Evaluation of background 2,3,7,8 - PCDD/F congener profiles in human serum collected during NHANES 2001/2002 using principal components analysis

Scott P[1] Haws L[2], Scott L[3], Harris M[3]
[1]ChemRisk, 20 Stanwix St, Pittsburgh, Pennsylvania USA; [2]Chemrisk, 3420 Executive Center Road, Austin, Texas USA; [3]10375 Richmond Avenue, Houston, Texas USA

Abstract
With the increasing use of PCDD/F serum measurements to evaluate exposure, it is important to understand how the congener profile in the potentially exposed population compares to the congener profile associated with background exposure. In this study, we have performed a principal components analysis (PCA) of the 2001/2002 NHANES PCDD/F data from the general U.S. population to determine if the background 2,3,7,8 PCDD/F congener profiles are affected by age, gender, or race. Overall, the results of the PCA indicated that the congener profile for the NHANES participants changes with age and gender. In general, the contribution of PCDFs decreases with increasing age and the contribution from PCDDs other than TCDD and OCDD increases with increasing age. In addition, males tend to have higher contributions of PCDD/Fs other than TCDD and OCDD than females of the same age. These results indicate that age and gender need to be accounted for when PCA is used to compare the congener profiles from the 2001/2002 NHANES data to those of a potentially exposed population.

Introduction
With the increasing use of serum measurements to evaluate populations potentially exposed to polychlorinated dibenzo-p-dioxins and dibenzofurans (PCDD/Fs) and dioxin-like polychlorinated biphenyls (PCBs) in the environment, the determination of background concentrations of PCDD/Fs and PCBs in the general population is important in order to put those measurements in context. Recently, Ferriby et al.[1] analyzed the weighted 2001/2002 National Health and Nutrition Examination Survey (NHANES) data to assess potential differences in mean total 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) toxic equivalent (TEQ) concentrations between various groups of individuals and to determine serum background TCDD TEQ concentrations for PCDD/Fs and dioxin-like PCBs in the general U.S. population. While it is important to compare the PCDD/F serum concentrations in a potentially exposed population to background concentrations on a TEQ basis, it is also important to understand how the relative distribution of the 17 2,3,7,8 PCDD/F concentrations in a potentially exposed group compares to the relative distribution in the general U.S. population. There may be individuals within a potentially exposed population who have above background TEQ concentrations because of higher dietary exposures resulting from regional differences in the type and amount of food consumed unrelated to the PCDD/F exposures of interest.

To compare the congener profile of an individual to the background congener profiles of the general U.S. population, multivariate statistical methods must be used. Principal components analysis (PCA) is a common statistical method used to evaluate multivariate environmental data sets. In this study, we will present the results of a PCA of the 2001/2002 NHANES 2,3,7,8-PCDD/F data and discuss how the congener profiles within the 2001/2002 NHANES data vary based on age, gender, and race.

Materials and Methods
PCA is a multivariate statistical technique that transforms a set of correlated variables into a subset of factors or principal components that are linear functions of the original variables and are uncorrelated.[2] The purpose of PCA is to reduce the dimensions of the data set so that trends in the data can be more easily examined. In terms of 2,3,7,8 PCDD/F congener data, because the concentrations many of the congeners are correlated with each other, PCA is able to transform a data set consisting of the 17 congeners into a subset of factors that is often less than 5. This method has been used in a similar manner at other sites to compare human serum PCDD/F congener profiles to environmental congener profiles.[3,4,5,6,7]

Only data from 2001/2002 NHANES participants with complete PCDD/F profiles were selected for this analysis

resulting in a sample set of 1,104 participants. This differs from the sample size of Ferriby et al.[1] of 1,081 participants because their participants were required to have complete PCDD/F and dioxin-like PCB profiles to be included, while for this current study focused only on PCDD/Fs. Congener concentrations that were reported to be below the limit of detection were included as the limit of detection divided by the $\sqrt{2}$. For the PCA, the relative fraction of each congener to the total 2,3,7,8 PCDD/F congeners was estimated for each NHANES participant by dividing the congener concentration by the sum of the 2,3,7,8 PCDD/F congener concentrations.[8] The relative fraction data were then range transformed as nonparametric method for normalization similar to what has been done in other studies.[9] The PCA was then performed on the covariance matrix of the range-transformed data.

Results and Discussion

The PCA of the 2001/2002 NHANES PCDD/F data resulted in three factors that described 82.2% of the variance in the original data set with the first factor describing 63.8%, the second factor describing 11.8%, and the third factor describing 6.6% of the variance in the original data set.

Factor 1 is characterized by high positive factor loadings for 1,2,3,7,8-PeCDF, 1,2,3,7,8,9-HxCDF, and 2,3,4,6,7,8-HxCDF, 1,2,3,4,7,8,9-HpCDF, and OCDF (0.18 – 0.24), and a high negative factor loading for OCDD (-0.1). For this factor, a positive factor score for a participant indicates that 1,2,3,7,8-PeCDF, 1,2,3,7,8,9-HxCDF, and 2,3,4,6,7,8-HxCDF, 1,2,3,4,7,8,9-HpCDF, and OCDF contribute more to the total 2,3,7,8 concentration than OCDD and negative factor scores indicate that OCDD contributes more than the PCDF congeners. Thus, this factor describes the relationship between the OCDD and PCDF contributions for the NHANES participants.

The factor 1 scores tend not only to decrease with increasing age but the variance of the factor 1 scores for a given age also tends to decrease with age (Figure 1). Males tend to have higher and more variable factor 1 scores than females for a given age (Figure 1). While non-Hispanic blacks have slightly lower factor 1 scores than the other race categories, there does not appear to be any substantial differences in factor 1 scores with race (Figure 2). Based on this factor, the relative contributions of PCDFs appear to decrease with age relative to OCDD in the NHANES study population, indicating that higher OCDD contributions are expected for older members of the general U.S. population.

Factor 2 is characterized by positive factor loadings of PeCDD, the HxCDDs, HpCDD, and 2,3,4,7,8-PeCDF (0.05 – 0.11) and negative factor loadings of OCDD (-0.11). For this factor, positive factor scores indicate that PeCDD, HxCDDs, and HpCDD contribute more to the total 2,3,7,8 congener concentration than OCDD. Conversely, negative factor scores indicate that OCDD contributes more. This factor describes the relationship between the relative contribution of the PeCDD, HxCDDs, and HpCDD versus the contribution of OCDD.

The factor 2 scores tend to increase with increasing age indicating that the contributions of PeCDD, HxCDDs, and HpCDD increase with increasing age (Figure 1). Similar to factor 1, males tend to have higher factor 2 scores than females at the same age (Figure 1). Non-Hispanic whites tend to have slightly higher factor 2 scores than Mexican-Americans, other Hispanics, or participants of other races (Figure 2).

Factor 3 is characterized by positive factor loadings of PeCDD, 1,2,3,6,7,8-HxCDD, OCDD, 2,3,4,7,8-PeCDF (0.027 – 0.047) and negative factor loadings of HpCDD (-0.13). For this factor negative factor scores indicate high contributions of HpCDD and positive scores indicate high contributions from PeCDD, 1,2,3,6,7,8-HxCDD, OCDD, 2,3,4,7,8-PeCDF. Factor 3 scores tend to slightly increase with increasing age, indicating increased contributions of PeCDD, 1,2,3,6,7,8-HxCDD, OCDD, 2,3,4,7,8-PeCDF (Figure 1). There is no apparent general difference in factor 3 scores based on gender although there appears to be a subpopulation of males with fairly low factor 3 scores across all ages indicating high contributions from HpCDD (Figure 1). Non-Hispanic blacks and whites tend to have slightly higher factor 3 scores than Mexican-Americans, other Hispanics, and participants of other races (Figure 2).

Figure 3 presents the factor scores plots for each pair of factors. The plot of factor 1 and 2 shows a clear trend

of decreasing factor 1 scores and increasing factor 2 scores with age. This indicates that the majority of the variance in the congener profiles of the NHANES participants is due to changes in the relative contributions of PeCDD, HxCDDs, HpCDD, PCDFs, and OCDD with increasing age. The plot of factors 1 and 3 indicates that HpCDD contributions tend to be uniform with increasing age and as the PCDF contributions decrease with increasing age. Based on the plot of factors 2 and 3, HpCDD contributions tend to become more variable with increasing PCDD contributions.

In summary, this analysis shows that there are substantial apparent differences in the 2,3,7,8 PCDD/F congener profiles with respect to age and gender and minor differences with respect to race. Because of these differences, any comparisons of the NHANES participants' congener profiles to a potentially exposed population need to account for the age and gender differences. This similar to what was observed when the dioxin-like PCB data for the 2001/2002 NHANES participants were evaluated[10]

References

1. Ferriby LL, Knutsen JS, Harris M, Unice KM, Scott P, Nony P, Haws LC, Paustenbach D. *J Expo Sci Environ Epidemiol* 2006; online July 5.
2. Manly, BF. *Multivariate Statistical Methods – a primer, 2nd edition*. Chapman & Hall, 1994, pgs 76-92.
3. Hsu JF, Guo YL, Yang SY, Liao PC. *Chemosphere* 2005, 61:1231-1243.
4. Lee CC, Chen HL, Su HJ, Guo YL, Liao PC. *Chemosphere*, 2005. 59:1465-1474.
5. Lee CC, Guo YL, Kuei CH, Chang HY, Hsu JF, Wang ST, Liao PC. *Chemosphere*, 2006, 65:436-448.
6. Orloff KG, Hewitt D, Metcalf S, Kathman S, Lewin M, Turner W. *J Expo Anal Environ Epidemiol.*, 2001,11:352-358.
7. Hedgemann E, Chang SC, Towey T, Demond A, Adriaens P, Chen Q, Franzblau A, Gillespie BW, Sima C, Garabrant D. *Organohalogen Compounds*, 2006.
8. Shields WJ, Tondeur Y, Benton L, Edwards MR. In: *Environmental Forensics, Contaminant Specific Guide*, Morrison RD, Murphy BL ed., Elsevier, Boston, 2006: 293-312.
9. Wenning R, Paustenbach D, Johnson G, Ehrlich R, Harris M, Bedbury H. *Chemosphere* 1993; 27:55-64.
10. Scott P, Haws L, Scott L, Harris M. *Organohalogen Cmpds*. 2007.

Figure 1. Factor scores with age and gender

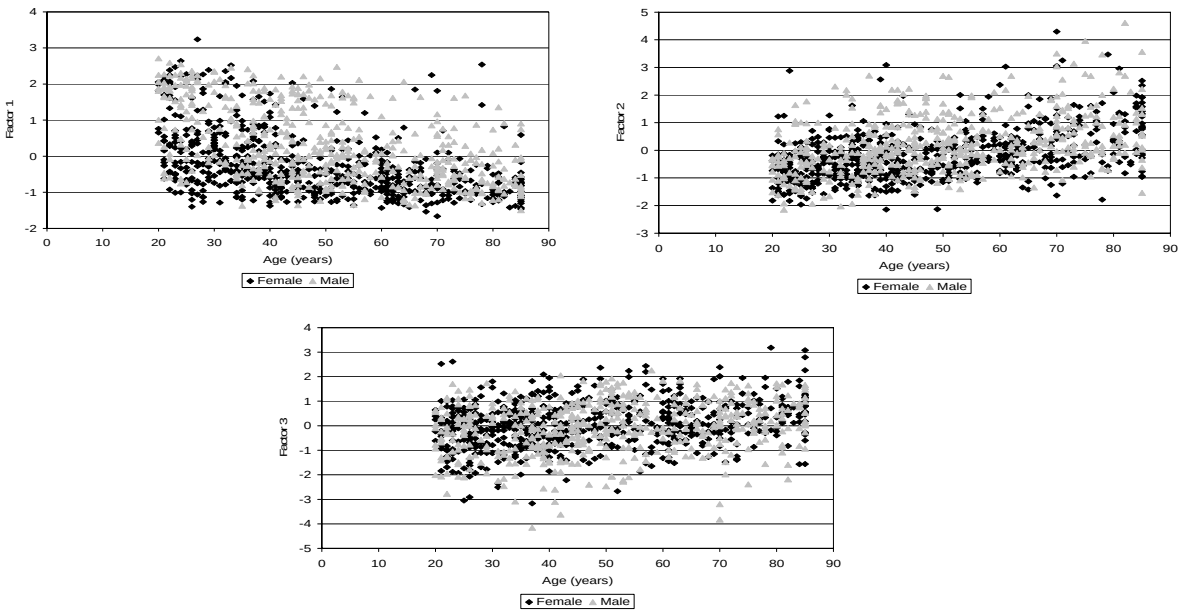

Figure 2. Box plots of factor scores by race category. MA is Mexican-American, NHB is non-Hispanic black, NHW is non-Hispanic white, OH is other Hispanic, and OR is other race.

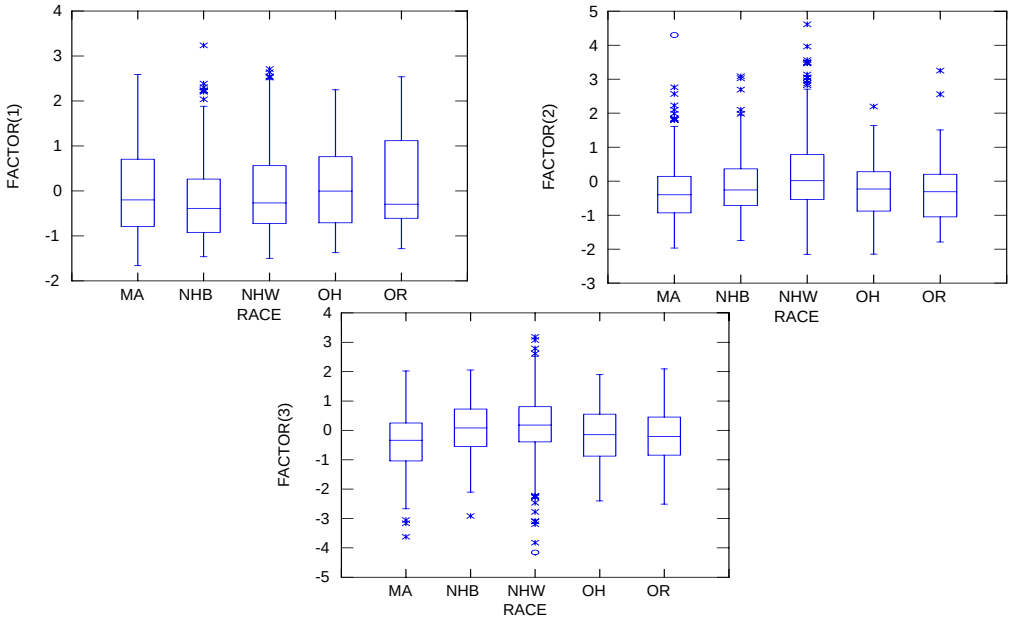

Figure 3. Factor Scores Plots of Factors 1 vs 2, 1 vs 3, and 2 vs 3.