansk E. Assessment of Water Use for Estimating Exposure to Tap Water Articles' Contaminants Environmental. Environ Heal. 1998; 106:55–59.

Shin H-M, Vieira VM, Ryan PB, Detwiler R, Sanders B, Steenland K, Bartell SM. Environmental fate and transport modeling for perfluorooctanoic acid emitted from the Washington Works Facility in West Virginia. Environ Sci Technol. 2011a; 45:1435–42.10.1021/es102769t [PubMed: 21226527]

Shin H-M, Vieira VM, Ryan PB, Steenland K, Bartell SM. Retrospective exposure estimation and predicted versus observed serum perfluorooctanoic acid concentrations for participants in the C8 Health Project. Environ Health Perspect. 2011b; 119:1760–5.10.1289/ehp.1103729 [PubMed: 21813367]

Shipton D, Tappin DM, Vadiveloo T, Crossley JA, Aitken DA, Chalmers J. Reliability of self reported smoking status by pregnant women for estimating smoking prevalence: a retrospective, cross sectional study. BMJ. 2009; 339:b4347.10.1136/bmj.b4347 [PubMed: 19875845]

Steenland K, Zhao L, Winquist A, Parks C. Ulcerative colitis and perfluorooctanoic acid (PFOA) in a highly exposed population of community residents and workers in the Mid-Ohio Valley. Environ Health Perspect. 2013; 121:900–905.10.1289/ehp.1206449 [PubMed: 23735465]

Tan YM, Clewell HJ, Andersen ME. Time dependencies in perfluorooctylacids disposition in rat and monkeys: A kinetic analysis. Toxicol Lett. 2008; 177:38–47.10.1016/j.toxlet.2007.12.007 [PubMed: 18242015]

Tsuchiya A, Duff R, Stern AH, White JW, Krogstad F, Burbacher TM, Faustman EM, Mariën K. Single blood-Hg samples can result in exposure misclassification: temporal monitoring within the Japanese community (United States). Environ Health. 2012; 11:37.10.1186/1476-069X-11-37 [PubMed: 22676231]

U.S. EPA (Environmental Protection Agency). Exposure Factors Handbook: 2011 Edition. 2011. Available: http://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=236252 [accessed 1 February 2015]

Watkins DJ, Josson J, Elston B, Bartell SM, Shin H, Vieira VM, Savitz DA, Fletcher T, Wellenius GA. Environmental health perspectives. Environ Health Perspect. 2013

Winquist A, Lally C, Shin HM, Steenland K. Design, methods, and population for a study of PFOA health effects among highly exposed Mid-Ohio Valley community residents and workers. Environ Health Perspect. 2013; 121:893–899.10.1289/ehp.1206450 [PubMed: 23735518]

Wu J, Jiang C, Jaimes G, Bartell S, Dang A, Baker D, Delfino RJ. Travel patterns during pregnancy: comparison between Global Positioning System (GPS) tracking and questionnaire data. Environ Health. 2013; 12:86.10.1186/1476-069X-12-86 [PubMed: 24107241]

Highlights

- PFOA exposure is linked to preeclampsia by a C8 Science Panel epidemiology study

- We study the impact of variability/uncertainty in independent exposure parameters

- We analyze the impact of water ingestion rates and PFOA pharmacokinetic parameters

- Our MC simulation shows a 25% bias towards the null in the AOR with preeclampsia

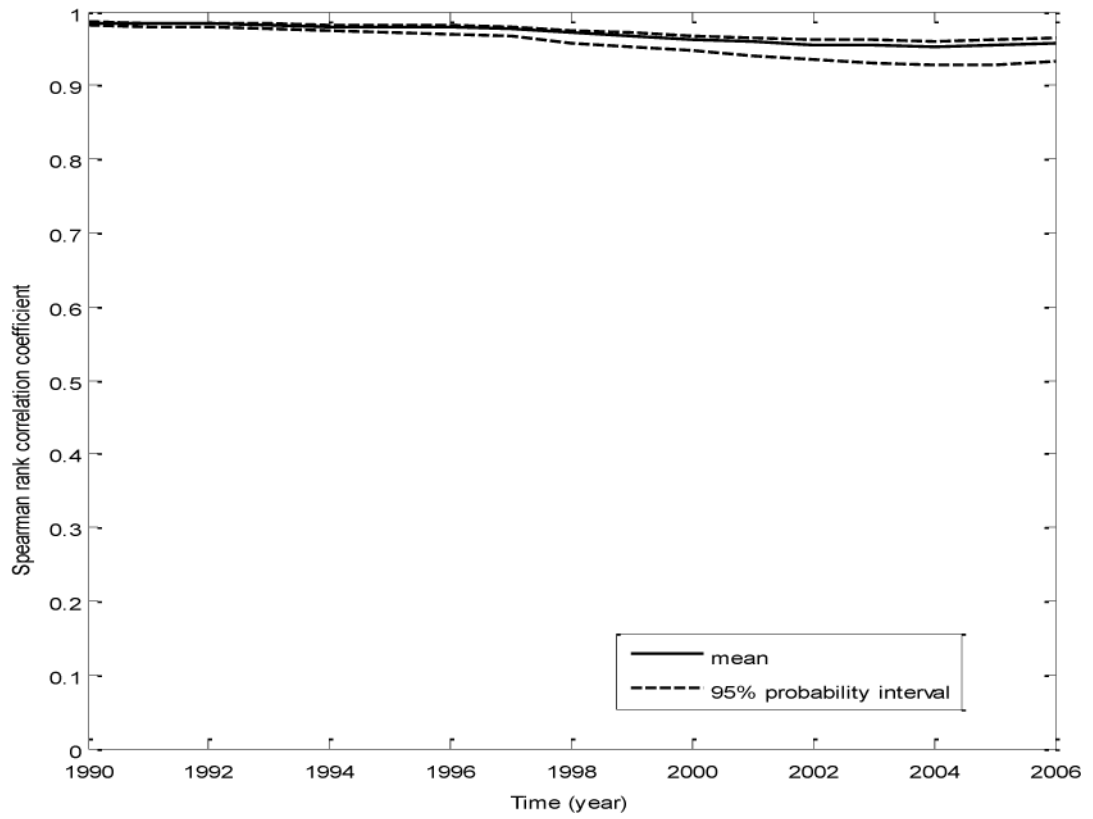- The true AOR in this population might be higher than the previously reported AOR

Figure 1.
A plot of the mean and the 95% probability interval of the correlation coefficient between the estimated serum concentrations for each Monte Carlo iterate and the original estimated serum concentrations, for all the participants, over time- MC simulation I (analysis of the impact of variability in independent input parameters)
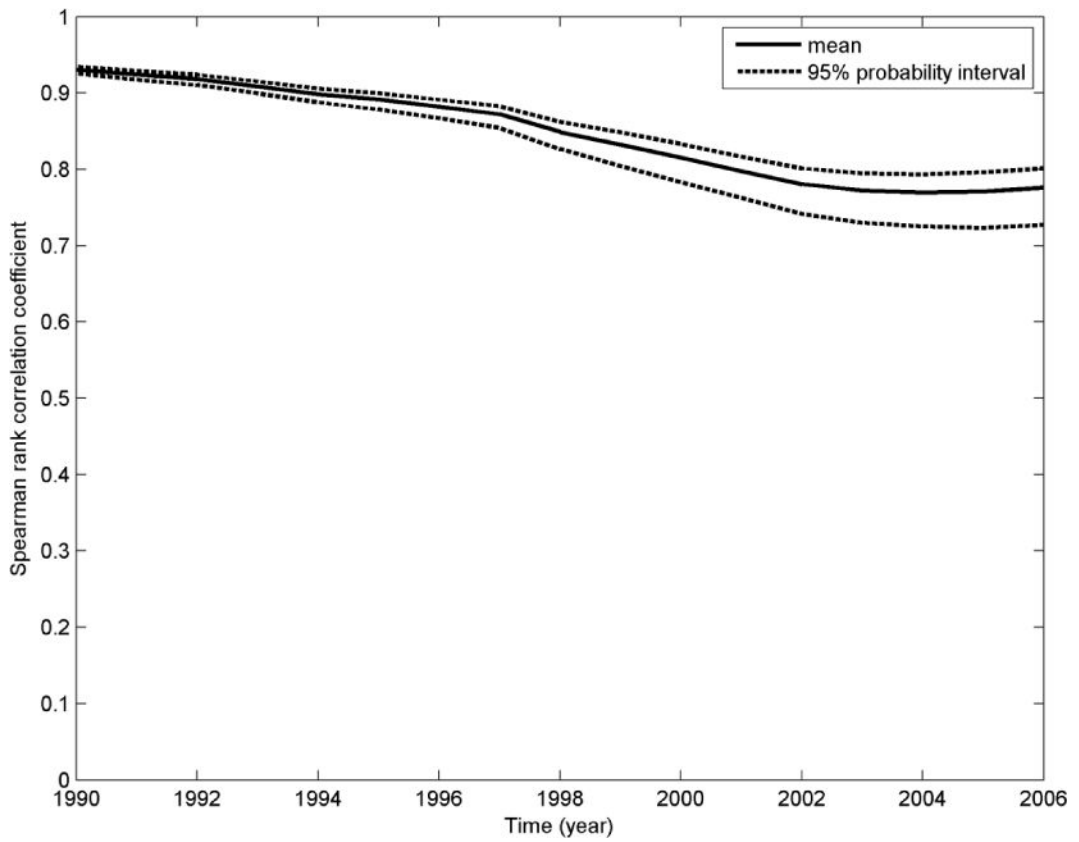
Figure 2.
A plot of the mean and the 95% probability interval of the correlation coefficient between the estimated serum concentrations for each Monte Carlo iterate and the original estimated serum concentrations, for all the participants, over time -MC simulation II (analysis of the impact of variability as well as epistemic uncertainty in independent input parameters)
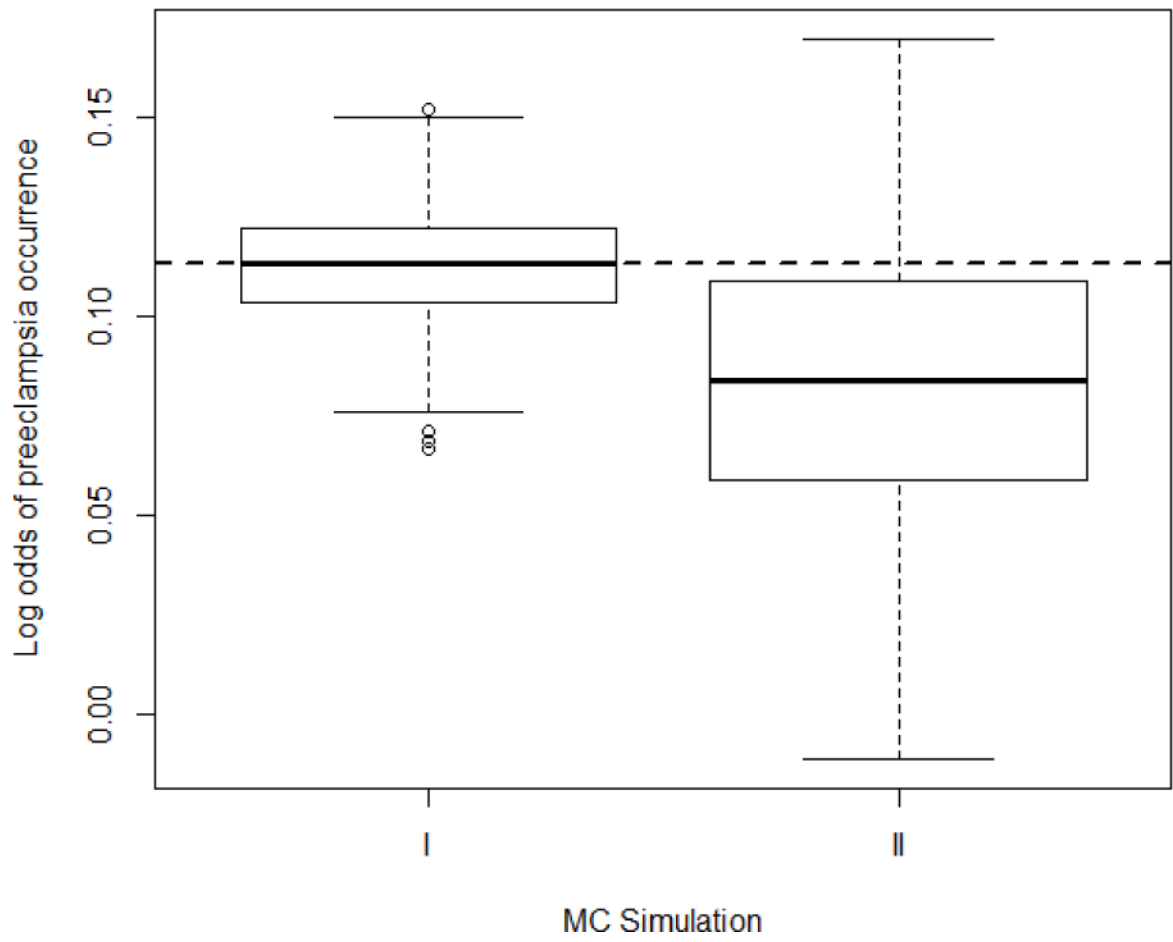
Figure 3.
A boxplot of the point estimates of the log odds of preeclampsia occurrence from the 300 iterations for MC simulation I and MC simulation II. The AOR using the original exposure assignments is shown as a dashed line.

Table 1

The parameters μ_a and σ_a of the log-normal distribution for all age categories of standard water ingestion rates (mL/day) based on the U.S. Environmental Protection Agency (EPA) in the Exposure Factors Handbook (U.S. EPA, 2011)

Age category (years)	Mean	95th Percentile	μ_a	σ_a
1 to < 2	308	893	−1.5697	0.8855
2 to < 3	356	912	−1.3044	0.7370
3 to < 6	417	1099	−1.1702	0.7688
6 to < 11	480	1251	−1.0199	0.7562
11 to < 16	652	1744	−0.7365	0.7859
16 to < 18	792	2002	−0.4941	0.7224
18 to < 21	895	2565	−0.4897	0.8704
21 to < 65	1183	2848	−0.0571	0.6709
65 or higher	1242	2604	0.0719	0.5381

Table 2

The mean and the 95% probability interval (PI) of the median, mean, 25th and 75th percentile serum concentrations at birth (ng/mL), for 10,149 participants for each of the 2 Monte Carlo simulations (300 iterations per simulation)

Simulation	Median (95% PI)	Mean (95% PI)	25th percentile (95% PI)	75th percentile (95% PI)
Modified original	9.4	51.1	5.1	32.5
MC simulation I	9.4 (6.8, 12.5)	58.4 (32.6, 93.0)	5.1 (4.9, 5.2)	33.2 (18.3, 51.3)
MC simulation II	8.7 (5.3, 13.3)	267.6 (63.5, 578.8)	4.9 (4.6, 5.1)	44.8 (12.0, 89.3)

Table 3

The mean and 95% probability interval (PI) of the IQR of the log serum PFOA concentration (ng/mL) across the 300 iterations for each of the 2 MC simulations

Simulation	Mean IQR (95% PI)
Modified original	1.85
MC simulation I	1.85 (1.32, 2.29)
MC simulation II	2.09 (0.96, 2.86)

Table 4

The mean and the 95% probability interval (PI) of the AOR and the percent contribution of exposure uncertainty to total uncertainty for each of the two MC simulations. The AOR (and 95% confidence interval computed from participant sampling variability only) using the original exposure assignments is 1.12 (1.00, 1.26)

Simulation	Mean AOR (95% PI)	Percent contribution of exposure uncertainty
MC simulation I	1.12 (1.00, 1.25)	6.9%
MC simulation II	1.09 (0.97, 1.23)	32.7%