Analysis of Dietary Intake of Selected Metals in the NHEXAS–Maryland Investigation

P. Barry Ryan,[1] Kelly A. Scanlon,[1] and David L. MacIntosh[2]

[1]Department of Environmental and Occupational Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA; [2]Department of Environmental Health Science, College of Agricultural and Environmental Sciences, University of Georgia, Athens, Georgia, USA

As part of a large pilot investigation of multimedia exposure to several classes of environmental contaminants, the National Human Exposure Assessment Survey (NHEXAS)–Maryland study, we collected 388 semiquantitative food checklists and duplicate diet solid food samples, analyzed for arsenic, cadmium, chromium, and lead concentrations, from 80 individuals in Maryland in 1995–1996 in a repeated measures design. Here we explore several methods to infer foods most strongly associated with concentrations of these metals observed in the duplicate diet in our data set. We employed two techniques in which logarithmically transformed metal concentrations in the duplicate diet were regressed on individual food item consumption using algorithms designed to identify the foods most associated with the observed duplicate diet concentrations. We also employed an alternative strategy in which foods to be used as independent variables in regression were selected using data collected in national food consumption and residue surveys, with regression procedures proceeding with the selected foods in a similar manner. The concordance of foods selected as major predictors among these three techniques is noteworthy and is discussed. Finally, the Dietary Exposure Potential Model (DEPM) was used with the Dietary Checklist data to predict duplicate diet concentrations within our sample. A comparison between the predicted values and those observed gave R^2 values of 0.180, 0.206, and 0.076 for As, Cd, and Pb, respectively ($p < 0.0001$ in all cases). We discuss the significance of these observations and the implications for dietary-exposure–based risk analysis and dietary intake epidemiology. *Key words*: arsenic, cadmium, chromium, chronic exposure, dietary exposure, lead, NHEXAS–Maryland. *Environ Health Perspect* 109:121–128 (2001). [Online 11 January 2001]
http://ehpnet1.niehs.nih.gov/docs/2001/109p121-128ryan/abstract.html

Many metals have received attention as both environmental contaminants and potential toxicologic hazards. Important among these are arsenic, cadmium, chromium, and lead due to their "relative uptake, accumulation, and toxicity to humans" (*1*). These four elements are naturally occurring and widely distributed in the environment (*2–5*). One may be exposed by ingesting foods containing these elements in addition to other environmental sources such as contaminated drinking water or airborne particulate matter.

The variety of the human diet and the frequency of consumption make food intake a potential major environmental influence on health (*6,7*). Food may be contaminated by direct uptake of contaminants by plants and animals or it may be contaminated incidentally as a result of growing, harvesting, processing, and distribution (*8*).

Traces of both inorganic and organic forms of As are found in many foods (*9*), but the inorganic form has been associated with human cancer and other adverse health effects. The highest concentrations of As found in food exist in several types of seafood as an organic form known as arsenobetaine or "fish arsenic" (*2*). Although this type of As is considered relatively nontoxic to humans, it is difficult to determine by means of simple food analyses what fraction of As in seafood takes this form.

Almost all foods have inherently low levels of Cd, with the highest levels occurring in grain products, leafy vegetables, potatoes and other root vegetables, organ meats, and shellfish (*3*). Similarly, Pb has been measured in a variety of foods, with the highest levels in vegetables, fruits, grains, meat, and seafood (*5*). Of special concern is the potential introduction of Cd and Pb compounds into the food web via fertilizers containing industrial wastes and municipal sewage.

Cr exists in several oxidation states, two of which are of practical public-health importance: trivalent Cr and hexavalent Cr (*10*). Total Cr concentrations in foods are generally low, with the highest levels found in acidic fruits and vegetables, meats, and seafood (*4*). The trivalent form is a nutrient essential to humans. In general, the hexavalent Cr compounds are more toxic to humans than the trivalent and are known carcinogens. However, laboratory analyses typically performed on dietary samples do not yield information on the type or amount of each Cr species present in foods.

To improve understanding of the role of dietary intake in total exposure to various chemical species, the Dietary Exposure Potential Model (DEPM) was developed by Technical Assessment Systems (TAS), Incorporated (*11*), for the U.S. Environmental Protection Agency's National Exposure Research Laboratory (U.S. EPA/NERL, Cincinnati, OH). The intent was to provide a modeling system that could estimate the dietary exposure to chemical contaminants by integrating dietary data and contaminant residue information (*11*). Version 2.4 includes consumption information from the Nationwide Food Consumption Survey (NFCS) for the 1977–1978 study and 1987–1988 study as well as residue data from the U.S. Food and Drug Administration's (FDA) Total Diet Study (TDS) 1986–1995.

Here we examine potential human exposure to As, Cd, Cr, and Pb from ingestion of food for participants in the National Human Exposure Assessment Survey (NHEXAS) pilot study in Maryland. All participants were informed of their rights as volunteers in the investigation before participation. An Institutional Review Board oversaw this research. Analysis of time-of-year food consumption as well as temporal and population variability of metal concentrations in food and dietary exposures has been presented elsewhere (*12*). Our goal was to identify foods that contribute the most exposure to metals through diet. We compare analyses of the dietary checklist and metal concentrations in food using alternative statistical tools. In addition, we compare measurements of dietary exposure to As, Cd, and Pb assessed using a duplicate diet approach to modeled estimates obtained from a market-basket approach and the DEPM.

Methods

We collected 388 observations of average daily food consumption and metal concentrations in composite samples of duplicate portions of solid food from a stratified random

Address correspondence to P.B. Ryan, Department of Environmental and Occupational Health, Rollins School of Public Health, Emory University, 1518 Clifton Road, Atlanta, GA 30322 USA. Telephone: (404) 727-3826. Fax: (404) 727-8744. E-mail: bryan@sph.emory.edu

This research was supported by the U.S. Environmental Protection Agency under cooperative agreement CR822038-1, the U.S. Department of Agriculture Hatch Project GEO00843, and the University of Georgia Research Foundation.

Received 18 July 2000; accepted 22 August 2000.

sample drawn from the state of Maryland (*13*). Details of the data collection methods are described elsewhere (*12*), so we discuss here only the information most relevant to the current analysis.

We selected an initial sample of 80 individuals. We collected information from this group 6 times over the course of 1 yr, from September 1995 through September 1996. We assessed food consumption using a food checklist modeled after the Willett food frequency questionnaire (*14,15*). The checklist contained 131 solid food items. Participants completed the checklist on four consecutive days during each of 6 monitoring periods. The average daily consumption rate for each food as determined from the 4 days of data is the basis for this analysis.

Additionally, we asked participants to prepare duplicate portions of each food item consumed on the same 4 days during which the checklist was completed. Duplicate portions of solid foods were collected and analyzed separately from duplicate beverage portions to minimize dilution effects on solid food residues by beverages. Duplicate portions were aggregated over the sampling period, yielding a single, 4-day composite sample. The composite samples were homogenized and analyzed for As, Cd, Cr, and Pb by inductively coupled plasma-mass spectrometry (ICP-MS). Analyses presented here focus on the solid food portions only. For As, Cd, and Pb, all measured food concentrations exceeded the limit of detection (*12*). For Cr, 52 samples contained concentrations of this metal below the limit of detection. These 52 observations were not included in the following analyses.

Scanlon et al. (*12*) provide details on the quality assurance program associated with both the dietary checklist data and the duplicate diet food samples. Briefly, chain-of-custody forms followed each type of sample from the field, through laboratory analysis, to final data sets. Of 404 possible checklists, five were assessed as invalid due to missing chains of custody or missing physical checklists. Similarly, 398 out of 403 possible duplicate-diet samples had valid metal residue concentrations as analyzed by the U.S. FDA using protocols described elsewhere (*16*). Limits of detection were determined for each sampling period and averaged 1.5, 1.2, 24.7, and 1.2 mg/kg for As, Cd, Cr, and Pb, respectively. Recovery fractions from spiked samples averaged 95.5%, 96.3%, 97.8%, 100.3% for the four metals over the course of the study. We used no field blanks in this investigation.

Major Contributors to Dietary Exposure

Our principal aim in this work was to identify foods most strongly associated with the metal concentrations measured in the duplicate diet samples. We denote such foods as "major contributors" to dietary exposure to each metal. Several methods can address this issue.

Empirical weights approaches. We performed a series of regression analyses to ascertain the relationship between the solid foods consumed as reported on the food checklists and the concentrations of metals in the duplicate solid food samples. The food consumption values for each food item were treated as the independent variables while the log-transformed (*12*) As, Cd, Cr, and Pb concentration served as the dependent variable. We used three different approaches to characterize the relationship between these variables and to determine the strength of their association: the bivariate, the stepwise-multivariate, and the contribution approaches (*1,17*).

Bivariate approach. Log-transformed metal concentrations were first regressed on the number of servings consumed for each of the 131 solid foods in a bivariate manner. The *p*-value for each of the 131 regressions was obtained, and foods for which the *p*-value was < 0.2 were retained for the next step. The retained foods were then included in a multivariate regression. The *p*-values for the individual foods within the multiple regression analysis were determined. Those maintaining the $p < 0.2$ criterion were retained for the final analysis. The final step was a multiple regression that used only the food items meeting all of the inclusion criteria: $p < 0.2$ in the bivariate regression, and $p < 0.2$ in the first multivariate regression. This last step is repeated if the *p*-value for any variable exceeds 0.2 in the multiple regression step after removal of the variable for which $p > 0.2$.

Stepwise-multivariate regression approach. In the second approach, the log-transformed metal concentrations were regressed on the 131 food items using the procedure PROC REG in SAS (*18,19*) with the stepwise selection criterion. The model was implemented using both forward and backward methods simultaneously with both entry and exit levels of significance set at 0.2. No variables were included initially, and variables were added and removed until no further modification of the variable list occurred under the entry and exit significance test.

Contribution approach. The desired outcome is to perform regressions similar to those discussed under the previous two headings, but with foods selected for inclusion in the model based on their expected contribution to the total dietary intake of the chosen contaminant. Estimates of expected contribution were obtained from the Dietary Exposure Potential Model (DEPM) (*11,20*). The DEPM is a modeling system that allows the user to select both a consumption profile and a residue profile to infer population dietary exposure patterns.

We selected the Nationwide Food Consumption Survey (NFCS) from 1977 as the consumption database. The NFCS was organized by the U.S. Department of Agriculture (USDA) and contains a comprehensive listing of core food items consumed by the general population (*21*). In the DEPM, specific age and gender groups, ethnic backgrounds, regions of the United States, and economic status can be selected to restrict the type of output to a focused subpopulation. For this study, NFCS data from the Northeast region of the United States were chosen because of the restricted locale of the NHEXAS–Maryland study (*13*). All age and gender groups, ethnicities, and economic levels were included in the model. To generalize the DEPM model run to the NHEXAS–Maryland foods, we matched the 131 foods in the NHEXAS–Maryland dietary checklist to foods in the NFCS database (*22*). The details of this matching process are given elsewhere (*23*) and can be obtained from the authors upon request. The NFCS foods that were matched with NHEXAS–Maryland checklist foods were used in all DEPM runs.

Metal concentrations measured in food items as part of the TDS were chosen as the source of residue data for the DEPM runs. The TDS database was selected because it represents approximately 10 years of residue information starting in 1986, and contains over 200 foods representing the core foods of the U.S. food supply as determined from the first NFCS (*24*). The TDS database contains information on concentrations in food items for three of the four metals included in the present study: As, Cd, and Pb, but not Cr.

Estimates of dietary exposure to contaminants based on a market-basket approach, like those of the DEPM, can be influenced by the treatment of samples with residue levels below the method detection limit (*20,25*). Fewer than 5% of the values for As, Cd, and Pb were below the level of detection (LOD) in our duplicate diet data set, suggesting that most diets, at least in Maryland, have sufficient amounts of these metals in them to measure. This may not be the case when individual food items are measured. Many items may be so low in concentration of these metals that none can be measured. Substituting one-half LOD for each of these is likely to bias the predicted values even higher. To minimize possible overestimation of residue levels in the DEPM exposure estimates, samples in the TDS database with nondetect concentrations were set to a value of zero.

We used the DEPM to estimate the population-mean dietary exposure to As, Cd, and Pb in the Northeast United States. We sorted food-specific exposure estimates for each metal in rank order and computed their relative contribution to total dietary exposure. Foods estimated to contribute at least 0.01% to total dietary exposure for each metal were identified as potential contributors and retained for inclusion in a multivariate regression model. The log-transformed metal concentrations in the NHEXAS–Maryland duplicate diet samples were regressed against the potential contributors identified by the DEPM. We determined the *p*-values for the individual foods within the multiple regression analysis. Those maintaining the $p < 0.2$ criterion were retained for the final analysis. The final step was a multiple regression that used only the food items meeting all of the inclusion criteria for the contribution approach: > 0.01% contribution to estimated mean dietary exposure for the Northeast United States, and $p < 0.2$ in the first multivariate regression. In this case, further removal of variables after the initial multivariate implementation step was not done, since the initial selection scheme was not based on purely statistical techniques.

Predicted versus observed intake. Exposure to contaminants in food is often assessed using models based on the market-basket approach. The use of such models may be considered an indirect assessment method. The collection of duplicate portions and analysis of contaminants therein is an alternative assessment technique. The duplicate portion approach may be considered a direct assessment method: Contaminants are measured in foods prepared by and collected from members of the population of interest. Comparison of predicted and observed dietary exposure is of interest but rarely possible. We used the DEPM model and metal concentrations measured in the NHEXAS–Maryland duplicate portion samples to determine the degree to which modeled exposures agree with observed exposures.

To calculate the values developed here, we applied two methods. The direct method for determining dietary intake uses the data from the duplicate diet and classifies individuals by the concentration of metal contamination multiplied by the mass of the food sample. The indirect method multiplies the consumption levels from the dietary checklists by residue data from the TDS. The data reported represent the geometric mean of the observed estimates from the duplicate diet samples for the quintiles of the predicted intake.

We modeled dietary exposure for each metal by multiplying the metal residue concentration for each item obtained from the DEPM by the corresponding consumption rates (number of servings consumed) as reported in the NHEXAS–Maryland checklist. The calculated values for average daily dietary intake for each of the three metals (all except Cr) were compared directly with the concentrations determined in the duplicate diet investigation. The ICP-MS analyses measured metal concentrations in the foods consumed by participants (*12*). These values, in mg metal/kg food, were multiplied by the mass of the duplicate samples saved by the participants, in kg, to calculate the actual intake of metal, in mg, from food ingestion. Although the duplicate diet methodology included instructions for each participant to save a portion of each food similar in size to that which they consumed, a degree of error may be introduced because not all foods may have been saved, as indicated by the 10% who reported they were not able to provide a duplicate of every food consumed. In addition, the weight of the samples was recorded only in cycles 2–6 and therefore does not represent the entire study period.

Unique household identification numbers (HIN) identified individuals in the study. The six repeated measurements are denoted by a variable called "cycle" that specifies the sampling time period (*12,13*).

Results

After quality assurance corrections were made to the database (*12*), the exposure database consists of 388 observations (HIN/cycle combinations) and 131 individual food items. Ninety-three percent of the study population participated in both the checklist and duplicate diet portions in four or more of the cycles.

Major Contributors to Dietary Exposure

Empirical weights approaches. Results for all of the regression analyses described above are found in Tables 1–4 for As, Cd, Cr, and Pb, respectively, and are discussed below. In those tables, food items found to be major contributors for one or more of the procedures are listed in the first column. The next two columns list the parameter estimates and *p*-values associated with that food item that were found in the bivariate approaches. The following two columns list similar information for the stepwise approach, and the final two columns for the contribution approach. If no entry appears for a given food item for one or more of the procedures, then there was no significant relationship for that food item by that procedure. Individual food items are ordered by the *p*-value for the

Table 1. Results representing major predictors of As derived from three regression approaches sorted by the stepwise *p*-value.

	Bivariate Adjusted R^2 = 0.3327[a]		Stepwise Adjusted R^2 = 0.4195[a]		Percent contribution Adjusted R^2 = 0.3004[a]	
Food Item	Par est	*p*-Value	Par est	*p*-Value	Par est	*p*-Value
Tuna	2.122	0.0001	2.458	0.0001	2.161	0.0001
Other fish	2.037	0.0001	1.930	0.0001	2.063	0.0001
Shrimp	1.012	0.0035	1.005	0.0031	0.929	0.0084
Potato, baked	–0.440	0.0235	–0.502	0.0087	–0.519	0.0088
Hamburger	–0.464	0.0285	–0.546	0.0088	–0.455	0.0361
Spinach, raw			2.880	0.0089		
Other fish, canned	2.794	0.0199	2.850	0.0155		
Orange			–0.461	0.0179		
White bread	–0.107	0.0390	–0.121	0.0180	–0.142	0.0080
Dark fish	0.719	0.0350	0.725	0.0286	0.773	0.0266
Bran			–0.888	0.0351		
Grapefruit			0.774	0.0392		
Mixed fruit, canned			–1.128	0.0447		
Yogurt			–0.320	0.0466		
Pepper shake			0.108	0.0479		
Oatmeal	–0.645	0.0113	–0.496	0.0491		
Pancake			–0.393	0.0600	–0.382	0.0769
Celery	–0.488	0.0312	–0.417	0.0670		
Oil and vinegar			0.339	0.0705		
Ice cream			0.327	0.0764		
Romaine lettuce			–0.264	0.0874		
Donut	0.202	0.1761	0.225	0.1250		
Brownie			–0.559	0.1289		
Chicken, canned			0.978	0.1682	0.013	0.1449
Spinach, canned			1.695	0.1494		
Cottage cheese	–0.622	0.0084	–0.351	0.1587		
Tomato, fresh			0.217	0.1831		
Lima beans, fresh			–0.506	0.1863		
Yams, canned					0.964	0.1952
Iceberg lettuce					0.171	0.3295

Par est, parameter estimates.
[a]Data from Draper and Smith (*26*).

stepwise-multivariate regression approach, with the smallest *p*-value coming first. We chose this order because the stepwise approach generally produced the most foods as major predictors, and most foods found by the other techniques to be major predictors were also found in the stepwise approach. Foods found by a technique other than stepwise are ordered as they appeared in the specific procedure.

Bivariate approach. In the bivariate regression approach we regressed the log-transformed metal concentrations on the 131 food items. The resulting statistics include parameter estimates and *p*-values. The parameter estimates represent an increase in the log-transformed metal residue for a unit increase (a serving per day) in consumption of the given food item. Note that negative coefficients imply that the log-transformed metal concentration decreases with consumption of the food item.

After the first bivariate regression, we identified 32 food items as having significant associations with As concentrations at $p <$ 0.2. Of these, 13 met the 0.2 criterion in the multivariate regression. We define these 13 foods as the major predictors of As contamination in the duplicate diet samples using this approach. Similarly, we identified 38 food items for Cd, with 21 classified as major predictors; 18 of 32 foods for Cr; and 20 of 31 foods for Pb.

Stepwise-multivariate regression approach. This model allowed inclusion of any of the 131 food items in a standard forward/backward approach with entry and exit significance levels of 0.2. Twenty-eight foods were retained as major predictors of As using this approach, whereas 34 were retained for Cd, 31 for Cr, and 40 for Pb

Contribution approach. Fifty-one food items were identified through DEPM modeling as contributing at least 0.01% to the intake of As. These foods accounted for 99.87% of the total estimated intake levels. Tuna, cereal, and shrimp account for approximately 60%, 14%, and 13% of the total consumption, respectively. Eighty-nine food items were identified as contributing at least 0.01% to the total intake of Cd, accounting for 99.94% of the levels. Pasta, cereal, iceberg lettuce, and baked potatoes were determined to contribute approximately 16%, 11%, 10%, and 10%, respectively, for a combined contribution of almost 50% of Cd intake from these four foods. Of the 131 foods, 104 were identified as contributing at least 0.01% to the total intake of Pb. These 104 foods accounted for 99.97% of the total estimated intake levels. Fresh and canned tomatoes, cereal, and hamburger contributed approximately 38%, 8%, and 7%, respectively, to the total Pb intake.

The food items identified as contributing 0.01% were included in a multivariate regression to determine the major predictors of As, Cd, and Pb intake based on the percent contribution regression approach. Ten of the 51 foods estimated to contribute 0.01% to the As level were found to be major predictors, whereas for Cd and Pb, 25 of the 89 foods and 30 of the 104 foods were retained, respectively (Tables 1, 2, 3, and 4).

Predicted versus observed intake. Because duplicate diet mass was not taken in the first cycle, only 314 HIN/cycle combinations were represented in this method. A series of scatter plots were produced based on the log-transformed intake values. Figure 1 illustrates the relationship between the predicted and observed values for As, Cd, and Pb, respectively. A regression line has been drawn through the points. The R^2 coefficients and the regression equation are listed on each subfigure.

We grouped the 314 observations into quintiles (denoted from lowest quintile, Q1, through the highest, Q5) based on the predicted intake values. We then computed the mean and standard error (SE) for the observed intake within each predicted quintile. These data were plotted with the quintile on the x-axis and the mean ± 2 SE on the y-axis. Again, the Cr observations were not plotted.

Discussion

Empirical Weights Approaches

The various procedures used to identify individual food items as major dietary pathway contributors to total dietary intake display strong concordance, especially among those food items most strongly associated with each metal. Generally, the percent contribution approach and the bivariate approach identify fewer foods as major contributors; the stepwise approach identifies more. In

Table 2. Results representing major predictors of Cd derived from three regression approaches sorted by the stepwise *p*-value.

	Bivariate Adjusted R^2 = 0.2858[a]		Stepwise Adjusted R^2 = 03902[a]		Percent contribution Adjusted R^2 = 0.2453[a]	
Food Item	Par est	*p*-Value	Par est	*p*-Value	Par est	*p*-Value
Spinach, fresh	0.941	0.0001	0.852	0.0001	0.896	0.0001
White rice	–0.167	0.0012	–0.216	0.0001	–0.227	0.0001
Orange	–0.204	0.0056	–0.281	0.0001	–0.241	0.0014
Grapefruit	–0.608	0.0001	–0.606	0.0001		
Oil and vinegar	0.210	0.0024	0.311	0.0001		
Potato chips	0.153	0.0039	0.205	0.0002	0.181	0.0012
Margarine	0.077	0.0047	0.091	0.0007		
French fries	0.248	0.0091	0.291	0.0024	0.241	0.0139
Banana	–0.115	0.0221	–0.144	0.0034	–0.095	0.0645
Potato, baked	0.157	0.0291	0.212	0.0041	0.134	0.0707
Pasta	0.131	0.0057	0.127	0.0072	0.132	0.0075
Tuna	0.193	0.0619	0.270	0.0072	0.294	0.0058
Tomato, canned			–0.268	0.0084		
Cooked cereal	–0.423	0.0257	–0.460	0.0151		
Chocolate bar			0.176	0.0253	0.515	0.0220
Crackers	0.051	0.0153	0.046	0.0268	0.051	0.0190
Beets, canned	–0.929	0.0734	–1.131	0.0271	–1.249	0.0307
Cantaloupe			–0.191	0.0282		
Cookies, commercial	0.044	0.0647	0.050	0.0356	0.055	0.0280
Beef	–0.201	0.0369				
Eggs	–0.144	0.0202	–0.129	0.0380	–0.154	0.0164
Cabbage	–0.193	0.0405	–0.197	0.0388		
Brownie	0.305	0.0272	0.260	0.0566	0.281	0.0521
Brown rice			0.196	0.0614		
Cake, homemade			–0.247	0.0668	–0.228	0.0778
Corn, canned	–0.363	0.0363	–0.296	0.0840	–0.378	0.0349
Hamburger			–0.125	0.1177		
Beans, canned			–0.200	0.1232		
Corn, fresh			0.142	0.1277	0.122	0.2082
White bread			–0.029	0.1291		
Spinach, raw			0.556	0.1329		
Dark fish			0.176	0.1501		
Tofu			–0.768	0.1534		
Biscuit			0.150	0.1848	0.209	0.0771
Pepper	–0.037	0.0784				
Romaine lettuce					0.122	0.0358
Cookies, homemade					–0.095	0.1859
Jam					0.062	0.1762
Raisin					–0.199	0.0098
Liver					0.202	0.3690
Peach, canned					–0.668	0.0174

Par est, parameter estimate.
[a]Data from Draper and Smith (*26*).

each procedure, however, relatively few foods are significantly associated with dietary intake. The results of the study reflect the dietary habits of a subpopulation of Maryland residents and should not be generalized to broader populations. Yet the practicality of both the dietary checklist and duplicate diet methods to estimate exposures should not be overlooked and can be used in different populations. Further, the fact that these diets are not market-basket–based but, indeed, duplicate actual diets should not be missed.

The parameter estimates indicate values that are both positive and negative. The sign of the parameter estimates is an indicator of how the specific food item contributes to the logarithm of the metal concentration in the duplicate diet samples. A positive parameter estimate suggests that the food item likely contains the metal in question, and the magnitude indicates the contribution to the logarithm of the concentration made by a single serving of that food item. The appropriate view of the food items with negative parameter estimates is less straightforward. Negative parameter estimates may be construed as representing substitution of the specific food item for another. Total food intake was relatively constant over the annual period measured (*12*), suggesting that consumption of one food item lessens the likelihood of consumption of another food item. One interpretation of the negative parameter estimates suggests that consumption of that specific food item may have resulted in less consumption of another food item containing the metal under study, thus lowering dietary intake of the metal. An alternative viewpoint—one that should be considered carefully in light of the relatively small sample size and likely autocorrelation of dietary intake within an individual across seasons—is statistical variability in intake coupled with heterogeneity in residue concentration in individual foods. Such difficulties may produce sign alternation among food items as the regression procedure attempts to fit what is little more than statistical "noise" in the data.

All three approaches estimated that seven of the 30 food items were predictors of As from food consumption, including tuna, other types of fish, and shrimp. Of the 40 food items determined to be major predictors of Cd intake levels, 15 were selected by all three approaches. Fresh spinach, white rice, and potato chips displayed the strongest overall associations with Cd levels in our duplicate diet samples. Seventeen of 32 food items were determined by the two approaches to be predictors of Cr intake levels. Some of the important predictors from this group include cereals, brownies, and celery. For Pb, 17 of the 43 total food items, including canned peaches, canned beets, and grapefruit, were determined by all three methods to be predictors.

It is noteworthy that for each metal studied the stepwise approach produced the maximum variance explained, followed by the bivariate approach. In the three cases for which the percent contribution approach was possible, results were comparable for the bivariate approach and the percent contribution approach, with the former explaining somewhat more of the variance than the latter for As and Cd. For Pb, the converse was true. The stepwise approach invariably selected more foods as major predictors and may be expected to describe more of the variance. However, we report adjusted R^2 (*26*) for this statistic, which attempts to account for the decreased degrees of freedom in such models. This observation suggests that the stepwise approach may lead to a fundamentally better method of selecting the major contributors to dietary intake for these metals.

It is of interest to note the food-item concordance between the two "pure" statistical techniques (bivariate and stepwise regressions), which attempt to describe the observed data in an optimum manner, and the percent contribution approach, which uses an external method of selecting appropriate food items. In the latter case, information on food-item residues measured in other studies has been used to select specific foods based on the expected contribution to total dietary intake. In the regression approaches, statistical procedures operating on the data collected in this study are used to select the food items best describing the variability in residue concentrations. That there is a large concordance between foods selected suggests that market-basket surveys, such as those contributing to the DEPM system, and the dietary checklist approach produce similar results with respect to the most highly influential dietary contributors for these food contaminants.

One must take care not to overinterpret these results. The parameter estimates and even the foods selected are unstable at the statistical significance margin. In our analyses, we observed several cases in the multiple regression phase of the bivariate approach and in the stepwise regression approach for which elimination of a variable by the $p > 0.2$ criterion followed by new multiple regression

Table 3. Results representing major predictors of Cr derived from two regression approaches sorted by the stepwise *p*-value.

	Bivariate Adjusted $R^2 = 0.1486^a$		Stepwise Adjusted $R^2 = 0.2647^a$		Percent contribution	
Food Item	Par est	*p*-Value	Par est	*p*-Value	Par est	*p*-Value
Cereal	0.175	0.0099	0.194	0.0042	NA	NA
Brownie	0.624	0.0118	0.705	0.0045		
Celery	–0.366	0.0198	–0.434	0.0052		
Kale	0.757	0.0128	0.829	0.0058		
Ice cream	0.250	0.0431	0.320	0.0139		
Shrimp			0.693	0.0262		
Cake, commercial	–0.414	0.0536	–0.474	0.0263		
Roll, commercial	0.467	0.0241	0.449	0.0296		
Banana	–0.165	0.0805	–0.211	0.0302		
Brown rice	–0.608	0.0201	–0.606	0.0303		
Canned meat	–0.718	0.0766	–0.902	0.0309		
Hotdog	–0.335	0.0183	–0.300	0.0314		
Lamb	–1.124	0.0635	–1.242	0.0408		
Pie, homemade	–0.967	0.0465	–0.995	0.0408		
Dark bread	0.014	0.0217	0.127	0.0439		
Yogurt			–0.217	0.0462		
Cake, homemade			–0.495	0.0526		
White rice			–0.219	0.0592		
Corn, canned	–0.599	0.0824	–0.645	0.0628		
Potato, baked			–0.245	0.0698		
Crackers	0.085	0.0323	0.072	0.0716		
Apple	–0.208	0.0616	–0.197	0.0822		
Margarine			0.087	0.0864		
Cooked cereal			0.627	0.0993		
Beans, canned			0.441	0.1018		
Turkey			–0.191	0.1226		
Blueberries, fresh			0.367	0.1291		
Bran			–0.420	0.1314		
Potato chips			0.141	0.1668		
Nuts			–0.219	0.1716		
Broccoli			0.197	0.1756		
Roll, homemade	0.217	0.1176				

Abbreviations: Par est, parameter estimate; NA, not available.
[a]Data from Draper and Smith (*26*).

resulted in a formerly significant (at the $p <$ 0.2 level) variable dropping below the significance level. This may be attributable to covariance among some variables due either to true associations between or among the variables describing dietary consumption or artifactual associations linked with limited sample sizes. We are currently exploring techniques designed to address these potential problems.

The selection of a p-value cutoff is of interest in these analyses. Previous researchers (*1*,*17*) adopted a similar criterion. The principal reason to adopt the $p <$ 0.2 criterion rather than a more strict cutoff stems from the desire to capture all foods that could possibly be major predictors of residue concentrations in duplicate diet samples. In the bivariate approach, the loose criterion simply offers more foods for consideration in the second, multivariate step. In the multivariate step, we report all p-values that can assess the efficiency of using various cutoff criteria. Of course, inclusion of more variables increases the variance explained in such a model.

Table 4. Results representing major predictors of Pb derived from three regression approaches sorted by the stepwise p-value.

	Bivariate Adjusted R^2 = 0.1987[a]		Stepwise Adjusted R^2 = 0.3546[a]		Percent contribution Adjusted R^2 = 0.2105[a]	
Food Item	Par est	p-Value	Par est	p-Value	Par est	p-Value
Broccoli	–0.323	0.0007	–0.392	0.0001		
Peach, canned	–0.983	0.0158	–1.362	0.0006	–1.198	0.0036
Beets, canned	1.353	0.0691	2.471	0.0013	2.314	0.0029
Grapefruit	–0.502	0.0102	–0.602	0.0016	–0.597	0.0025
Chocolate bar			0.894	0.0019	0.802	0.0074
Cooked cereal	0.560	0.0347	0.806	0.0023		
Romaine lettuce	0.135	0.1004	0.251	0.0030	0.191	0.0237
Tomato, fresh	–0.285	0.0007	–0.229	0.0078	–0.236	0.0054
Ice cream	0.190	0.0369	0.240	0.0079	0.239	0.0104
Cottage cheese	0.260	0.0476	0.342	0.0088	0.363	0.0063
Sherbet	0.207	0.1137	0.329	0.0102	0.309	0.0234
Yams, fresh			0.566	0.0112	0.400	0.0762
Pancake			–0.271	0.0132	–0.243	0.0328
Peach, fresh			–0.371	0.0159	–0.324	0.0340
Oil and Vinegar	0.268	0.0061	0.229	0.0173	0.231	0.0185
Beef	0.315	0.0185	0.141	0.2817		
Canned meat	0.754	0.0165	0.687	0.0236	0.627	0.0476
Pasta	–0.125	0.0609	–0.147	0.0250	–0.187	0.0052
Oatmeal			–0.287	0.0262	–0.214	0.1124
Turkey			–0.195	0.0303		
Beef sandwich			0.187	0.0323	0.130	0.1372
Lima beans, fresh			0.394	0.0340	0.341	0.0727
Salt shake			–0.079	0.0363		
Other fish	0.193	0.0982	0.254	0.0422	0.276	0.0230
Liver	0.522	0.0904	0.586	0.0491	0.789	0.0202
Cookies, homemade	–0.239	0.0174	–0.201	0.0512	–0.302	0.0027
Brussels sprouts	1.349	0.0552	1.276	0.0597		
Tofu			1.382	0.0656		
Pork			–0.177	0.0666		
Orange			0.183	0.0684		
Dark fish, canned			0.528	0.0693		
Celery	0.219	0.0713	0.214	0.0723	0.333	0.0069
Mayonnaise			–0.101	0.0815		
Cake, commercial-made			0.317	0.0840	–0.239	0.1330
Mixed veg, canned	0.934	0.0961	0.934	0.0920	1.103	0.0508
Margarine			–0.061	0.1055		
Roll, homemade			–0.160	0.1148		
Pie, homemade			–0.537	0.1586		
Shrimp			0.223	0.1806		
Spinach, raw	0.926	0.0776	0.657	0.1921	0.870	0.0964
Chicken					–0.107	0.1361
Watermelon					–0.260	0.1073
Coffee whitener					–0.043	0.2525
Chicken, canned					–0.779	0.0542

Abbreviations: Par est, parameter estimate; veg, vegetables.
[a]Data from Draper and Smith (*26*).

Predicted versus Observed Intake

The DEPM and NHEXAS–Maryland database provide excellent companion data streams for cross-validation. The DEPM was developed to provide a standard methodology for assessing human exposure to contaminants from food ingestion. Working with several consumption and contaminant databases, it uses a market-basket framework to assess exposures to individuals and populations and is representative of residue data in individual food items in large populations. The duplicate diet residue levels determined in the NHEXAS–Maryland database represents the entire diet, not individual meals or food items, albeit on a limited sample of individuals and descriptive of only the smaller geographic area surrounding Baltimore, Maryland. Further, the NHEXAS–Maryland database includes use of the dietary checklist, which affords a tie-in with a much larger database, the semiquantitative food frequency questionnaire (*12*,*14*,*15*).

The DEPM model prediction of residue levels was only modestly successful. As indicated in Figure 1, the DEPM results, although statistically significant at the $p <$ 0.0001 level for all metals, predicted only a small fraction of the total variance in observed values when both were regressed as log-transformed variables. This result suggests that an approach using food checklist data, collected using the NHEXAS–Maryland checklist, coupled with residue data taken from national-scale market-basket studies, can lead to some predictive power in estimating

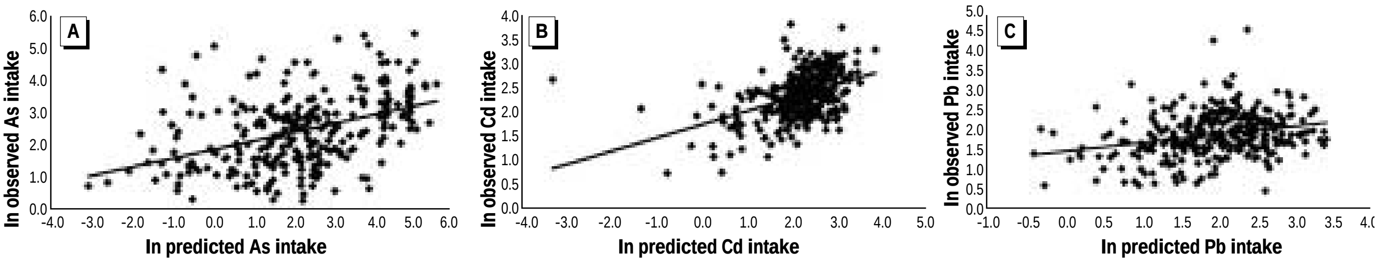

Figure 1. Observed natural log (ln) of metal concentration in duplicate diet samples versus predicted ln of metal concentration using DEPM. (*A*) Arsenic; ln observed = 0.276 ln predicted + 1.80; R^2 = 0.180. (*B*) Cadmium; ln observed = 0.276 ln predicted + 1.71; R^2 = 0.206. (*C*) Lead; ln observed = 0.224 ln predicted + 1.40; R^2 = 0.076.

exposure. However, much of the variance is still not explained. Such missing variance may come from specific variability in small-population residue data not accounted for in larger-region market-basket approaches using a relatively small number of samples to represent a region.

One possible explanation for the differences between values predicted through the two-step modeling procedure and those actually observed is that the food preparation step modifies the concentration observed in the duplicate diet samples. Although possible, such a modification is unlikely. First, TDS foods are prepared foods. While the preparation may not be identical to that done in the NHEXAS–Maryland homes, the TDS preparation techniques are designed to be similar to standard preparation. Second, we are evaluating elemental metals in this study. Thus, while we might expect concentration differences to be manifested due to changes in water content or other food attributes, unlike organic contaminants, volatilization, changes in solubility, or degradation are unlikely to play a significant role. More likely explanations include variation in residue levels among different samples of specific food items, and variation in serving sizes among the food items.

An alternative way of looking at these data is presented in the quintile plots in Figure 2. Often it is sufficient in epidemiologic investigations to separate individuals into a low-exposure group and a high-exposure group—say, the lowest quintile of exposure and the highest quintile of exposure. A statistically significant group separation may be accompanied by a concomitant separation in disease outcome, for example. Therefore, is the indirect method outlined above sufficient to categorize individuals in a manner consistent with the direct method? Inspection of these figures suggests that such a categorization can be effected. For As, a monotonic association is noted; increasing predicted quintile number is associated with increasing mean observed residue concentration. There is some overlap between quintiles, suggesting incomplete separation. However, Q1 and Q5 are well separated, suggesting that classification into a lowest exposure group and a highest exposure group may be cleanly effected for this metal using the indirect, modeling approach. Similar results can be seen for Cd exposure, with a similar monotonic relationship between predicted quintile and observed residue concentration. Separation is effected for Q1 and Q5 for this metal as well.

The results for Pb exposure are more problematic. The monotonic relationship between predicted and observed residue levels does not exist for this metal in our data. Despite this, however, one still can say with 95% confidence that Q1 and Q5 represent different residue levels, and an examination of differences in disease outcome for these two groups may be a fruitful path of investigation.

Daily Intake

The final step of this assessment examines the potential for adverse health effects in the study population due to dietary ingestion of these metals. The daily intake levels and the concentrations of metals as measured in the NHEXAS duplicate diet samples have been compared with intake levels and concentrations in foods as published in the Toxicological Profiles (*2–5*). The comparison is displayed in Table 5. This table is designed to relate the results of this study to a broader base of values and determine the extent to which our results can be generalized to a larger population. The concentrations of As, Cd, and Pb detected in NHEXAS food samples fall within the range of concentrations published by ATSDR, while the Cr concentration is somewhat greater than the range published in the Toxicological Profiles. The intake values for As and Cd are below the published intake values, and for Pb, the intake value as determined from the NHEXAS data falls midway in the range published in the Toxicological Profile. While these values are consistent with those given in the ATSDR reviews, it is important to note that, particularly for As, not all species are toxicologically equivalent. Arsenobetaine, for example, offers little risk. However, our data do not support understanding the speciation of As or the other metals. Further research in this area is warranted. These results suggest that the intake of metals from ingestion of these duplicate diets would not likely pose a serious threat to human health.

Other Issues

The data collected in the NHEXAS–Maryland study included such demographic information as social/racial/economic data that may afford assessment of difference in predictive power for subgroups within the study. However, the NHEXAS–Maryland investigation was a pilot-level study with a relatively small number of participants. Subpopulation analysis was not done due to small subgroup numbers.

Table 5. Comparison of values from NHEXAS diet study with data published in Toxicological Profiles.[a]

	NHEXAS		Toxicological profiles	
Metal	Concentration in food (ppb)[b]	Intake (mg/day)[b]	Concentration in food (ppb)	Intake (mg/day)
As	31.1	27.82	20–140	50
Cd	16.4	10.44	2–40	30
Cr	64.8	—[c]	20–50	25–224
Pb	11.0	8.14	3–83 dairy 2–159 meat/fish 2–136 grain/cereal 5–649 vegetables 5–223 fruit	5–11

[a]Agency for Toxic Substances and Disease Registry (*2–5*). [b]Intake values are based on the combination of NHEXAS–MD consumption levels and TDS residue levels. [c]Cr intake levels were not calculated.

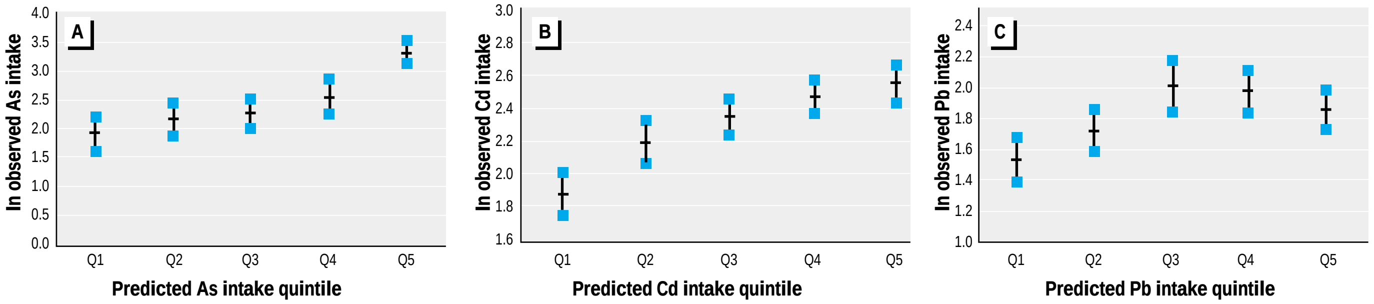

Figure 2. Plot of mean of the observed natural log duplicate-diet metal concentrations (± 2 SE) versus quintile of predicted ln metal concentration. (*A*) Arsenic. (*B*) Cadmium. (*C*) Lead.

REFERENCES AND NOTES

1. MacIntosh DL, Williams PL, Hunter DJ, Sampson LA, Morris SC, Willett WC, Rimm EB. Evaluation of a food frequency questionnaire-food composition approach for estimating dietary intake of inorganic arsenic and methylmercury. Cancer Epidemiol Biomark Prev 6:1043–1050 (1997).
2. ATSDR. Toxicological Profile for Arsenic. Atlanta, GA:Agency for Toxic Substances and Disease Registry, 1993.
3. ATSDR. DRAFT Toxicological Profile for Cadmium. Atlanta, GA:Agency for Toxic Substances and Disease Registry, 1997.
4. ATSDR. Toxicological Profile for Chromium. Atlanta, GA: Agency for Toxic Substances and Disease Registry, 1993.
5. ATSDR. DRAFT Toxicological Profile for Lead. Atlanta, GA:Agency for Toxic Substances and Disease Registry, 1997.
6. Kohlmeier L. Future of dietary exposure assessment. Am J Clin Nutr 61(suppl):702S–709S (1995).
7. Kohlmeier L, Bellach B. Exposure assessment error and its handling in nutritional epidemiology. Annu Rev Public Health 16:43–59 (1995).
8. Berry MR. Advances in dietary exposure research at the United States Environmental Protection Agency–National Exposure Research Laboratory. J Expo Anal Environ Epidemiol 7:3–16 (1997).
9. Klaassen CD, Amdur MO, Doull J. Casarett and Doull's Toxicology: The Basic Science of Poisons. New York:McGraw-Hill, 1996;83–86.
10. Adams MA, Bolger MB, Carrington CD, Coker CE, Cramer GM, Dinovi MJ, Dolan S. Guidance document for chromium in shellfish. Available: http://vm.cfsan.fda.gov/~frf/guid-cr.html [cited 8 March 1998].
11. TAS. Dietary Exposure Potential Model. Washington, DC:Technical Assessment Systems, Inc., 1996.
12. Scanlon KA, MacIntosh DL, Hammerstrom KA, Ryan PB. A longitudinal investigation of solid-food based dietary exposure to selected elements. J Expo Anal Environ Epidemiol 9:485–493 (1999).
13. Ryan PB. Quality System Implementation Plan for the NHEXAS–Maryland Investigation. Washington, DC:U.S. Environmental Protection Agency, 1995.
14. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, Hennekens CH, Speizer FE. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 122:51–65 (1985).
15. Rimm EB, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC. Reproducibility and validity of an expanded, self-adminstered food frequency questionniare among male health professionals. Am J Epidemiol 135:1114–1126 (1992).
16. FDA. Standard Operating Procedures for Determination of Trace Elements in NHEXAS Food Composites by Inductively Coupled Plasma Mass Spectrometry. Washington, DC:U.S. Food and Drug Administration, Center for Food Safety and Nutrition, 1997.
17. Romieu I, Stampfer MJ, Stryker WS, Hernandez M, Kaplan L, Sober A, Rosner B, Willett WC. Food predictors of plasma beta-carotene and alpha-tocopherol: validation of a food frequency questionnaire. Am J Epidemiol 131:864–876 (1990).
18. SAS Institute Inc. Proprietary Software. Cary, NC:SAS Institute Inc., 1992.
19. SAS Institute Inc. Version 6.12 for Windows. Cary, NC:SAS Institute, Inc., 1996.
20. Tomerlin JR, Berry MR, Tran NL, Chew S-B, Petersen BJ, Tucker KD, Fleming KH. Development of a Dietary Exposure Potential Model for evaluating dietary exposure to chemical residues in food. Expo Anal Environ Epidemiol 7:81–101 (1997).
21. Longnecker MP, Harper JM, Kim S. Eating frequency in the Nationwide Food Consumption Survey (U.S.A.), 1987–1988. Appetite 29:55–59 (1997).
22. NDC. Nutrition Basics. Rosemont, IL:National Dairy Council, 1994;3–30.
23. Scanlon KA. Analyses for dietary intake of selected elements. In: Environmental and Occupational Health. Atlanta, GA:Emory University, 1998.
24. Pennington JAT, Schoen SA. Total Diet Study: estimated dietary intakes of nutritional elements, 1982–1991. Int J Vitam Nutr Res 66:350–362 (1996).
25. MacIntosh DL, Spengler JD, Özkaynak H, Tsai L-H, Ryan PB. Dietary exposures to selected metals and pesticides. Environ Health Perspect 104:202–209 (1996).
26. Draper NR, Smith H. Applied Regression Analysis. New York:John Wiley & Sons, 1981.