PRINCIPAL COMPONENTS ANALYSIS OF HOUSEHOLD DUST CONCENTRATIONS OF PCDDs, PCDFs, AND PCBs FROM A COMMUNITY IN MICHIGAN, USA

Zwica L[1], Chang S-C[2], Towey T[2], Knutson K[1], Adriaens P[2], Demond A[2], Chen Q[3], Gillespie BW[3], Franzblau A[1], Garabrant D[1]

[1]Department of Environmental Health Sciences, University of Michigan School of Public Health, 109 S Observatory, Ann Arbor, MI 48109; [2]Department of Civil and Environmental Engineering, University of Michigan College of Engineering, 1351 Beal, Ann Arbor, MI 48109; [3]Department of Biostatistics, University of Michigan School of Public Health, 109 S Observatory, Ann Arbor, MI 48109

Introduction

"Dioxins" or dioxin-like compounds are a family of structurally related chemicals including polychlorinated dibenzodioxins (PCDDs), polychlorinated dibenzofurans (PCDFs) and polychlorinated biphenyls (PCBs). The University of Michigan Dioxin Exposure Study (UMDES) was undertaken in response to concerns among the population of Midland and Saginaw Counties that dioxin-like compounds from the Dow Chemical Company facilities in Midland have resulted in contamination of soils in the Tittabawassee River flood plain and areas of the City of Midland. There is concern that people's body burdens of PCDDs, PCDFs and PCBs may be elevated because of environmental contamination. A central goal of the UMDES is to determine the factors that explain variation in serum congener levels of PCDDs, PCDFs, and PCBs, and to quantify how much variation each factor explains. Household dust concentration and loading were included in the list of potentially explanatory factors to investigate. The descriptive statistics and distributions of the resulting household dust concentration and loading data from the sampled communities have been described elsewhere.[1]

In order to better understand the distribution of PCDDs, PCDFs and PCBs in the household dust of UMDES participants, an analysis of congener patterns in samples was performed using multivariate chemometric methods. Principal component analysis (PCA) and hierarchical cluster analysis (HCA) were performed on the complete UMDES household dust data set. The PCA and HCA output were used to generate a concentration heatmap and cluster-centroid pattern profiles. Additionally, the geographic distribution of the clusters was evaluated using a geographic information system. From the analysis, clusters emerged with distinct characteristics in congener patterns, locations, and concentration ranges.

Materials and Methods

Household Dust Sampling and Analysis: Vacuum sampling for household dust was conducted in the home of each UMDES participant. The sampling protocol was based, with minor modifications, on the American Society for Testing and Materials (ASTM) method "Standard Practice for Collection of Floor Dust for Chemical Analysis".[2] The sample was taken from two sampling locations that presented the highest potential for human contact with household dust and dirt. The locations were generally a frequently occupied living space (e.g., living or family room) and a high traffic hallway or pathway. Samples were taken from both hard and soft surfaces and were not taken of undisturbed dust in generally inaccessible areas. The total surface area of all of the sampling areas was recorded, as well as the surface types from which the samples were taken. Analyses were performed by Alta Analytical Laboratory, Inc. (El Dorado Hills, California, USA) for the World Health Organization designated 29 PCDD, PCDF, and PCB congeners[3] using US EPA methods 8290[4] and 1668[5].

Data Cleaning: All data received from the analytical laboratory went through a data cleaning step to ensure data quality. All values below limit of detection (LOD) were replaced with the LOD divided by the square root of two ($LOD/\sqrt{2}$) to create the untransformed household dust data set.

Data Transformation: Since congener data exhibited log-normal distributions, a natural logarithm transformation of ln(x+1) was undertaken. A constant-row-sum transformation was used to convert the sum of each row to unity and the transformed concentration value of each congener in each sample was converted to a fraction of unity. Finally, a range transformation was applied to each column of the dataset to ensure the variation within each congener would be similar. This final step kept the PCA from being driven by several congeners with extreme variation.[6]

Principal Component Analysis: PCA was performed using Minitab software (Minitab, Inc., State College, Pennsylvania, USA). A Scree plot, a cumulative variance plot, and principal component loading graphs were generated. The principal components that accounted for 95% of the cumulative variance were selected for further use in the HCA.[7, 8]

Hierarchical Cluster Analysis: Using the selected principal components, HCA was performed based on a correlation matrix and average linkage of Euclidean distance between samples.[7] Each household dust sample in the dataset was assigned a cluster membership. A dendrogram, indicating the similarity between clusters, was generated. Also, two and three-dimensional principal component score plots - grouped by cluster membership, were produced.

Heatmap Representation: The constant-row-sum transformed data (not range transformed) was sorted according to cluster membership. Using the sorted data, a heatmap was generated to represent the congener patterns of all household dust samples in a single graph. In addition, supplemental information items, such as untransformed household dust samples exceeding a particular threshold concentration, were indicated in columns adjacent to the congener pattern for each sample. From this visual representation, clear differences in the congener patterns of the clusters emerge.

Centroid Pattern Representation: Minitab software allows for the creation of a distance to cluster-centroid matrix. The sample with the smallest value for the distance to each centroid was selected to represent that cluster. The congener pattern of each cluster centroid was produced using the untransformed household dust data to create 100% stacked bar graphs. Both original concentration and TEF-weighted patterns were produced.

Geographical Representation: Clustered data were projected onto a map to evaluate the geographic distribution of the clusters using ArcGIS 9.1 (ESRI, Inc., Redlands, CA, USA). Inferences regarding sources of congener pattern variability were made.

Results and Discussion

Results and discussion will not be available until after complete study results have been presented to the affected communities in August of 2006.

Acknowledgments

The authors acknowledge the Dow Chemical Company for funding the study and Ms. Sharyn Vantine for her continued assistance.

References

1. Zwica L, Knutson K, Towey T, Hedgeman E, Franzblau A, Chen Q, Lee S-Y, Sima C, Gillespie BW, Adriaens P, Demond A, Lepkowski J, Ward B, Ladronka K, Olson K, Sinibaldi J, Chang S-C, Gwinn D, Swan S, Garabrant D. *Organohalogen Comp* 2006 (forthcoming).
2. American Society for Testing and Materials, Standard Practice for Collection of Floor Dust for Chemical Analysis, Designation D 5438-00, Reprinted from the Annual Book of ASTM Standards, Philadelphia, PA.
3. Van den Berg M, Birnbaum L, Bosveld AT, Brunstrom B, Cook P, Feeley M, Giesy JP, Hanberg A, Hasegawa R, Kennedy SW, Kubiak T, Larsen JC, van Leeuwen FX, Liem AK, Nolt C, Peterson RE, Poellinger L, Safe S, Schrenk D, Tillitt D, Tysklind M, Younes M, Waern F, and Zacharewski T. *Environmental Health Perspectives* 1998;106:775.
4. United States Environmental Protection Agency. Method 1668, Revision A: Chlorinated biphenyl congeners in water, soil, sediment, and tissue by HRGC/HRMS. Washington, DC: Office of Water, 1999.
5. United States Environmental Protection Agency (US EPA). Method 8290: Polychlorinated dibenzodioxins (PCDDs) and polychlorinated dibenzofurans (PCDFs) by high-resolution gas chromatography/high-resolution mass spectrometry (HRGC/HRMS). Washington, DC: Office of Solid Waste and Emergency Response, 1994.
6. Johnson GW, Ehrlich R. State of the art report on multivariate chemometric methods in environmental forensics. *Environmental Forensics* 2002; 3:59.
7. Jolliffe IT. *Principal Component Analysis*. Springer-Verlag, New York. 1986; 8.
8. Hair Jr JF, Black WC. Cluster analysis. In: *Multivariate Data Analysis*, Hair Jr JF, Anderson RE, Tatham RL, Black WC. (ed.), Prentice-Hall Inc, New Jersey, 1998: 147.