Chemosphere 79 (2010) 243–252

Contents lists available at ScienceDirect

Chemosphere

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

Regression models to estimate total polychlorinated biphenyls in the general US population: 2001–2002 and 2003–2004

Ram B. Jain *, Richard Y. Wang

Centers for Disease Control and Prevention, Mail Stop F-47, 4770 Buford Highway, Chamblee, GA 30329, United States

ARTICLE INFO

Article history:
Received 28 October 2009
Received in revised form 4 February 2010
Accepted 5 February 2010
Available online 26 February 2010

Keywords:
PCB
Congeners
Limit of detection
Geometric mean
NHANES

ABSTRACT

Certain polychlorinated biphenyls (PCB) have long half-lives and, despite the regulatory bans on the industrial pollutants that expose humans to PCB, are detectable in human serum. However, many of them are not detectable because of the small quantities that may be present in body fluids. For this reason, attempts have been made to estimate the total concentration of PCB (ΣPCB) using the relationship between ΣPCB and the concentrations of a few of the PCB congeners which can be reliably measured at detectable levels. PCB 153 or a combination of PCB 153, 138, and 180 have previously been used for this purpose. However, because of the unique populations investigated in these studies, the results are not necessarily applicable to the racially/ethnically heterogeneous US population. We defined ΣPCB as the sum of the concentrations of 12 PCB congeners, and sum of 33 PCB congeners for NHANES 2001–2002 and 2003–2004 respectively. We built regression models in a step-wise fashion using ΣPCB as the dependent variable and age, race/ethnicity, and gender as the covariates for both whole-weight and lipid-adjusted data. In addition, concentration of PCB 153 was used as the continuous independent variable for 2001–2002 models, and PCB 153 and PCB 180 for 2003–2004 models respectively. R^2 for both models for NHANES 2001–2002 was >86%. The R^2 for both NHANES 2003–2004 models was >81%. Thus, the estimate of ΣPCB for the general US population can be improved by considering common demographic variables, such as race/ethnicity, and selected congeners.

Published by Elsevier Ltd.

1. Introduction

Polychlorinated biphenyls (PCB) is a class of persistent organic pollutants listed for elimination by the Stockholm Convention on Persistent Organic Pollutants (http://en.wikipedia.org/wiki/stockhom_convention_on_persistent_organic_pollutants). Estimation of the total concentration of PCB (ΣPCB) in serum from the concentrations of a few specific PCB congeners has been a topic of interest because many of the PCB congeners are present in extremely low concentrations and are not measurable even using current, highly sensitive techniques (Needham et al., 2002, 2007). Some of these estimates are based on descriptive data, for example, the proportion of ΣPCB contributed to by the concentrations of specific congeners like 153. Other approaches are based on regression modeling. Longnecker et al. (2003) fitted models between ΣPCB and PCB 153 across ten studies because it was always present at the highest concentration and was highly-correlated with ΣPCB (R^2 = 0.89 – 0.96). Van der Ven et al. (1992) used the sum of concentrations of PCB congeners 28, 52, 101, 118, 138, 153, and 180 to represent ΣPCB. Glynn et al. (2000), in a study of serum concentrations among Swedish men found that PCB congeners 138, 153, and 180 represented 80% of ΣPCB while ΣPCB was defined as the sum of the concentrations of 10 PCB congeners. Using regression modeling, they also found that the concentrations of PCB 153 alone could be used to represent ΣPCB (R^2 = 0.96) as well as the sum of the concentrations of *di-ortho* PCBs (R^2 = 0.97).

Barr et al. (2006), in a study of 7- and 14-year old children from the Faroe Islands, estimated ΣPCB: (i) as the sum of the concentrations of 11 PCB congeners; (ii) as the sum of the concentrations of PCBs 138/158, 153, and 180 multiplied by 2; and (iii) as determined by the equation attributed to Schulte and Malisch (1984) ((7.03 * [PCB 138] + 6.64 * [PCB 153] + 11.86 * [PCB 180])/3). Using regression modeling, they found that the concentration of PCB 153 comprised about 25% of ΣPCB and the sum of the concentrations of PCBs 138/158, 153, and 180 comprised about 65% of ΣPCB. They also found that ΣPCB, determined as the sum of the concentrations of PCBs 138/158, 153, and 180 multiplied by 2 and by the equation of Schulte and Malisch (1984) over-estimated ΣPCB. Needham et al. (2005) determined ΣPCB as the sum of the concentrations of PCBs 138/158, 153, and 180 multiplied by 1.54 instead of 2.0. Wolff et al. (2005) used the sum of the concentrations of congeners 118, 138, 153, and 180 to represent ΣPCB for both wet weight and

* Corresponding author. Tel.: +1 770 488 0162; fax: +1 770 488 0181.
E-mail address: rij0@cdc.gov (R.B. Jain).

0045-6535/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.chemosphere.2010.02.013

lipid-adjusted concentrations. In a study of fishermen's wives on the Swedish east coast, Grimvall et al. (1997) determined ΣPCB as the sum of the concentrations of *non-ortho* congeners 126 and 169; *mono-ortho* congeners 105, 114, 118, 156, 157, and 167; *di-ortho* congeners 138, 153, 170, and 180; and other congeners 28 and 101 for both whole-weight and lipid-adjusted concentrations. They found a high correlation ($R^2 = 0.99$) between the concentrations of PCB 153 and ΣPCB. Fitzgerald et al. (2005) used wet weight and lipid-adjusted concentration of nine congeners to determine ΣPCB.

DeVoto et al. (1997) reported results from the analysis of PCB data from three studies conducted in the United States. The authors reported that a panel consisting of the concentrations of congeners 74, 118, 138, 153, and 180 might essentially yield the same results as a larger panel consisting of all quantified PCBs because of the high correlations between specific congeners. Daniels et al. (2003) summed whole-weight concentrations in maternal serum of 11 PCB congenersto determine ΣPCB.

Most of the investigations described above studied special populations only. Because race and ethnicity may affect the relationship between the concentrations of the specific PCB congeners and ΣPCB, these studies of specific populations may not be applicable to the heterogeneous US population. Age also affects the relationship between the concentrations of the specific PCB congeners and ΣPCB as shown by Grimvall et al. (1997). McGraw and Waller (2009) found that females of African American ancestry had higher levels of PCB 101 and 118 than females of mixed race/ethnicity. These racial/ethnic differences were attributed to CYP 450 mediated metabolic differences. Patterson et al. (2009), using National Health and Nutrition Examination Survey (NHANES) 2003–2004 data reported significantly lower levels of total PCB concentrations among Mexican Americans than both non-Hispanic whites and non-Hispanic blacks. Patterson et al. (2009) also reported that males appear to have slightly higher levels of total PCB concentrations than females. In our preliminary investigation, using NHANES 2001–2002 data (http://www.cdc.gov/nchs/about/major/nhanes/nhanes2001-2002/lab01_02.htm), we determined that race/ethnicity and age were statistically significant factors in defining the relationship between ΣPCB and the concentrations of the specific PCB congeners. We also found that age was a better predictor of ΣPCB than race.

The 10 studies reviewed above used both wet or whole-weight and lipid-adjusted PCB concentrations. Thus, it is important that the regression model used to evaluate the relationship between ΣPCB and its congeners is built separately for whole-weight ΣPCB_{WW} and lipid-adjusted ΣPCB_{LA}. It is possible that the models for ΣPCB_{WW} and ΣPCB_{LA} may be similar or identical because of the high correlation ($R^2 = 0.90$) between ΣPCB_{WW} and ΣPCB_{LA} as shown by Grimvall et al. (1997), but this needs further investigation.

The objective of this investigation was to use linear regression modeling to estimate serum ΣPCB in the general US population. Linear regression models were developed, independently using NHANES 2001–2002 data (http://www.cdc.gov/nchs/about/major/nhanes/nhanes2001-2002/lab01_02.htm) and NHANES 2003–2004 data (http://www.cdc.gov/nchs/about/major/nhanes/nhanes2003-2004/lab03_04.htm) because of differences between these datasets in detection frequencies for each PCB congener and the number of PCB congeners available for analysis. Separate models were built for whole-weight data, ΣPCB_{WW}, and for lipid-adjusted data, ΣPCB_{LA}.

In order to align this manuscript with Stockholm Convention on Persistent Organic Pollutants (http://www.pops.int/documents/ConvText/ConvText_en.pdf), we also present data on the sum of seven indicator PCBs, $\sum PCB_7$, namely, PCB congeners 28, 52, 101, 118, 138/158, 153, and 180.

2. Material and methods

We selected available data from NHANES 2001–2002 and 2003–2004. NHANES used a complex, stratified, multistage, probability sampling designed to be representative of the civilian, noninstitutionalized US population based on age, sex, and race/ethnicity (CDC, 2004). We downloaded publically available data from PCB files (http://www.cdc.gov/nchs/about/major/nhanes/nhanes2001-2002/lab01_02.htm and http://www.cdc.gov/nchs/about/major/nhanes/nhanes2003-2004/lab03_04.htm) and demographic files (http://www.cdc.gov/nchs/nhanes/demo01_02.htm and http://www.cdc.gov/nchs/nhanes/demo03_04.htm). The concentrations of PCB reported in these datasets were measured in serum by gas chromatography/isotopic dilution-high resolution mass spectrometry (CDC, 2005). We calculated the serum concentration of total lipids by using the concentrations for triglycerides and total cholesterol measured by an enzymatic method: total lipids = (2.27 * total cholesterol + triglycerides + 62.3) mg dL^{-1} (Phillips et al., 1989).

For the NHANES 2001–2002, after the two files were merged and pregnant females were excluded, data from a total of 2427 subjects aged 12 and older were available for analysis. In the sample, there were 1184 males, 1243 females; 1086 non-Hispanic whites (NHW), 556 non-Hispanic blacks (NHB), 596 Mexican Americans (MA), and 189 from other races/ethnicities (OTH); and 1106 subjects in the age group 12–29 years, 555 in the age group 30–49 years, and 766 in the age group 50+ years. For the 2003–2004 dataset, out of the total sample size of 2194, there were 1129 males and 1065 females; 1000 NHW, 522 NHB, 502 MA, and 170 OTH; 959 subjects in the age group 12–29 years, 479 in the age group 30–49 years, and 756 in the age group 50+ years.

For 2001–2002, 31 PCB congeners were available for analysis (Table 1). However, for 17 of these congeners, the percent observations which were below the limit of detection (LOD) were >70% (Table 1) and were not considered to significantly contribute to ΣPCB. A total of 14 PCB congeners with adequate numbers above the LOD were available. For most of these congeners the percent of missing values was about 10%, but for PCB 126 and for PCB 169, the percent of missing values was more than 50%. Thus, including these two congeners to estimate ΣPCB would have resulted in loss of more than half the data. For this reason, these two congeners were also excluded from further analysis in any of the models built for this study. All values below the LOD were imputed as $LOD/\sqrt{2}$. For 2003–2004, using the same criterion as above, 33 congeners were available for analysis (Table 1).

For 2001–2002, ΣPCB_{WW} was computed as the sum of the whole-weight concentrations of 12 PCB congeners 74, 99, 118, 138, 146, 153, 170, 180, 187, 194, 196, and 199. ΣPCB_{LA} was computed as the sum of the lipid-adjusted concentrations of the same 12 PCB congeners. The distributions of both ΣPCB_{WW} and ΣPCB_{LA} were highly skewed to the right. Thus, we log transformed both ΣPCB_{WW} and ΣPCB_{LA} before using them as the dependent variable in the regression models. Gender (males, females), race/ethnicity (NHW, NHB, MA, OTH) and age group were used as covariates in the models. These covariates were based on our preliminary analysis in which we found statistically significant differences between MA when compared with NHW and NHB and between different age groups. To simplify and avoid over-parameterization of the models, based on similar increases from one age category to another for each race/ethnicity, we collapsed age into three or four categories. Thus, as a starting point, each model had gender, race/ethnicity, and age as the covariates. For 2003–2004, ΣPCB_{WW} was computed as the sum of the whole-weight concentrations of 33 PCB congeners 28, 44, 49, 52, 66, 74, 87, 99, 101, 105, 110, 118, 138/158, 146, 149, 151, 153, 156, 157, 167, 170, 172, 177, 178, 180, 183, 187, 194, 195, 196, 199, 206, 209.