A TOXICOKINETIC MODELING APPROACH TO INVESTIGATE THE PREDICTORS OF SERUM TCDD CONCENTRATION

Hao W[1], Jiang X[1], Wenger Y[1], Chang C-W[1], Gillespie BW[2], Towey T[3], Chen Q[2], Hong B[1], Franzblau A[1], Lepkowski J[4], Adriaens P[5], Demond A[5], Garabrant DH[1], Jolliet O[1]

[1]University of Michigan School of Public Health, Environmental Health Sciences, Ann Arbor, Michigan, USA; [2]University of Michigan School of Public Health, Biostatistics, Ann Arbor, Michigan, USA; [3]Limno-Tech, Ann Arbor, Michigan, USA; [4]Univerisity of Michigan Institute for Social Research , Ann Arbor, Michigan, USA; [5]University of Michigan College of Engineering, Civil and Environmental Engineering, Ann Arbor, Michigan, USA.

Introduction

The University of Michigan Dioxin Exposure Study (UMDES) was carried out to investigate the relationship between historic releases of dibenzo-p-dioxins (PCDDs) and polychlorinated dibenzofurans (PCDFs) by the Dow Chemical Company and serum dioxin concentrations among the population in Midland and Saginaw, Michigan. The principal goals of the study are to determine the potential pathways through which the dioxins may have entered the target population and to differentiate past and present exposures. Because time-related factors are often correlated, approaches that reduce these correlations can be valuable for clarifying the temporal sequence of exposure events. In this article, we develop an approach which enables past and present exposures to be examined in relation to serum dioxin levels, accounting for time-related factors including change of concentrations in food sources in different historic periods, change in elimination half-lives with age, and change in amount of food consumed with age. We apply this approach to pretreat input variables by calculating the toxicokinetic cumulated intakes or exposures over the lifetime of each individual in our study. We perform here a linear regression analysis of serum concentration of 2,3,7,8 tetrachloro-*p*-dibenzodioxin (TCDD) as a function of these toxicokinetic-transformed variables and additional demographic variables to determine the regional predictors that contribute significantly to serum concentrations.

Materials and methods

The main idea is to convert the historical dioxin intake into a 2005 lifetime cumulated dioxin intake, then regress the blood measurements on these pretreated dioxin exposure predictors and demographic predictors to identify the factors that predict the serum. As discussed by Jolliet et al.[1,2], the formula to calculate the lifetime cumulated variables is shown below: For a given individual *i* and dioxin congener *j*, the 2005 cumulated intake through exposure pathway *k* is expressed as sum of quantity consumed (or other information for nonfood exposure variables) over years till 2005 adjusted by 3 correction factors.

$$Q^{i,2005}_{\text{lifetime cumulated } k,j} = \sum_{yr=\text{birth year}}^{2005} Q^{i,2005}_{\text{lifetime cumulated } k,j}(yr)$$
$$= \sum_{yr=\text{birth year}}^{2005} Q_k^{i,2005} \cdot CF^{\,i}_{food,k}(age(yr)) \cdot CF_{conc,j}(yr) \cdot CF^{\,i}_{decay,j}(age(yr)) \cdot Exposed(yr) \quad (1)$$

CF_{food} : correction factor for the changes in the amount of food consumed at different ages;

CF_{conc} : correction factor for the relative change in congener concentration in the food with calendar year;

CF_{decay} : correction factor for the metabolic decay in the body between year of intake *yr* and year of blood measurement;

Exposed: A 0/1 indicator describing whether or not if the participant was exposed to a potential dioxin pathway for each year from birth to 2005.

For fish intake variables, we combine the fish meals consumed with the TCDD concentration in the fish to obtain the dioxin intake from fish in ng/year, then cumulate it over the lifetime to get the actualized 2005 cumulated intake in ng. The game intakes variables are created similarly. For other food variables, $Q^{i,2005}_{lifetime\ cumulated\ k,j}$ is maintained in units of the number of meals because relevant concentration data to generate the intake variables in ng are not available.

A linear stepwise regression is carried out for blood (lipid adjusted) concentrations as a function of the lipid adjusted toxicokinetic cumulated exposure variables (intake for fish and game) over the participant's lifetime. We also let additional demographic predictors enter the model to control for effects that were not fully adjusted in the correction factors of equation (1):

$$C_{blood\ lipid\ adjusted,\ j} = b_0 + \sum_k b_k \frac{Q^{i,2005}_{lifetime\ cumulated\ k,j}}{W_{lipid}} + \sum_l d_l \cdot demographic\ predictors \qquad (2)$$

where $k = 1, \cdots, 48; l = 1, \cdots, 14; j = 1, \cdots, 29; i = 1, \cdots, 946;$

The regression analyses are tested using 2 main approaches: with clustering and strata, without survey weights (unweighted model) and with survey weights (weighted model). Both models have advantages and limitations: the unweighted model provides minimum variance on the parameter estimates in the contaminated region of interest because of oversampling, such as the effect of consuming fish from the Tittabawassee River (TR), and the Saginaw River/Bay (SR/SB); but it does not allow inferences on the overall population to be made. In contrast, the population-weighted model is representative of the population of the Midland/Saginaw and the referent area of Jackson/Calhoun counties and allow inferences to be made about these populations, but may result in wider (or narrower) confidence limits on parameter estimates.

We use forward stepwise variable selection procedure to select predictors of serum concentrations from pretreated exposure variables and demographic variables. Then we test the significance of a meaningful interaction term if both of the main effect demographic variables enter the model (e.g., interaction term between age and sex). In the third step, we trim down the model by removing variables that have implausible negative coefficients (such as intake variables that are inversely associated with serum TCDD concentrations) in the order of the descending *p*-values. Then we allow each variable that is not in the previous model to enter and examine whether the resulting model has a better overall fit to the data, using the AIC and overall model adjusted R^2 as criteria for evaluating goodness of fit. If a significant predictor is identified from the previous non-selected variables, it is then added to the preliminary final model. This step is repeated until no more predictors can be added and the model obtained is the final model. To assess the stability of the model and the effect each individual could have on the predictor, the dfbeta statistics are computed for each observation for each factor in the final model, and those factors that cease to be statistically significant ($p>0.05$) after the removal of 3 or fewer observations are defined as unstable factors.

Results and discussion

Table 1 presents parameter estimates and their *p*-values, and the adjusted R^2 for both unweighted and population-weighted models. The coefficient of lipid adjusted intake through lifetime fish consumption from the Tittabawassee River and the Saginaw River/Bay in the unweighted model is 0.43 which corresponds to the fraction of the 2005 lipid adjusted cumulated intake (in ppt) that is absorbed after cooking and ingestion. Note that lipid adjusted lifetime meat consumed from the contaminated area in both models is labeled as an unstable factor, however, we are cautious about this conclusion as there are only 25 participants in the study who have non-zero values for lipid adjusted lifetime meat consumption from the contaminated region. The adjusted R^2 for the unweighted and weighted models are respective 45.1% and 55.8%, meaning that both models explain well the variation in serum concentration. These two adjusted R^2 are not directly comparable and most of their difference can be explained as follows: The observation with high residual tends to have lower weight than average. Thus the weighted R^2 is higher than the unweighted one.

Table 1: Unweighted and weighted linear regression model results summary for serum TCDD

Model	Unweighted (adjusted R^2: 45.1%)		Weighted (adjusted R^2: 55.8%)	
Factor	Estimate	*p*-Value	Estimate	*p*-Value
Residence factors[a]				
No. of years lived in Midland/Saginaw in 1940-1983	43.9	<0.0001	40.6	0.0007
No. of years of air concentration in 1940-1983	5120	0.02	4170	0.0003
Fish, meat and root vegetables consumption[a]				
Fish intake from Tittabawassee River, Saginaw River or Saginaw Bay	0.43	0.03		
No. of meals of meat from Tittabawassee River, Saginaw River or Saginaw Bay	0.03[b]	<0.0001	0.03[b]	<0.0001
No. of meals of root vegetables property			0.001[b]	0.01
Work history[a]				
No. of years worked in paper industry	1.32	0.03		
Demographic factors				
Age at interview-50	0.01	0.62[c]	0.02	0.03
Sex (1 for female; 0 for male)	1.36	<0.0001	1.17	<0.0001
Sex ×(age at interview-50)	0.11	<0.0001	0.07	<0.0001
Lifetime smoking status(1 for ever; 0 for never)	-0.48	0.01		
Total no. of months all children were breast-fed			-0.02	0.01
Intercept	1.42	<0.0001	1.29	<0.0001

[a]Residential factors/Fish, meat and root vegetables consumption/Work history are all toxicokinetic transformed factors.

[b]Unstable factors which lose significance after the removal of up to 3 observations.

[c]Age centered at 50 was included since the interaction term with sex is significant.

To quantify the influence of individual factors on the serum concentration, we calculate the predicted effect for each participant by multiplying the participant's value of each factor by its parameter estimate. For each observation, this is plotted as a vertical bar that expresses the amount by which the serum concentration is increased by that factor. These are then plotted from lowest to highest. The following figures show the contribution of 4 regional predictors that reflect potential exposure pathways related to Midland/Saginaw contamination to the serum concentration TCDD based on the unweighted model. The contribution to serum TCDD from living in Midland/Saginaw in 1940-1983 ranges from 0 to a maximum of 10 ppt, exposure to TCDD in air from the Dow incinerator in 1940-1983 ranges from 0 to 10 ppt, and lifetime consumption of fish from the contaminated area ranges from 0 to 7 ppt, and lifetime consumption of meat from the contaminated area ranges from 0 to 8 ppt. The number of participants affected by each factor varies substantially: living in Midland or Saginaw in 1940-1983 has an estimated effect on over 500 of the 946 participants, whereas lifetime consumption of meat from the contaminated area has an estimated effect on 25 people, of whom 3 are expected to have an increase in serum TCDD of > 2 ppt. It should be noted that these models do not differentiate yet between current (ongoing) consumption of fish and meat from the contaminated areas and historic consumption earlier in the participant's lifetime. Analyses to differentiate between historic and current exposures are underway.

Figure 5 presents the predicted serum TCDD concentration as sum of contributions of all predictors that are statistically significant in the unweighted model, ranked by the total effect of all factors combined. The combined effect of all factors relevant to the contaminated area results in up to 18 ppt increase in serum concentration. For most observations, living in Midland/Saginaw in 1940-1983 and living at high modeled air concentration from the Dow incinerator in 1940-1983 dominate. The toxicokinetic based household dust and soil factors show no significant associations with the serum TCDD, indicating that they may not be important determinants of serum TCDD concentrations as long as food pathways are avoided.

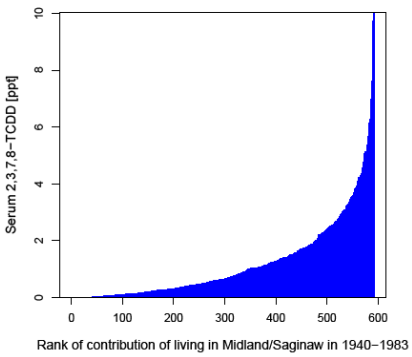

Figure 1. Contribution of living in Midland/ Saginaw in 1940-1983 to serum TCDD

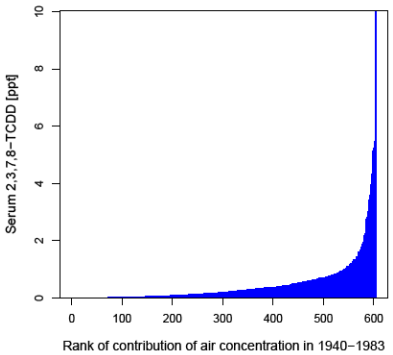

Figure 2. Contribution of air concentration in 1940-1983 to serum TCDD

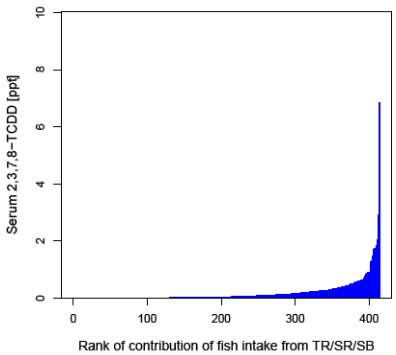

Figure 3. Contribution of fish consumption from contaminated area to serum TCDD

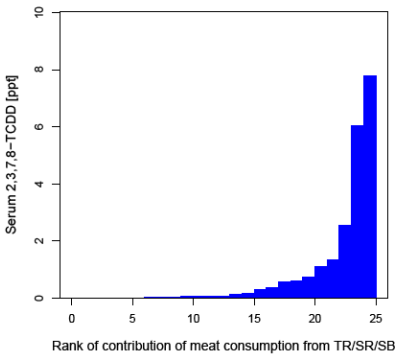

Figure 4. Contribution of meat consumption from contaminated area to serum TCDD

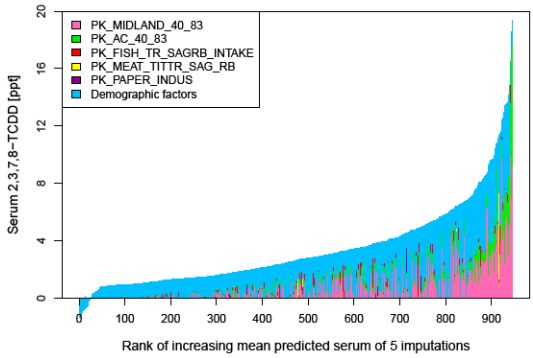

Figure 5. Contribution of demographic and toxicokinetic adjusted exposure factors to serum TCDD

Both unweighted and weighted models explain well the variation in serum concentration TCDD. The population graphs (Figures 1-5) are a powerful representation which can also include the population weights, although they have not been presented here. For the consumption of fish and meat from the contaminated area, the additive and explicit nature of the toxicokinetic modeling approach makes it plausible to assess further the consumption by year and differentiate between past exposure that cannot be changed and present exposure that could be mitigated.

Acknowledgements

Financial support for this study comes from the Dow Chemical Company through an unrestricted grant to the University of Michigan. The authors acknowledge Drs. L. Birnbaum, R. Hites, P. Boffetta, and M. H. Sweeney for their guidance as members of our Scientific Advisory Board, and S. Vantine for her assistance.

References

1. Jolliet O, Wenger Y, Milbrath MO, Chang C-W, Chen Q, Franzblau A, Garabrant D, Towey T, Adriaens P, Gillespie BW. (2007); *Organohalogen Compounds*. 69: 226-229

2. Gillespie BW, Jolliet O, Chang C-W, Milbrath MO, Wenger Y, Chen Q, Towey T, Lee S-Y, Hong B, Garabrant D, Franzblau A Hedgeman E, Knutson K, Adriaens P, Demond A, Trinh H, Lepkowski J, Olson K, Luksemburg W. (2007); *Organohalogen Compounds.* 69: 2260-2263

3. Garabrant D, Franzblau A, Lepkowski J, Gillespie BW, Adriaens P, Demond A, Hedgeman E, Knutson K, Zwica L, Olson K, Towey T, Chen Q, Hong B, Chang C-W, Lee S-Y, Ward B, LaDronka K, Luksemburg W, Maier M. (2009); *Environ Health Perspect*. 117 : 818-824

4. Milbrath MO, Wenger Y, Chang C-W, Emond C, Garabrant D, Gillespie BW, Jolliet O. (2009): *Environ Health Perspect.* 117:417-425