LINEAR REGRESSION MODELING TO PREDICT HOUSEHOLD DUST PCDF CONGENER CONCENTRATIONS

Knutson K[1], Zwica L[1], Lee S-Y[2], Hong B[1], Chen Q[2], Towey T[3], Gillespie BW[2], Demond A[3], Adriaens P[3], Lepkowski J[4], 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 Biostatistics, University of Michigan School of Public Health, 109 S Observatory, Ann Arbor, MI 48109; [3]Department of Civil and Environmental Engineering, University of Michigan College of Engineering, 2340 G.G. Brown Building, Ann Arbor, MI 48109; [4]Survey Research Center, Institute for Social Research, University of Michigan, 426 Thompson Street, Ann Arbor, Michigan 48104

Abstract
Household dust samples were obtained from the residences of eligible participants in the University of Michigan Dioxin Exposure Study (UMDES). Linear regression models were used to determine which factors explained the variations in polychlorinated dibenzofuran (PCDF) congener concentrations in household dust measured from the participants in the UMDES and to quantify how much of the variation each factor explained. Potential dust predictor variables were derived from the UMDES questionnaire, dust sampling field data notes and results from the household dust and soil sampling laboratory analyses. Household dust samples were obtained from 764 study participants over the course of three sample seasons. In general, older floor surfaces and higher PCDF congener concentrations in the soil samples were associated with higher PCDF congener concentrations in household dust. Sampling season and region were generally associated with lower PCDF congener concentrations in household dust. PCDF congener concentrations in the soil samples and sampling season provided the greatest explanation of the variance in the models.

Introduction
The University of Michigan Dioxin Exposure Study (UMDES) was conducted to explain the variation in serum levels of polychlorinated dibenzodioxins (PCDDs), polychlorinated dibenzofurans (PCDFs) and polychlorinated biphenyls (PCBs) and to quantify how much of the variation each factor explained. The study was undertaken in response to concerns of environmental contamination in the Tittabawassee River floodplain in Midland and Saginaw Counties in Michigan and in areas within the City of Midland, Michigan. Household dust was determined to be one of many potential explanatory factors for the variation in serum levels of PCDDs, PCDFs and PCBs and household dust samples were obtained from eligible UMDES study participants. The purpose of the present study is to describe which factors explained the variations in PCDF congener concentrations in household dust measured from participants in the UMDES and to quantify how much of the variation each factor explained.

Materials and Methods
UMDES study participants were recruited from five populations, described elsewhere[1], from areas in Midland, Saginaw, Bay, Jackson and Calhoun Counties in Michigan using a two-stage area probability household sample design. In order to be eligible for household dust sampling, a participant had to be eligible to provide a serum sample, had to be the owner of their residence and had to have lived in their residence for at least five years. Detailed methods for sample selection, recruitment, eligibility and consent are described elsewhere.[2]

The household dust sampling protocol was based on the American Society for Testing and Materials (ASTM) method "Standard Practice for Collection of Floor Dust for Chemical Analysis".[3] A High Volume Small Surface Sampler (HVS3), capable of capturing 99.95% of particles above 0.3μm aerodynamic mean diameter, was used to collect each household dust sample. Sample locations were selected by the samplers after entering a residence and were generally taken from one or two locations with a high potential for human exposure. Bulk dust and filter samples were combined and a single household dust sample from each residence was analyzed by Vista

Analytical Laboratory, Inc. (El Dorado Hills, California, USA) for the WHO designated 29 PCDD, PCDF, and PCB congeners[4] using US EPA methods 1668[5] and 8290[6]. The full UMDES protocol, including the household dust sampling protocol[7], is available online at www.umdioxin.org.

Congener concentration is the outcome of interest presented in this report. Congener concentration is the amount of a PCDD, PCDF or PCB congener per gram of dust, measured in pg/g or parts per trillion (ppt). Household dust levels can also be measured in terms of congener loading. Congener loading is the amount of a PCDD, PCDF or PCB congener per square meter of floor surface, measure in pg/m^2. Congener loading was also calculated and analyzed and the results of those analyses will be presented in future reports.

Potential household dust predictor variables were derived from the UMDES questionnaire (i.e., work history, property use and remediation variables), household dust sampling field data notes (i.e., sampling season and floor surface variables) and results from the household dust and soil sampling laboratory analyses. Continuous variables were reported in years or days to account for a participant's total cumulative exposure to a potential source of contamination while living in their current residence. Based on the data available, several potential predictor variables were also created to group participants into different exposure categories. Categorical variables were created to explain sample season, floor surface sampled and floor replacement after a flood. Sample season was determined based on the month the household dust was sampled. Each floor surface sampled was classified as a high pile carpet (i.e., velvet carpet), low pile carpet (i.e., berber carpets) or a hard floor surface. A variable was also created for residences that had been affected by river flooding to differentiate between floor surfaces that had been replaced after the flooding and floor surfaces that had not been replaced after the flooding.

SAS version 9.1 statistical software was utilized to complete all analyses.[8] The PCDF congener concentration outcome variables were transformed to a log base10 scale. Histograms confirmed that the distributions of the outcome variables were extremely skewed in the original scale and were normalized after transformation. Descriptive statistics were calculated using both unweighted data and survey sample weighted data. Graphs were also created to explain important relationships between household dust and soil PCDF congener concentrations. Tables and boxplots, available on the UMDES website, www.umdioxin.org, provide descriptive statistics and distributions by region for the TEQ and 29 dioxin congeners. [Note: TEQ values presented were based on the World Health Organization 2005 TEFs].[4]

Linear regression modeling used data from five imputed data sets and survey sample weights generated based on the 764 subjects who had household dust measurements taken. Final variable selection was based on forward and backward selection techniques after testing for interactions and collinearity. For example, sampling time was not included in the regression models because sampling time was directly related to dust loading or how much dust there was on a floor surface per square meter of floor surface. The final variable list only included a soil concentration variable and other statistically significant variables.

Overall adjusted r-square values were calculated for each PCDF congener concentration regression model to describe how well each model explained the total variation. Additionally, the change in the adjusted r-square (ΔR^2) was calculated after removing each significant variable one at a time to determine which predictor variables provided the greatest explanation of the variance in each model.

Results and Discussion

Household dust samples were obtained from 764 study participants over the course of three sample seasons. 205 of those samples were from the Floodplain region, 161 from the Near Floodplain region, 168 from other parts of Midland/Saginaw not in the floodplain, 32 from the Midland Plume and 198 from the Jackson and Calhoun region. 224 household dust samples were taken in Fall 2004, 432 samples were taken in Summer 2005 and 108 samples were taken in Spring 2005. The mean age of the floor surfaces sampled was 12.2 years. Prior to sampling, five residences had been remediated as a result of dioxin contamination.

For all congeners, scatterplots showed a positive correlation between the soil PCDF concentrations around the house and the household dust PCDF concentrations. As the soil PCDF concentrations around the house increased, the household dust PCDF concentrations also increased. Scatterplots are not included in this report.

Regression models, using both forward and backward selection methods, produced similar results and therefore, only the backward selection method is presented. Table 1 shows the parameter estimates and the overall adjusted r-square values for the PCDF congener concentration regression models. Across all congeners (expect OCDF), older floor surface were consistently associated with higher PCDF congener concentrations in household dust. PCDF concentrations in the soil samples, in general, were still associated with higher congener concentrations in household dust after controlling for all other important covariates in the model. For 12378PeCDF, 123768 HxCDF, and 1234678 HpCDF, remediation for dioxins was associated with higher congener concentrations in household dust. For most PCDF congeners, sampling in the summer and region were associated with lower congener concentrations in household dust. Higher dust loading levels were also associated with lower congener concentrations in household dust for the majority of the PCDF congeners.

The overall adjusted r-square values show that the backward selection models explained about 17% to 36% of the variation in household dust PCDF congener concentrations. Based on the change in the adjusted r-square values, the PCDF concentrations in the soil samples and sampling season generally provided the greatest explanation of the variance in the models. The change in the adjusted r-square (ΔR^2) values will be described in a future report. Congener concentration and congener loading are both important outcomes for household dust studies and in future reports, congener loading will also be reported.[9]

Acknowledgments
Financial support for this study comes from the Dow Chemical Company through an unrestricted grant to the University of Michigan. The authors acknowledge Ms. Sharyn Vantine for her continued assistance and Drs. Linda Birnbaum, Ron Hites, Paolo Boffetta and Marie Haring Sweeney for their guidance as members of our Scientific Advisory Board.

References

1. Franzblau A, Garabrant, D, Adriaens P, Gillespie B, Lepkowski J, Olson K, Lohr-Ward B, Ladronka K, Sinibaldi J, Chang S-C, Chen Q, Demond A, Gwinn D, Hedgeman E, Hong B, Knutson K, Lee S-Y, Sima C, Towey T, Wright D, Zwica L. *Organohalogen Comp* 2006:68;208.
2. Lepkowski J, Olson K, Ward B, Ladronka K, Sinibaldi J, Franzblau A, Adriaens P, Gillespie BW, Chang S-C, Chen Q, Demond A, Gwinn D, Hedgeman E, Knutson K, Lee S-Y, Sima C, Swan S, Towey T, Zwica L, Garabrant D. *Organohalogen Comp* 290006:68;212.
3. 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.
4. Van den Berg M, Birnbaum L, Denison M, De Vito M, Farland W, Feeley M, Fiedler H, Hakansson H, Hanberg A, Haws L, Rose M, Safe S, Schrenk D, Tohyama C, Tritscher A, Jouko T, Tysklind M, Walker N, Peterson RE. *Toxicological Sciences* 2006;93(2):223.
5. 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.
6. 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.
7. Zwica L, Luksemburg W, Kennington B, Watka M, Shinn N, Chen J, Wenning R, Knutson K. *Organohalogen Comp* 2006:68;1372.
8. SAS Institute. SAS/STAT User's Guide Version 8. Cary, NC: SAS Institute Inc., 1999.
9. O'Connor R, Sabrsula J. *Environmental Forensics* 2005:6;283.

Table 1: Significant Parameter Estimates and Overall Adjusted R-square Values for PCDF Congener Concentrations in Household Dust**

(Shaded estimates: var. associated with higher levels, Non-shaded estimates: var. associated with lower levels)

	2378 TCDF	**12378 PeCDF**	**23478 PeCDF**	**123478 HxCDD**	**123678 HxCDF**	**123789 HxCDF**	**234678 HxCDF**	**1234678 HpCDF**	**1234789 HpCDF**	**OCDF**
Soil (log ppt)										
Soil dioxin conc. around house				0.286	0.240	0.221	0.284	0.281	0.215	0.281
Soil dioxin conc. from garden					0.033		0.030	0.021	0.055	
Max soil dioxin conc. on property	0.424	0.275	0.197							
Region										
M/S floodplain vs. J/C					-0.217		-0.178	-0.219	-0.272	-0.238
M/S near floodplain vs. J/C	0.216			-0.232	-0.273	-0.195	-0.236	-0.299	-0.300	-0.284
M/S out of floodplain vs. J/C	0.103						-0.066	-0.093	-0.100	-0.121
Midland plume vs. J/C										
Work History										
Worked a job such as waste disposal, wastewater, foundries, etc. (yrs*)								-0.007		-0.008
Worked at jobs with chemicals such as chlorophenol, agent orange, etc. (yrs*)			0.009					0.011		0.013
Lived with a Dow worker (yrs*)				0.009						
Lived with a worker of another chemical co. (yrs*)						-0.017				
Indoor Conditions										
Any cats at the residence		-0.086				-0.093			-0.100	
Children in the house (yrs*)										0.005
Floor Surface										
Age of floor surface (yrs*)	0.006	0.007	0.007	0.007	0.008	0.006	0.007	0.008	0.009	
Some velvet carpet (high pile) vs. hard surface							0.327	0.293		
All carpet not velvet (low pile) vs. hard surface										
Floor replaced after flooding vs. not flooded							-0.149	-0.148	-0.310	
Floor not replaced after flooding vs. not flooded								0.269		
Dust Load (log g/m^2)		-0.167		-0.109	-0.121	-0.152	-0.108	-0.116		-0.168
***Sampling Season*†**										
Sampling in spring vs. fall								-0.140		
Sampling in summer vs. fall		-0.195	-0.196	-0.275	-0.263	-0.324	-0.243	-0.293		-0.324
Remediation										
House remediated for dioxins		0.185			0.163			0.143		
Overall Adjusted R-square (%)	**36.42**	**34.39**	**27.46**	**22.72**	**18.25**	**23.21**	**26.17**	**24.54**	**20.54**	**17.18**

* Years/days in which subjects lived at the sampled residence

** All parameter estimates are significant at alpha=0.05 level

† Spring samples were taken from March-May 2005, summer samples were taken from June-August 2005 and fall samples were taken from September-December 2004