# PREDICTORS OF 8 FURAN CONGENERS IN BACKGROUND U.S. POPULATIONS: DATA FROM TWO MICHIGAN COUNTIES AND THE U.S. NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY (NHANES)

Gillespie BW[1,2], Reichert H[2], Chang C-W[3], Hedgeman E[3], Hong B[3], Chen Q[1], Jolliet O[3], Knutson K[3], Lee S-Y[1], Lepkowski J[4], Olsen K[3], Adriaens P[5], Demond A[5], Towey T[5], Ward B[4], Luksemburg W[6], Maier M[6], Franzblau A[3], and Garabrant D[3]

[1]Department of Biostatistics, University of Michigan School of Public Health, 109 S Observatory, Ann Arbor, MI 48109 USA; [2]Center for Statistical Consultation and Research, University of Michigan, 3550 Rackham Building, 915 E. Washington St., Ann Arbor, MI 48109; [3]Environmental Health Sciences, University of Michigan School of Public Health, 109 S Observatory, Ann Arbor, MI 48109 USA; [4]Institute for Social Research, 426 Thompson St., University of Michigan, Ann Arbor, MI 48104 USA; [5]Department of Civil and Environmental Engineering, University of Michigan College of Engineering, 1351 Beal, Ann Arbor, MI 48109; [6]Vista Analytical Laboratory, Inc., 1100 Windfield Way, El Dorado Hills, CA 95762 USA;.

**Keywords**: Human samples, Serum, Environmental

**Introduction:** Serum furan concentrations in populations exposed to industrial or other sources of pollution are often compared to serum concentrations in populations exposed only to background levels. However, data from appropriate comparison populations are not always available. Modeling background furan concentration as a function of demographic and other variables can allow prediction of background levels for an individual with specific characteristics. Such models could also elucidate the routes of furan exposure and elimination.

This investigation modeled serum concentrations for 8 furan congeners (TCDF, 12378 PeCDF, 23478 PeCDF, 123478 HxCDF, 123678 HxCDF, 234678 HxCDF, 1234678HpCDF, and OcCDF) in populations exposed only to background levels of exposure. Two additional congeners (123789 HxCDF and 1234789 HpCDF) were measured, but had over 95% of values below the limit of detection (LOD) in both UMDES and NHANES. Two data sources were used: (1) the University of Michigan Dioxin Exposure Study (UMDES), and (2) the National Health and Nutrition Examination Survey (NHANES). The UMDES collected serum samples from subjects living in Michigan, USA in areas potentially exposed to furan-like compounds as well as areas presumably exposed only to background levels of these compounds. The NHANES collected demographic and health data from a sample of non-institutionalized residents of the U.S., and measured serum furan levels in a subset of these subjects. This study presents prediction models for serum levels of eight furan congeners in both the UMDES background region and the NHANES study.

**Materials and Methods:** The UMDES was carried out in Michigan, USA, in Midland, Saginaw and parts of Bay County (all potentially exposed areas) and in Jackson and Calhoun Counties (control area, and the focus of this analysis). A two-stage probability household sampling design was used.[1] Eligible subjects were at least 18 years of age, lived in their current residence for at least 5 years, and provided written informed consent. A detailed exposure questionnaire asked several hundred questions including demographics, smoking history, pregnancy history, occupational exposure, food consumption, and other questions possibly related to human body burden. Serum samples were collected in 2005-2006 from subjects who consented and were medically eligible to give blood as defined by the American Red Cross. House dust, soil, and vegetation samples (not reported here) were collected from the homes and property of consenting subjects who owned their homes and/or properties. Chemical analyses were performed by Vista Analytical Laboratory, Inc. (El Dorado Hills, California, USA) for the World Health Organization designated 29 PCDD, PCDF, and dioxin-like PCB congeners using US Environmental Protection Agency (EPA) methods 8290 and 1668.[2,3] Lipid-adjusted concentrations were reported; values below the limit of detection (LOD) were identified, and the LOD was given in each case.

NHANES data are available as a downloadable database of national health and vitality information for a sample of the United States population[4]. The NHANES data for this study were taken from the 2001-2002 sample

release. All individuals answered a general questionnaire covering health, diet and social-demographics; medically eligible persons were asked to donate a blood specimen, and a sub-sample of 1228 persons aged 20 years and older was selected for additional analysis of serum dioxins and furans. Chemical analyses and lipid measurements were performed by the Centers for Disease Control and Prevention laboratories. Lipid-adjusted concentrations were given; results below the limit of detection (LOD) were identified, and reported as LOD/√2.

**Statistical Methods**: The UMDES and NHANES data were analyzed separately. For each, the $\log_{10}$ serum furan level was the outcome variable in a likelihood-based regression for left-censored (below LOD) data, assuming a lognormal distribution. Models for congeners >99% below LOD are not shown. For Figure 1, predicted values were based on models that included only age, age$^2$, sex, race, and interactions as significant. For Table 2, potential predictors included demographic (age, sex, race, BMI, and BMI loss or gain in the past year), and lifestyle (pack-years of smoking, breastfeeding, numbers of complete and incomplete pregnancies). Selected two-way interactions were tested. Backward selection was used to remove non-significant variables. The regression models used survey weights to adjust for the probabilities of sample selection and non-response, and allow inference to the entire two-county population (UMDES) or the U.S. population (NHANES). Multiple imputation was used to handle missing data in UMDES. We used the generalized $R^2$ to estimate explained variation. We used the same covariates in the UMDES and NHANES models, with three exceptions: First, although the NHANES sample was designed with substantial racial diversity, the UMDES sample was over 90% Caucasian. Therefore, race effects were tested in the NHANES models, but not in the UMDES models. Second, the UMDES collected the months of breastfeeding for each child (which was summarized as number of months breastfed after 1980), and NHANES collected the number of children breastfed. Third, the UMDES collected data on lifetime consumption of specific foods, but NHANES collected data on past-24-hour food consumption for a wider variety of specific foods. (Food results reported separately.) SAS® version 9.1 (SAS Institute Inc., Cary, NC, USA, 2007) was used for data management, and the Stata version 9.2 intreg procedure (StataCorp., College Station, Texas, USA, 2007) was used for left-censored regression.

**Results and Discussion:** The UMDES interviewed 1324 subjects, 946 of whom provided serum samples, 251 from Jackson or Calhoun Counties. NHANES had ~1158 subjects with serum dioxin measures, with slight variation in sample size by congener. Characteristics of subjects from the two groups are given in Table 1 below. The two groups are reasonably similar with respect to demographic characteristics.

**Table 1. Characteristics of subjects from two background population groups (survey weighted).**

| | **UMDES (n=251):** Two Michigan Counties | **NHANES (n=1158):** U.S. Population |
|---|---|---|
| | ***Mean (s.e.) or %*** | ***Mean (s.e.) or %*** |
| Age (years) | 49.9 (1.3) | 46.3 (.76) |
| Sex: Female | 61.9% | 51.6% |
| Body Mass Index (BMI) | 28.7 (.52) | 28.0 (.27) |
| BMI loss in past year | 1.1 (.13) | 0.7 (.06) |
| Smoking (packyears) | 12.6 (1.4) | 10.1 (.81) |
| Breastfeeding (% of women) | 16.8% | 23.4% |
| # Incomplete Pregnancies | 0.5 (.13) | 0.3 (.04) |
| Race: Caucasian (non-Hispanic) | 93.6% | 73.2% |
| Black (non-Hispanic) | 4.8% | 10.2% |
| Hispanic (UM) / Mexican American (NH) | ~~.?% or~~ 1.2% | 6.9% |

The percent below the LOD and the median LOD are given in Table 2 for each congener and study. The percents below LOD ranged from 1% to 95% in UMDES, and 11% to 89% in NHANES. The median LODs ranged from 0.4 to 2.7 ppt in UMDES, and from 2.0 to 2.4 ppt in NHANES. Table 2 below shows the final models for each congener and group. Both models show effects of age, sex, race (NHANES only), BMI, BMI loss and/or gain, smoking and breastfeeding. Due to lower LODs, the UMDES model had more statistical power to detect interactions and other effects.

**Table 2. Congener concentration percentiles, half-lives, and regression results predicting furan levels from two background populations: UMDES (UM; two MI counties), and the National Health and Nutrition Examination Survey (NHANES, NH). (Not shown if >99% below LOD.)**

| | TCDF | 1,2,3,7,8 PeCDF | 2,3,4,7,8 PeCDF | | 1,2,3,4,7,8 HxCDF | | 1,2,3,6,7,8 HxCDF | | 2,3,4,6,7,8 HxCDF | | 1,2,3,4,6,7,8 HpCDF | | OcCDF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Median (75th, 90th)** | 0.0 (0.6, 1.3) | 0.0 (0.0, 1.0) | 5.5 (7.7, 11.6) | 5.3 (10, 15.8) | 5.5 (7.5, 10.4) | 5.4 (8.2, 12.9) | 5.6 (7.4, 10.4) | 4.5 (7.2, 11.2) | 0.0 (.8, 2.0) | 2.1 (2.7, 3.4) | 7.0 (9.7, 13.5) | 10.1 (14.5, 21.4) | 0.0 (0.0, 0.0) |
| **% below LOD (median LOD)** | 66.5% (.4) | 75.7% (.4) | 1.2% (2.5) | 34.5% (2.0) | 5.2% (2.7) | 18.0% (2.4) | 4.8% (2.0) | 36.7% (2.2) | 74.5% (.9) | 89.2% (2.0) | 9.2% (1.8) | 10.5% (2.4) | 94.8% (2.6) |
| **Half-life in human body** | 2.1 yrs | 3.5 yrs | 7.0 yrs | | 6.4 yrs | | 7.2 yrs | | 2.8 yrs | | 3.1 yrs | | 1.4 yrs |
| | **UM Beta (p-val)** | **UM Beta (p-val)** | **UM Beta (p-val)** | **NH Beta (p-val)** | **UM Beta (p-val)** | **NH Beta (p-val)** | **UM Beta (p-val)** | **NH Beta (p-val)** | **UM Beta (p-val)** | **NH Beta (p-val)** | **UM Beta (p-val)** | **NH Beta (p-val)** | **UM Beta (p-val)** |
| **Age** [a] | .005 (.272) | .005 (.144) | .012 (<.001) | .014 (<.001) | .009 (<.001) | .007 (<.001) | .009 (<.001) | .008 (<.001) | .007 (.072) | .006 (<.001) | -.001 (.285) | .0009 (.239) | -.0003 (.956) |
| **Age$^2$** [a] | | -.0004 (.028) | -.0001 (.001) | -.0001 (.072) | | | | | | | | | |
| **Sex (female)** | -.164 (.024) | -.269 (.006) | -.003 (.884) | .044 (.202) | .003 (.913) | -.012 (.564) | .002 (.943) | .015 (.583) | -.029 (.825) | -.104 (.006) | .067 (.104) | -.029 (.183) | .155 (.482) |
| **BMI** [a] | | .014 (.031) | | -.005 (.033) | .006 (<.001) | .004 (.003) | .006 (.015) | | | | | | |
| **BMI loss (past yr)** | | | .014 (<.001) | | .014 (.005) | | .019 (<.001) | .003 (.680) | | .034 (.029) | | | |
| **BMI gain (past yr)** | -.102 (<.001) | | -.022 (.004) | | | -.013 (.005) | | -.018 (.006) | | | | -.021 (.003) | |
| **Children** | | | | -.039 (.005) | | | | | | | | | |
| **Breastfeeding** [b] | | | | | | | | -.019 (<.001) | | | | | |
| **Pack Years** | | -.009 (.003) | | | | | | | -.022 (<.001) | -.006 (<.001) | | | |
| **Race** [c] **(Mex. American) (NHANES only)** | | | | -.146 (.011) | | -.119 (<.001) | | -.146 (<.001) | | -.057 (.435) | | | |
| **Race** [c] **(Non-Hispanic Black) (NHANES only)** | | | | | | .048 (.045) | | .140 (.004) | | | | .134 (<.001) | |
| **Age** [a]***Race** [c] **(Non-Hispanic Black (NHANES)** | | | | | | .004 (.001) | | .004 (.001) | | | | | |
| **Sex** [a]***Race** [c] **(Mex. American) (NHANES)** | | | | | | | | | | -.290 (.007) | | | |
| **Sex*BMI loss** | | | | | | | | .023 (.021) | | | | | |

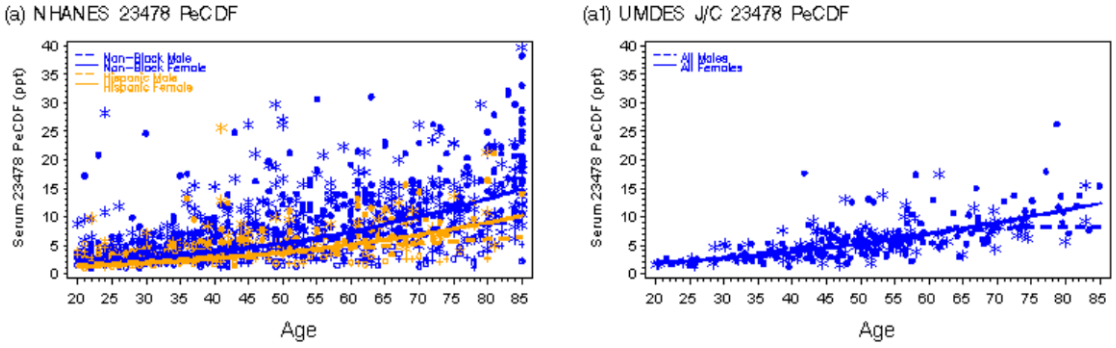
[a] Age and BMI are centered around the mean values (UMDES Mean Age=52, Mean BMI=29; NHANES Mean Age=38, Mean BMI=27).
[b] Breastfeeding definition varies by population (UMDES Total months breastfeeding after 1980; NHANES Number of children breastfed at least 1 month).
[c] The reference group for race is Caucasian. [d]Furans not shown were not modeled due to the high proportion of censored data.

Higher serum concentrations were significantly associated with older age for the 3 congeners with the longest half-lives (6.4 to 7.2 years), and one additional congener (half-life 2.8 years). Gender effects were weaker than those seen for dioxins, and in the opposite direction. Women had significantly lower levels of furans for only 3 congeners (half-lives of 2.1 to 3.5 years). Higher BMI and recent weight loss were associated with higher serum concentrations for 3 and 4 congeners, respectively, with most significant effects among congeners with longer half-lives. Weight gain was associated with lower serum concentrations for 5 congeners. Children and breastfeeding were each associated with lower serum concentrations for a single congener (half-lives 7.0 and 7.2, respectively). Smoking was associated with lower serum concentrations for two furans (half-lives 2.8 and 3.5). Hispanics had lower serum concentrations for 4 congeners, and blacks had higher serum concentrations for 3 congeners, with race effects seen mostly for congeners with longer half-lives.

Figure 1 (below) shows raw data with predicted mean values by age, gender and race (NHANES only) for 23478 PeCDF. (Other congeners not given due to space limitations.) In general, estimates are similar in UMDES and NHANES; race effects in NHANES are mainly seen at older ages.

A strength of both studies is the population-based sampling, allowing valid inference back to the source populations. The low LOD in UMDES allowed more detailed modeling. Limitations include the high LOD values in NHANES 2001-2002 data. Future steps include modeling of food variables, and analysis of the newly released 2003-2004 NHANES data (released May 2007).

**Figure 1**. Plots of serum congener concentrations of 23478 PeCDF and model predicted means by age, gender and race (NHANES only) for both (a) UMDES and (b) NHANES for TCDD. For values below the LOD, the LOD itself is plotted (grey circles).




**Acknowledgements:** 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. Garabrant DH, Franzblau A, Lepkowski J, Gillespie BW, Adriaens P, Demond A, Ward B, LaDronka K, Hedgeman E, Knutson K, Zwica, Olson K, Towey T, Chen Q, Hong B. The University of Michigan Dioxin Exposure Study: Methods for an environmental exposure study of polychlorinated dioxins, furans and biphenyls. (In preparation)
2. 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, 19997.
3. United States Environmental Protection Agency. 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.
4. National Health and Nutrition Examination Survey (NHANES) website (accessed 5/10/2008): http://www.cdc.gov/nchs/nhanes.htm.