Biasing Factors for Simple Soil Ingestion Estimates in Mass Balance Studies of Soil Ingestion

Edward J. Stanek III,1 Edward J. Calabrese,2 and Martha Zorn1

1 Department of Biostatistics and Epidemiology; 2Department of Environmental Health, University of Massachusetts at Amherst, Amherst, Massachusetts 01002

ABSTRACT

Soil ingestion estimates from mass balance soil ingestion studies can be used in Monte Carlo Risk assessment. We develop and describe a simulation model based on four mass balance soil ingestion studies that enables food ingestion, soil ingestion, and transit time to be mimicked. We use the simulation to evaluate potential biases that exist in current estimates of the distribution of daily soil ingestion in children (constructed from subject specific average daily estimates). The simulation identifies the importance of the study duration on the bias in the upper percentile soil ingestion estimates, indicating that the 95% soil ingestion estimate may be positively biased by over 100%. Misspecification of play areas for soil sampling is shown to have no biasing effect, and absorption of trace elements in food of up to 30% is shown to bias the soil ingestion distribution by less than 20 mg/d. The results, based on Al and Si trace element estimates, define the limits of previously published soil ingestion estimates, and provide insight for future study design and estimation methods.

INTRODUCTION

Possible ingestion of contaminated soil is often a driving factor in setting clean up levels based on a risk assessment. Estimates of exposure (via soil ingestion) are key to evaluating risk. The standard preferred method of estimating the quantity of soil ingested by children and adults is the trace element mass-balance approach. Mass balance soil ingestion studies are conducted by measuring trace element input (from food and water) and output (from urine and feces) over comparable time periods on subjects, and then converting the difference to the amount of soil ingested using element concentrations in soil. Although this approach has been shown to be valid among adults at relatively high soil ingestion levels, when soil...
ingestion is low, many factors may create difficulty in separating the soil ingestion signal from the noise. Five mass balance soil ingestion studies have been conducted in children to date (Binder et al. 1986; Calabrese et al. 1989; Davis et al. 1990; Calabrese et al. 1997a,b). Estimates from these studies agree broadly, with mean daily soil ingestion estimates ranging from 3 to 181 mg/d for Al, and ranging from 0 to 184 mg/d for Si. However, percentile estimates, and estimates based on other trace elements differ more widely.

It is well recognized that soil ingestion will vary both from day to day for a subject, and also vary from subject to subject. The natural variability in soil ingestion is accounted for in deterministic risk assessments via use of a high-end percentile estimate (usually the 95th percentile) from study samples. However, the differences in soil ingestion estimates between trace elements (based on the same subjects over the same time period) are evidence of uncertainty. The extent of these differences has prompted much research in strategies to minimize the uncertainty, and arrive at more accurate estimates (Stanek and Calabrese 1991; Calabrese and Stanek 1991; Stanek and Calabrese 1995a,b; Stanek and Calabrese 2000). Experimental studies in adults have not eliminated uncertainty as a source of concern, since reliability was reasonable only at high soil ingestion levels. Generalization of these results to children depends on the similarity between children and adults.

The current regulatory environment is placing increasing attention on the use of Monte Carlo methods to quantify variability in risk based on variability in multiple exposure variables. In a probabilistic risk assessment, it is important to be able to distinguish variability (which the Monte Carlo study should incorporate), from uncertainty (which would result in over stating both high and low ingestion estimates). While experimental soil ingestion studies in children have the potential of quantifying various sources of uncertainty (such as sample loss, total recovery, absorption, and other sources of trace elements); such studies are costly and may be impractical.

We present results that quantify uncertainty in soil ingestion estimates based on mass-balance study designs. The results identify the importance of three possible sources of uncertainty: the length of the study; ingestion of non-sampled soil; and bioavailability of food. For each source, the bias that results from estimating the cumulative soil ingestion distribution by subject specific average estimates is quantified. We use these results to discuss the extent to which variability can be separated from uncertainty in the context of available children’s soil ingestion studies. The results are achieved by specifying a distribution of soil ingestion (known at the start), and then simulating other factors that would interfere with a simple objective measure of soil ingestion.

MATERIALS AND METHODS

Much has been learned about factors that affect estimates of soil ingestion from a mass-balance study. The distribution of the amount of food ingested, the trace element concentration in food, the trace element concentration in soil, and the timing of fecal samples have been recorded in five children studies. These distributions characterize what has been found in the past, and may encompass the distribution of these factors likely to occur in similar future mass-balance studies. There may
be considerable question about the relative contributions of uncertainty and variability when characterizing the distribution of soil ingested between children and days.

We simulate other factors using extensive data on the two most promising trace elements (Al and Si) collected from four mass-balance studies. For each subject, we develop distributions of trace element food intake and fecal output that mirror results seen in previous research. The distributions are used to simulate food intake and fecal output data for a hypothetical population of subjects using a study design, as well as trace element concentrations in soil. We then postulate a distribution of soil ingestion. A simulation is conducted that results in a snapshot of the hypothetical study data for the design. Since the true soil ingestion of the subjects underlying the data is known, we use the results to evaluate the accuracy and reliability of current estimates. By varying the assumed soil ingestion distribution, we characterize the extent to which current studies can separate variability from uncertainty when estimating soil ingestion. The simulation also allows these factors to be investigated at minimal cost.

There are four stages in this study. First, using previous mass-balance study data, we characterized the distribution of trace element concentrations in food and soil along with the distribution of quantity of food ingested, and develop a simulation program that will produce a hypothetical set of trace element amounts in food and fecal samples for a specified study protocol. Second, we verify that the simulations produce results that are similar to those reported in previous mass-balance soil ingestion studies. We do so by comparing the distributions resulting from the simulations with distributions from actual mass-balance studies, and by evaluating whether or not with an adequate study design, the simulation will reproduce a hypothetical soil ingestion distribution.

Next, we specify six hypothetical soil ingestion distributions, using current estimates of children’s soil ingestion to bound the soil ingestion distributions. Then, using these distributions, we describe the ability of a particular study design (based on a 4-day and 7-day design) and estimation strategy (based on an overall average per subject) to characterize the hypothetical soil ingestion distribution (separating uncertainty from variability). In this context, we also assess the sensitivity of the uncertainty estimates to various assumptions including ingested soil from a yard other than the subject’s yard, or absorption of trace elements.

**STUDIES AND TRACE ELEMENTS**

The basic data on trace element ingestion from food and soil concentrations are derived from the four published mass-balance studies of soil ingestion in children that included measures of trace element intake from food (Calabrese et al. 1989; Davis et al. 1990; Calabrese et al. 1997a,b). The Amherst (Calabrese et al. 1989) soil ingestion study collected samples on 64 preschool children for 2 weeks (with 3 days of food, and 4 days of fecal samples each week). The Washington State (Davis et al. 1990) soil ingestion study collected duplicate food samples and all feces and urine over a four consecutive day period. Samples were pooled over the time period, with the pooled sample weighed and analyzed for element concentrations for the 101 study subjects. The Anaconda soil ingestion study (Calabrese et al. 1997b) collected daily duplicate food samples and feces on 64 children for seven consecutive days.
The children's pica (Calabrese et al. 1997a) soil ingestion study collected daily duplicate food samples and feces samples on twelve children for seven consecutive days. All four studies included three trace elements (Al, Si, and Ti). We limited the investigation to the trace elements Al and Si since they had relatively high concentrations (and uniform concentration) in soil; relatively low amounts (in terms of food/soil equivalents) in food; and relatively low apparent intake of these elements from other sources. We first characterize these distributions so as to be able to simulate trace element ingestion from food, and fecal transit times.

Determining Al and Si Concentration Distributions in Food and Soil

Four concentration distributions [Al and Si in food and soil] were characterized. First, each concentration distribution was summarized via histograms and descriptive statistics for each study by pooling daily estimates for subjects in the study. Concentrations were above the detection limit for all subjects and days except for 9% of estimates for Al from the Amherst study, where the concentration was set to the detection limit (0.050 μg/g). Although the concentration distributions differed between studies, we chose to pool all study data to develop basic models of element concentrations in food, accounting for the study differences via a random study effect. When the distribution did not appear to be normal (based on normal probability plots) a Box/Cox transformation was fit in STATA (Statacorp 1997) to identify a normalizing power transformation. The estimate of the power was used to transform the data. The normality of the transformed data was assessed via normal probability plots.

Next, nested mixed models were fit using SAS (SAS Institute 1996) to the transformed food concentrations data for each study (and overall studies), with variance components estimated for studies, subjects, and days. The child's age (in months) was included as an independent regression variable in these analyses. Age had a very modest effect on trace element concentrations (increasing the concentrations by 2% per year for Al, and decreasing the concentration by 1% per year for Si). Parameter estimates from these studies were used in subsequent simulations of trace element concentrations in food.

Determining Freeze-Dried Food Weight Distributions

A similar procedure was followed to characterize the freeze-dried food weight. For the Washington State study, the corrected freeze-dried food weight was divided by the length of the collection period to obtain a daily estimate of freeze dried food weight. Inspection of the distribution of daily freeze dried food weight indicated that the weight was approximately normally distributed. Age was accounted for in the mixed models that were fit to estimate variance components between subjects and days for weight. These models formed the basis of parameter estimates that were used to simulate the quantity of food ingested.

Simulating Transit Time from Food to Fecal Samples

The final parameter that was characterized was the distribution of days between fecal samples. Data from two studies (Calabrese et al. 1989; Calabrese et al. 1997b) that collected daily measures was used to characterize this distribution. First, the
distribution of fecal sample intervals was tabulated for each study. A simple Poisson model was fit to the pooled data to estimate fecal intervals. Differences by age were also examined.

**Linking Food and Fecal Trace Element Amounts**

The results of these analyses were used to simulate food intake of trace elements over subjects and days. We link the trace element intake in food to subsequent fecal output by a transit time representing the time period it takes for the food to appear in the fecal sample. The simulation assumes that once the initial time period is set, subsequent food intake would pass continuously to fecal samples with the same time lag. However, due to the interval distribution of fecal samples, the continuous stream of fecal material will be seen in lumps (depending on the fecal sample interval) and not as a steady sample.

We simulate the initial transit time by generating a random number from the Poisson model fit to fecal sample intervals, where the model is expressed in hours (not days). Given the number of transit hours, we partition the trace element amounts from food into equivalent amounts on subsequent days. Then, selecting another random Poisson variable based on the Poisson distribution for fecal sample intervals expressed in days, we lump the fecal samples to simulate the actual observed fecal amounts.

**Selection of Hypothetical Soil Ingestion Distributions**

We select hypothetical soil ingestion distributions that span the plausible soil ingestion distributions suggested by the literature. The soil ingestion distributions that we consider are the following:

1. All subjects ingest 100 mg of soil on all days.

2. All subjects ingest 50 mg of soil on all days.

3. All subjects have a mean soil ingestion of 100 mg/day, but daily soil ingestion is normally distributed, with a standard deviation of 50 mg/d.

4. All subjects have a mean soil ingestion of 50 mg/day, but daily soil ingestion is normally distributed, with a standard deviation of 75 mg/d.

5. Daily soil ingestion is normally distributed, with mean 18 mg/d, subject standard deviation of 56 mg/d, and daily variance of 108 mg/d (based on the average of estimates from mixed models fit to Al, Si, Y, and Zr from the Amherst and Anaconda studies (Stanek and Calabrese 2000)). The distribution is based on an average of the mean and variance component estimates from the Amherst and Anaconda studies, as developed in Stanek and Calabrese (2000).

6. An empirical soil ingestion distribution based on Amherst and Anaconda subjects (excluding the pica subject), with normally distributed daily variation in estimates. The daily standard deviation is estimated from a simple linear
regression model of the standard deviation on the mean (intercept=63.5, slope=0.834). Subject values were based on the average of the median daily estimate (assuming a 28hr lag) for Al, Si, Y, and Zr. The empirical distribution is based on data from the same studies using the common trace elements (except Ti), with no outlier criteria applied.

RESULTS

We first describe the basic results on food, soil, and transit time distributions derived from previously reported soil ingestion studies that form the basis of the simulations. Details of the modeled distributions are presented. We evaluate the appropriateness of these distributions by superimposing simulated data (using results of the modeling exercise) on data obtained from the soil ingestion studies. Following this discussion, we quantify the importance of three possible sources of uncertainty: the length of the study; bioavailability of food; and ingestion of non-sampled soil.

The first step in designing the simulations was characterization of the food and soil concentration distributions of Al and Si. Concentration distributions of Al and Si in food were skewed to the right in all studies. Box/Cox analyses indicated raising concentrations to the power 0.1369 and -0.1915 for Al and Si, respectively. The estimated mean and variance components from mixed models fit to these concentration data are summarized in Table 1. We included a study component of variance in the modeling as in a meta analysis approach. For both trace elements, this component of variance was less than the component of variance from day to day.

Al and Si concentrations in soil were only slightly skewed to the right. As a result, the soil concentrations were not transformed. Parameter estimates from the mixed models are also given in Table 1. Food sample freeze-dried weight was normally distributed in the studies (after accounting for age) and hence not transformed, with mixed model parameter estimate given in Table 1. Finally, two studies contained daily data allowing an estimate of the time interval in fecal samples for children. In the Amherst study, out of a total time span of 448 days, there were 328 1-day intervals, 48 2-day intervals, and 8 3-day intervals. In the 7-day soil ingestion studies at Anaconda, out of a total time span of 427 days, there were 256 1-day intervals, 58 2-day intervals, 14 3-day intervals, 2 4-day intervals and 1 5-day interval. The mean interval between days was ranged from 1.15 to 1.29 in the studies. A simple Poisson model fit these data with Poisson parameter 1.22. There was evidence that younger children (between 12 months and 18 months of age) delivered fecal samples more regularly than children over 18 months of age. However, since there were few children in this age range in the studies, we assumed a common Poisson parameter for the simulations.

Figures 1 to 4 overlay the empirical distribution of trace element concentrations observed in the four studies with the simulated distributions based on the modeling results given above. The empirical data are based on the four studies with a total of 241 subjects, and 4 to 8 days of observation per subject. The simulation is based on 241 subjects and 5 days so as to roughly match the size of the empirical studies. In all cases, there is a close overlap between histograms of the observed and simulated distributions. Figure 5 presents the empirical and simulated distribution of freeze-
Biases in Soil Ingestion Studies

Table 1. Summary of parameter estimates from mixed models fit to concentrations of Al and Si in food and soil, and freeze dried food weight from pooled data on 241 children (ages 1 to 7 yrs) with 1016 observations (Calabrese et al. 1989; Davis et al. 1990; Calabrese et al. 1997a,b).

<table>
<thead>
<tr>
<th></th>
<th>Power</th>
<th>Mean</th>
<th>Slope (age in Mths)</th>
<th>Subject Variance</th>
<th>Study Variance</th>
<th>Day Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al in Food µg/g</td>
<td>0.1369</td>
<td>1.27</td>
<td>0.0022</td>
<td>0.0172</td>
<td>0.0105</td>
<td>0.0478</td>
</tr>
<tr>
<td>Si in Food µg/g</td>
<td>-0.1915</td>
<td>0.455</td>
<td>-0.0004</td>
<td>0.0</td>
<td>0.0006</td>
<td>0.00185</td>
</tr>
<tr>
<td>Al in Soil µg/g</td>
<td>1</td>
<td>56.8</td>
<td>-0.007</td>
<td>46.8</td>
<td>64.7</td>
<td>na</td>
</tr>
<tr>
<td>Si in Soil µg/g</td>
<td>1</td>
<td>292.1</td>
<td>-0.076</td>
<td>1645</td>
<td>1462</td>
<td></td>
</tr>
<tr>
<td>Frz. dried food wt (g)</td>
<td>1</td>
<td>125.2</td>
<td>2.13</td>
<td>49.3</td>
<td>1146</td>
<td>2031</td>
</tr>
</tbody>
</table>

a p<0.05
b p<0.001

dried food weights. In a similar manner, Figure 6 illustrates the results of simulated fecal sample intervals compared to actual intervals observed in the two studies with daily measures. Once again, the results indicate a good ability to generate data consistent with what has been observed in past soil ingestion studies.

For clarity, we illustrate generation of the Al data for a simulated subject with 100 mg/d soil ingestion followed in a 7-day soil ingestion study, and a 36-hour initial transit time in Table 2. A 5-day "burn in" period of food ingestion is included prior to the start of the ingestion study. This allows for the initial transit time, and enables trace elements to build up in lumped fecal samples. Note that we allow the initial transit time to vary hourly, but afterwards consider the fecal amounts to be identical to the food plus soil amounts, apart from the transit time difference. The differences in amounts between the fecal sample passing through and a simple sum of the (food + soil) amounts occurs since the transit time assumption may fractionate trace element amounts within a day. The lumped amounts result from irregular fecal output, as generated by the fecal interval random variable in the simulation.

We characterize the four stochastic long term average soil ingestion distributions used in this study by simulating soil ingestion for 365 days for a large number of children (5000) for each distribution, and then tabulating the cumulative distribution (Figure 7). We illustrate the distribution based on estimates from Al, but note that similar results occur when estimates are based on Si. The empirical distribution of 6- to 8-day average soil ingestion from the 63 Amherst and 64 Anaconda children (based on the average of the median daily estimate [assuming a 28 hour lag] for Al, Si, Y, and Zr) is included in Figure 7. The near vertical line at 100 mg/day soil ingestion represents the distribution of 365 day average soil ingestion for 5000 children assuming each child ingests an average of 100 mg of soil per day (when averaged over a very long time period). The bending of the tails of the cumulative
Figure 1. Plot of empirical and simulated distributions for Al food concentrations.
Figure 2. Plot of empirical and simulated distributions for Si food concentrations.
Figure 3. Plot of empirical and simulated distributions for Al soil concentrations.
Figure 4. Plot of empirical and simulated distributions for Si soil concentrations.
Figure 5. Plot of empirical and simulated distributions for freeze dried food weight.
Figure 6. Chart of empirical and simulated distributions for fecal sample intervals.
Table 2. Illustration of simulated Al food intake and fecal output for a child ingesting 100 mg/day of soil (with an Al soil concentration of 60 mg/g) over a 7-day study assuming a 36-hour initial transit time.

<table>
<thead>
<tr>
<th>Day</th>
<th>Al Conc in Food (ug/g)</th>
<th>Freeze Dried Food Weight (g)</th>
<th>Total mg Al in Food ingested from Soil</th>
<th>Total mg Al ingested</th>
<th>Al in Fecal Sample (mg)</th>
<th>Al (mg) in Lumped Fecal Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>-4</td>
<td>22.36</td>
<td>326</td>
<td>7.29</td>
<td>6</td>
<td>13.29</td>
<td></td>
</tr>
<tr>
<td>-3</td>
<td>0.60</td>
<td>365</td>
<td>0.22</td>
<td>6</td>
<td>6.22</td>
<td>6.65</td>
</tr>
<tr>
<td>-2</td>
<td>0.96</td>
<td>277</td>
<td>0.26</td>
<td>6</td>
<td>6.26</td>
<td>9.75</td>
</tr>
<tr>
<td>-1</td>
<td>27.99</td>
<td>213</td>
<td>5.97</td>
<td>6</td>
<td>11.97</td>
<td>6.24</td>
</tr>
<tr>
<td>0</td>
<td>33.18</td>
<td>278</td>
<td>9.23</td>
<td>6</td>
<td>15.23</td>
<td>9.12</td>
</tr>
<tr>
<td>1</td>
<td>7.33</td>
<td>242</td>
<td>1.78</td>
<td>6</td>
<td>7.78</td>
<td>13.60</td>
</tr>
<tr>
<td>2</td>
<td>1.35</td>
<td>257</td>
<td>0.35</td>
<td>6</td>
<td>6.34</td>
<td>11.50</td>
</tr>
<tr>
<td>3</td>
<td>11.93</td>
<td>290</td>
<td>3.45</td>
<td>6</td>
<td>9.45</td>
<td>7.06</td>
</tr>
<tr>
<td>4</td>
<td>11.71</td>
<td>316</td>
<td>3.70</td>
<td>6</td>
<td>9.70</td>
<td>7.90</td>
</tr>
<tr>
<td>5</td>
<td>1.74</td>
<td>358</td>
<td>0.62</td>
<td>6</td>
<td>6.62</td>
<td>9.58</td>
</tr>
<tr>
<td>6</td>
<td>0.55</td>
<td>236</td>
<td>0.13</td>
<td>6</td>
<td>6.13</td>
<td>8.16</td>
</tr>
<tr>
<td>7</td>
<td>5.09</td>
<td>302</td>
<td>1.54</td>
<td>6</td>
<td>7.54</td>
<td>6.38</td>
</tr>
</tbody>
</table>
Figure 7. Cumulative distribution of four stochastic long-term average soil ingestion distributions. Empirical distribution of soil ingestion from 63 Amherst and 64 Anaconda children is overlaid. Stochastic distributions based on simulations for 5000 subjects, 365 days/6, using Al as trace element.
distribution is due to the daily variation in soil ingestion, assumed to have a standard deviation of 50 mg/d. With this standard deviation, the standard error of the mean over 365 days is 2.6 mg/day. This standard error accounts for the bending of nearly 5 mg/day at the 2.5th percentile and 97.5th percentile of the cumulative distribution.

A similar spread of the distribution is evident when each child ingests an average of 50 mg/d (when averaged over a very long time period). The daily standard deviation is set to 75 mg/d, resulting in a standard error of the 365 day mean of 3.9 mg/d. Note that the mean of the distribution of soil ingestion does not appear to be 50, but is actually 61 mg/d. This higher soil ingestion average is a result of setting soil ingestion values to zero when the simulated soil ingestion is less than zero. Such apparent negative soil ingestion can occur in the simulation since when a daily deviation in soil ingestion has a negative value larger than the true long run mean. Forcing all simulated soil ingestion day values to be at least zero shifts up the lower part of the distribution, and truncates the low values at zero.

The third simulated soil ingestion distribution displayed in Figure 7 corresponds to the soil ingestion distribution estimated based on mixed models fit to the Amherst and Anaconda data (distribution #5). If the estimates of soil ingestion from these studies did correspond to the true soil ingestion distribution of children, the mean soil ingestion would be 18 mg/d, with subject and day standard deviations of 56 mg/d and 108 mg/d, respectively. The combined effect of the subject and daily standard deviations would result in a standard error for a subject of 5.6 mg/d, spreading the tails of the distribution (at the 2.5th percentile and 97.5th percentile) by approximately 11 mg/d. The impact of forcing all daily soil ingestion to be at least zero shifts the lower portion of the distribution up, resulting in the observed 50th percentile being 53 mg/d (as opposed to 18 mg/d) and the 75th percentile being 77 mg/d (as opposed to 65 mg/d). A similar, but less dramatic effect is evident for the fourth soil ingestion distribution in Figure 7. The small shift in the lower end of the distribution results from fewer simulated soil ingestion daily values being forced to be greater than zero.

The final distribution displayed in Figure 7 is the empirical distribution based on the Amherst and Anaconda studies. This distribution differs from the other simulated distributions since it represents 7-day average soil ingestion, as opposed to 365-day average soil ingestion. Since soil ingestion varies from day to day for a subject, this distribution is more spread out than the distribution of average soil ingestion for children in the long 365-day period.

The Impact of the Study Duration

We discuss how the study duration impacts the spread of the empirical distribution. We limit the presentation to two soil ingestion distributions, noting that similar results occur for other distributions. For most comparisons, we focus on the results based on soil ingestion distribution #6 (median: 26mg/d; 90th percentile: 106 mg/d; 95th percentile: 149 mg/d), since this distribution comes closest to the soil ingestion observed in previous studies, and overlaps the empirical soil ingestion distribution at upper percentiles of soil ingestion. For some comparison, we include the simpler soil ingestion distribution #4 (mean soil ingestion of 50 mg/d with a daily standard deviation of 75 mg/d).
Biases in Soil Ingestion Studies

First, we consider the impact of study duration on the distribution of average daily soil ingestion by simulating soil ingestion distributions over different study duration times. Figure 8 presents results that illustrate the effect of study duration on the estimated soil ingestion distribution based on soil ingestion distribution #4, while Figure 9 present similar results based on distribution #6. Each cumulative distribution is obtained from a soil ingestion study of the specified duration by tabulating the average soil ingestion per person for 5000 subjects. The results in Figure 8 graphically dramatize the impact of study duration on the resulting distribution of average soil ingestion. Although the true long run soil ingestion for each subject is 50 mg/d, the upper percentiles of the average soil ingestion distribution are much higher, particularly for studies of length 4 or 7 days. Using the 4-day or 7-day averages to estimate the 95th percentile of the long run soil ingestion distribution will over estimate the true 95th percentile long run soil ingestion. Similar, but less dramatic results are illustrated in Figure 9 using soil ingestion distribution #6.

Figures 8 and 9 illustrate that short study designs with the soil ingestion distribution estimated by the average soil ingestion will over estimate the upper tails of the distribution (and underestimate the lower tails of the distribution). We summarize the expected impact on estimates of percentiles of these distributions in Table 3 for Al and Si. Focusing on the 95th percentile estimate in a 4-day study where the true long run soil ingestion was 50 mg/d (daily standard deviation 75 mg/d), the 95th percentile estimate will be positively biased by 112% for Al, and 91% for Si. For the same true long-run soil ingestion distribution in a 7-day study, the 95th percentile estimate will be positively biased by 70% for Al, and 60% for Si.

Smaller biases occur when the results are tabulated based on the true long run soil ingestion distribution #6. In a 4-day study, the 95th percentile estimate will be positively biased by 24% for Al, and 16% for Si. For the same true long-run soil ingestion distribution in a 7-day study, the 95th percentile estimate will be positively biased by 11% for Al, and 10% for Si.

The bias is due to an increased spread of the estimated soil ingestion distribution. This bias can be summarized by comparing estimates of the between subject variance with the true between subject variance when the estimated variance is based on average soil ingestion estimates from different duration study designs. Using the soil ingestion distribution #6 and a 4-day study, the variance between subjects will be overestimated by 186% and 116% in a 4-day study based on Al and Si, respectively, and by 88% and 66% in a 7-day study based on Al and Si, respectively.

The Impact of Ingestion of Soil from Neighbor's Yards and/or Absorption of Trace Elements from Food

Mass balance soil ingestion studies have based soil ingestion estimates on trace element concentrations in soil collected from play areas identified by children's guardians. It is possible that soil ingested by a child may not have come from these identified areas, but rather from a neighbor's or friend's yard. We evaluate the potential impact of using the wrong soil concentration for a trace element by simulating soil ingestion for a child based on a specified soil concentration, but selecting an independent soil concentration (to represent the trace element in the mis-specified play area) for estimation of soil ingestion. We illustrate the results of
Figure 8. Plot of simulated distributions of soil ingestion by length of trial for 5000 children assuming a mean = 50 mg/d, daily std. = 75 mg/d, and using Al as a trace element.
Figure 9. Plot of simulated distribution of soil ingestion by length of trial for 5000 children assuming the distribution from the Amherst/Anaconda study, using Al as a trace element.
Table 3. Cumulative distributions of average soil ingestion developed from simulations using various soil ingestion distributions for 4- and 7-day studies.

<table>
<thead>
<tr>
<th>Soil Ingestion Distribution</th>
<th>Trace Element</th>
<th>Subjects</th>
<th>Days</th>
<th>Mean</th>
<th>Std</th>
<th>P01</th>
<th>P05</th>
<th>P10</th>
<th>P25</th>
<th>P50</th>
<th>P75</th>
<th>P90</th>
<th>P95</th>
<th>P99</th>
</tr>
</thead>
<tbody>
<tr>
<td>#4a</td>
<td>AL</td>
<td>5000</td>
<td>4</td>
<td>61.1</td>
<td>53.3</td>
<td>-80.5</td>
<td>-12.5</td>
<td>8.7</td>
<td>34.0</td>
<td>58.7</td>
<td>86.7</td>
<td>116.5</td>
<td>141.1</td>
<td>216.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5000</td>
<td>7</td>
<td>61.4</td>
<td>33.5</td>
<td>-21.6</td>
<td>14.9</td>
<td>26.2</td>
<td>42.0</td>
<td>60.0</td>
<td>80.2</td>
<td>99.2</td>
<td>113.4</td>
<td>149.7</td>
</tr>
<tr>
<td></td>
<td>True Ing.</td>
<td>365</td>
<td>61.4</td>
<td>3.1</td>
<td>54.0</td>
<td>56.2</td>
<td>57.4</td>
<td>59.3</td>
<td>61.4</td>
<td>63.5</td>
<td>65.4</td>
<td>66.6</td>
<td>68.7</td>
<td></td>
</tr>
<tr>
<td>#4a</td>
<td>SI</td>
<td>5000</td>
<td>4</td>
<td>61.5</td>
<td>38.4</td>
<td>-21.9</td>
<td>2.9</td>
<td>16.6</td>
<td>36.5</td>
<td>59.8</td>
<td>84.5</td>
<td>108.9</td>
<td>126.7</td>
<td>162.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5000</td>
<td>7</td>
<td>61.4</td>
<td>26.3</td>
<td>2.8</td>
<td>21.1</td>
<td>29.6</td>
<td>43.4</td>
<td>60.2</td>
<td>78.0</td>
<td>95.1</td>
<td>106.2</td>
<td>127.4</td>
</tr>
<tr>
<td></td>
<td>True Ing.</td>
<td>365</td>
<td>61.4</td>
<td>3.1</td>
<td>54.2</td>
<td>56.3</td>
<td>57.4</td>
<td>59.3</td>
<td>61.4</td>
<td>63.5</td>
<td>65.5</td>
<td>66.5</td>
<td>68.6</td>
<td></td>
</tr>
<tr>
<td>#6b</td>
<td>AL</td>
<td>5000</td>
<td>4</td>
<td>55.3</td>
<td>70.0</td>
<td>-82.9</td>
<td>-21.8</td>
<td>-4.7</td>
<td>15.3</td>
<td>41.0</td>
<td>81.1</td>
<td>139.5</td>
<td>185.2</td>
<td>298.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5000</td>
<td>7</td>
<td>55.8</td>
<td>56.8</td>
<td>-44.6</td>
<td>-4.1</td>
<td>6.4</td>
<td>21.3</td>
<td>41.7</td>
<td>76.5</td>
<td>126.0</td>
<td>165.8</td>
<td>257.8</td>
</tr>
<tr>
<td></td>
<td>True Ing.</td>
<td>365</td>
<td>55.8</td>
<td>41.4</td>
<td>21.6</td>
<td>23.2</td>
<td>24.1</td>
<td>26.1</td>
<td>36.0</td>
<td>67.0</td>
<td>105.9</td>
<td>149.0</td>
<td>194.0</td>
<td></td>
</tr>
<tr>
<td>#6b</td>
<td>SI</td>
<td>5000</td>
<td>4</td>
<td>55.1</td>
<td>60.9</td>
<td>-34.5</td>
<td>-8.4</td>
<td>0.5</td>
<td>16.2</td>
<td>40.8</td>
<td>76.2</td>
<td>130.3</td>
<td>172.6</td>
<td>276.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5000</td>
<td>7</td>
<td>55.9</td>
<td>53.3</td>
<td>-13.1</td>
<td>2.7</td>
<td>9.5</td>
<td>22.3</td>
<td>41.0</td>
<td>73.2</td>
<td>120.2</td>
<td>163.4</td>
<td>254.4</td>
</tr>
<tr>
<td></td>
<td>True Ing.</td>
<td>365</td>
<td>55.8</td>
<td>41.4</td>
<td>21.6</td>
<td>23.2</td>
<td>24.1</td>
<td>26.1</td>
<td>36.0</td>
<td>66.9</td>
<td>105.7</td>
<td>149.1</td>
<td>194.2</td>
<td></td>
</tr>
</tbody>
</table>

a All subjects have a mean soil ingestion of 50 mg/day, but daily soil ingestion is normally distributed, with a standard deviation of 75 mg/day.

b An empirical soil ingestion distribution based on Amherst and Anaconda subjects (excluding the pica subject), with normally distributed daily variation in estimates based on standard deviation modeled from a simple linear regression on the mean (intercept = 63.5, slope = 0.834).
such mis-specification for a 7-day study design, where the true soil ingestion is given by distribution #6. The impact is assessed by tabulating the distribution of the average 7-day soil ingestion estimates for Al and Si from a simulated sample of 5000 subjects (Table 4), and comparing the estimated distribution with the similar distribution where the trace element concentration in soil is correct.

The first two rows of Table 4 for Al and Si illustrate the impact of mis-specification of the trace element concentration in soil. For either trace element, there is minimal impact on the estimated soil ingestion distribution due to mis-specification of soil concentrations. Since the estimated soil ingestion distribution appeared to be insensitive to misspecification of the child's play area, we examined the impact on the soil ingestion distribution if a single trace element soil concentration was used for each child. These results are given in the third row of Table 4 for Al and Si, and illustrate that the estimated soil ingestion distribution is insensitive to use of a common soil concentration. This indicates that for studies similar to the ones evaluated, a simple reliable average soil concentration estimate is adequate. Based on the soil data in Table 1, average soil concentration will have a coefficient of variation of less than 5% by taking nine random soil samples.

The results in Table 4 also describe the impact of absorption of trace elements from food. Although trace elements selected for use in mass-balance studies are thought to have low absorption, only limited studies have been conducted that characterize the actual absorption. If trace elements ingested in food are absorbed and the absorption is not accounted for when estimating soil ingestion, the soil ingestion distribution will be under-estimated. We evaluate the impact of food absorption by assuming 30% of each trace element in food is absorbed, and examining the impact on the distribution of seven day average soil ingestion. Such results are given in rows 4 and 5 of Table 4 for each element.

In the simulations, the average soil equivalent amount of Al in food is 64 mg/d, while the average soil equivalent amount of Si in food is 58 mg/d. The assumption of 30% absorption from food will reduce the soil equivalent amounts from food in fecal samples by 19 mg/d and 17 mg/d for Al and Si, respectively. These values correspond closely to the observed bias of the soil ingestion distribution in row 4 of the panels for Al and Si in Table 4. Row 5 in Table 4 presents similar results when there is absorption of food, and the play area is misspecified for the soil sample. Once again, since the impact of missing the play area is negligible, the results are similar to those in row 4.

Finally, note that the last row in each panel of Table 4 presents the actual distribution of long-term soil ingestion (calculated by simulating soil ingestion for 365 days on each of 5000 subjects). These results are identical to the last rows in each of the last two panels of Table 3. It is this distribution that we attempt to estimate in a soil ingestion study. In summary, Table 4 illustrates that the most important impact on estimates of the distribution of soil ingestion is the duration of the study design. There is little impact on the distribution of misspecifying the child’s play area. Food absorption of 30% will bias the distribution estimates down by less than 20 mg/d.

DISCUSSION

The development of a simulation model for mass balance soil ingestion studies is a powerful tool for evaluating biases in previous soil ingestion estimates, evaluat-
Table 4. Cumulative distributions of average soil ingestion developed from simulations using soil ingestion distribution #6\textsuperscript{a}, taking into account absorption and ingesting different soil.

<table>
<thead>
<tr>
<th>Trace</th>
<th>Element</th>
<th>Abortion</th>
<th>Soil</th>
<th>Subjects</th>
<th>Days</th>
<th>Mean</th>
<th>Std</th>
<th>P01</th>
<th>P05</th>
<th>P10</th>
<th>P25</th>
<th>P50</th>
<th>P75</th>
<th>P90</th>
<th>P95</th>
<th>P99</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL</td>
<td>None</td>
<td>Same</td>
<td>5000</td>
<td>7</td>
<td>55.8</td>
<td>56.8</td>
<td>-44.6</td>
<td>-4.1</td>
<td>6.4</td>
<td>21.3</td>
<td>41.7</td>
<td>76.5</td>
<td>126.0</td>
<td>165.8</td>
<td>257.8</td>
<td></td>
</tr>
<tr>
<td>AL</td>
<td>None</td>
<td>Different</td>
<td>5000</td>
<td>7</td>
<td>54.7</td>
<td>58.8</td>
<td>-37.8</td>
<td>-3.5</td>
<td>5.5</td>
<td>19.9</td>
<td>41.0</td>
<td>73.0</td>
<td>123.9</td>
<td>162.3</td>
<td>272.4</td>
<td></td>
</tr>
<tr>
<td>AL</td>
<td>None</td>
<td>Constant</td>
<td>5000</td>
<td>7</td>
<td>55.9</td>
<td>57.7</td>
<td>-43.4</td>
<td>-4.0</td>
<td>6.2</td>
<td>20.7</td>
<td>41.4</td>
<td>76.6</td>
<td>127.7</td>
<td>168.0</td>
<td>261.9</td>
<td></td>
</tr>
<tr>
<td>AL</td>
<td>30% food</td>
<td>Same</td>
<td>5000</td>
<td>7</td>
<td>36.2</td>
<td>59.4</td>
<td>-92.1</td>
<td>-36.0</td>
<td>-17.5</td>
<td>5.0</td>
<td>26.2</td>
<td>57.9</td>
<td>106.5</td>
<td>146.3</td>
<td>233.7</td>
<td></td>
</tr>
<tr>
<td>AL</td>
<td>30% food</td>
<td>Different</td>
<td>5000</td>
<td>7</td>
<td>36.1</td>
<td>60.6</td>
<td>-91.0</td>
<td>-36.5</td>
<td>-15.5</td>
<td>4.5</td>
<td>25.1</td>
<td>55.9</td>
<td>106.0</td>
<td>143.8</td>
<td>248.0</td>
<td></td>
</tr>
<tr>
<td>AL</td>
<td>None</td>
<td>True soil</td>
<td></td>
<td></td>
<td>53.8</td>
<td>41.4</td>
<td>21.6</td>
<td>23.2</td>
<td>24.1</td>
<td>26.1</td>
<td>36.0</td>
<td>67.0</td>
<td>105.9</td>
<td>149.0</td>
<td>194.0</td>
<td></td>
</tr>
<tr>
<td>SI</td>
<td>None</td>
<td>Same</td>
<td>5000</td>
<td>7</td>
<td>55.9</td>
<td>53.3</td>
<td>-13.1</td>
<td>2.7</td>
<td>9.5</td>
<td>22.3</td>
<td>41.0</td>
<td>73.2</td>
<td>120.2</td>
<td>163.4</td>
<td>254.4</td>
<td></td>
</tr>
<tr>
<td>SI</td>
<td>None</td>
<td>Different</td>
<td>5000</td>
<td>7</td>
<td>55.1</td>
<td>54.2</td>
<td>-11.5</td>
<td>2.2</td>
<td>9.4</td>
<td>21.2</td>
<td>39.8</td>
<td>71.7</td>
<td>119.7</td>
<td>162.0</td>
<td>249.8</td>
<td></td>
</tr>
<tr>
<td>SI</td>
<td>None</td>
<td>Constant</td>
<td>5000</td>
<td>7</td>
<td>55.7</td>
<td>53.7</td>
<td>-11.8</td>
<td>2.5</td>
<td>9.2</td>
<td>22.2</td>
<td>40.8</td>
<td>73.8</td>
<td>121.5</td>
<td>159.4</td>
<td>256.0</td>
<td></td>
</tr>
<tr>
<td>SI</td>
<td>30% food</td>
<td>Same</td>
<td>5000</td>
<td>7</td>
<td>38.1</td>
<td>53.6</td>
<td>-40.4</td>
<td>-16.8</td>
<td>-8.6</td>
<td>5.0</td>
<td>24.3</td>
<td>55.9</td>
<td>103.2</td>
<td>144.7</td>
<td>236.8</td>
<td></td>
</tr>
<tr>
<td>SI</td>
<td>30% food</td>
<td>Different</td>
<td>5000</td>
<td>7</td>
<td>37.1</td>
<td>54.1</td>
<td>-37.7</td>
<td>-17.7</td>
<td>-9.1</td>
<td>4.3</td>
<td>22.7</td>
<td>53.9</td>
<td>101.9</td>
<td>143.9</td>
<td>228.2</td>
<td></td>
</tr>
<tr>
<td>SI</td>
<td>None</td>
<td>True soil</td>
<td></td>
<td></td>
<td>53.8</td>
<td>41.4</td>
<td>21.6</td>
<td>23.2</td>
<td>24.1</td>
<td>26.1</td>
<td>36.0</td>
<td>66.9</td>
<td>105.7</td>
<td>149.1</td>
<td>194.2</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a} An empirical soil ingestion distribution based on Amherst and Anaconda subjects (excluding the pica subject), with normally distributed daily variation in estimates based on standard deviation modeled from a simple linear regression on the mean (intercept = 63.5, slope = 0.834).
ing the impact of assumptions inherent in the estimates, and identifying important issues for future study designs. We briefly summarize the strengths of this approach, and comment on limitations and directions for future work.

The simulation model developed in this paper builds on data collected in four mass-balance soil ingestion studies (Calabrese et al. 1989; Davis et al. 1990; Calabrese et al. 1997a,b) conducted in the United States, and summarizes the basic distributions of trace element concentrations in food and soil, along with the distribution of freeze dried food weight and transit time from food to fecal samples in these studies. The similarity of results from the four mass-balance studies on many parameters for the distributions provides a broad basis for generalizability of the results. Inclusion of estimates of between study variance for each parameter allows examination of study heterogeneity.

Concentration distributions of the trace elements in food are likely to be similar in future studies with similar protocols. Although it is possible that new protocols may be developed that are successful in minimizing targeted food items high in the trace elements, the distributions characterized in this study will place bounds on the gains that can be expected from such efforts.

The distribution of freeze dried food weight characterized for the simulation is based on observed freeze-dried weights for 241 children between the ages of 1 and 7 years on over 1000 days. Age is accounted for in the distribution, and hence the distribution can simulate realistic food ingestion for children in that age range. Since this age range covers ages where soil ingestion is thought to be of most concern, the results have broad relevance.

The distribution of trace elements in soil is based on soil from the states of Massachusetts, Montana, and Washington. There were fairly large differences (as much as 25%) in the average concentration of trace elements in soil between states, although the distributions of soil concentrations for different soil samples in a state were similar. This difference may be important when evaluating future study designs. Areas with higher trace element concentrations in soil may result in more reliable soil ingestion estimates.

A valuable feature of the simulation construction is the ability to mimic the food/soil passage from ingestion to fecal samples. Using fecal sample interval data from three daily soil ingestion studies (Calabrese et al. 1989; Calabrese et al. 1997a,b), the transit time distribution that has been observed can be simulated using a simple Poisson distribution. Beginning the simulation prior to the study start allows possible lumping of food and fecal samples on the first fecal collection day to be simulated. In addition, use of the transit time interval on subsequent days allows fecal sample collections to be lumped in a manner consistent with what has been observed in past studies. While the simulation of food passage time and fecal aggregation broadly characterizes what has been observed, it should be noted that the simulation does not account for age differences in food transit time, and considers fecal sample periods as discrete day intervals. A day's fecal sample can only represent discrete day's samples, and not a portion of the fecal amount on a given day.

A broad range of hypothetical soil ingestion distributions can be used in the simulation. We have selected soil ingestion distributions to illustrate the impact of the mass-balance study design, as well as to bound plausible distributions based on
previous soil ingestion studies. There is no limitation on the type of soil ingestion distribution that can be postulated. It is important to note that the objective of simulating realistic data from a mass-balance soil ingestion study with a known (and artificial) soil ingestion distribution is to refine the study design and estimation methods for capturing the artificial distribution, and not on the artificial soil ingestion distribution itself.

This study identifies the bias on soil ingestion distribution estimates that will result from various assumptions and characteristics of soil ingestion studies. The results apply to soil ingestion distribution estimates based on the cumulative distribution from a tabulation of simple average soil ingestion estimates for subjects over the study period. This estimation strategy has been the dominant strategy used in mass-balance soil ingestion studies. One of the most striking conclusions of this investigation is the impact of the study duration. Estimated soil ingestion distributions based on short-term (4- or 7-day) study designs will have variances that are too large, with biases between 66% and 186%. The biased variance will result in upper percentiles of the estimated soil ingestion distribution being too high. The extent of the bias will depend on the actual true soil ingestion distribution and the duration of the study. In a 4-day mass-balance study, estimates of the 95th percentile true soil ingestion may be positively biased by more than 100%. While falling short of identifying the extent of the bias in current soil ingestion study estimates (since the true soil ingestion distribution is unknown), the results make clear that apart from other problems, previous estimates of the cumulative soil ingestion distribution are upper bounds on actual long-run soil ingestion.

The simulation studies place bounds on the sensitivity of the estimated soil ingestion distribution to two other assumptions in the mass-balance approach. First, properly accounting for soil concentrations in a child’s play area is not a biasing factor when estimating the soil ingestion distribution. Differential soil concentrations by play area may affect the reliability of subject specific estimates of average soil ingestion, but will not bias the results. Second, absorption of trace elements from food will uniformly bias the estimated soil ingestion distribution down. The extent of the bias will be proportional to the degree of absorption, and can be estimated based on the estimated average trace element intake from food (expressed as the soil equivalent amount). For 30% absorption of trace elements from food, the negative bias on the Al and Si soil ingestion distribution is likely to be less than 20 mg/d.

The simulation study offers an objective way of evaluating different strategies for estimating the distribution of soil ingestion, and interpreting the estimates. The study is limited to consideration of two trace elements, Al and Si. These two elements were included in all mass-balance soil ingestion studies, have high concentrations in soil, and appear to have been minimally ingested from non-food, non-soil sources. Inclusion of other trace elements (such as Ce, La, Nd, Y, and possibly Zr) is desirable since these elements have been collected in some studies. However, possible ingestion of non-food, non-soil sources is likely to be a concern for Ce, La, and Nd, and limitations in chemical analysis a concern for Zr. For this reason, simulating realistic ingestion and fecal output of these trace elements is more complex.
Biases in Soil Ingestion Studies

There are other important limitations to this investigation. These limitations arise from differences between the assumptions underlying the simulations and the available data used to set parameters for the simulations. The simulations and conclusions are valid under the assumption that the simulated values of soil ingestion are the result of a simple random sample of soil ingestion days for a simple random sample of subjects. In practice, studies of soil ingestion have used available subjects (volunteers or subjects who agree to participate), and specified a set of consecutive study days for observation. This fact implies that soil ingestion parameters estimated from the study data will likely not match parameters for soil ingestion in a population. Specifically, variability between subjects in soil ingestion studies reflects the variability evident for these subjects, but not necessarily children in general. One can readily speculate as to reasons why estimates from the available data are too large or too small. Additional studies are needed to further define the adequacy of the simulation parameters.

A particular limitation in interpretation of the simulations relates to the sample days. The simulations were conducted assuming that the soil ingestion observed days are a simple random sample of possible days. In contrast, soil ingestion studies have measured soil ingestion in consecutive periods only in the summer or early fall. This time period clearly is not comparable to days in the winter (especially in the northern U.S.), and the days are clearly not random. As a result, the simulation results do not capture the real world seasonal variability. Furthermore, soil ingestion studies have measured subjects on days only during a single year. There is currently no data to quantify the degree of tracking of soil ingestion across subjects from year to year. Once again, additional data are needed to address this limitation.

The focus of this study is on biasing factors for the distribution of simple average estimates of soil ingestion. We have limited the discussion to three possible biasing factors (study duration, play area misspecification, and trace element absorption from food). Other factors that may affect the soil ingestion estimates (including absorption of trace elements from soil, fecal sample loss, and non-food, non-soil trace element ingestion) were outside the scope of this investigation. The simulation framework lends itself to development and evaluation of methods that may minimize problems resulting from such factors. This is a topic for future investigation.

The results in this study describe properties of estimates based on simulations of a study of 5000 subjects. The large number of subjects was used to obtain stability in the cumulative distribution of soil ingestion across subjects. This stability was necessary to identify biases in the distribution, and not confuse these biases with possible differences due to reliability. No soil ingestion study to date has included such a large number of subjects. Instead, estimates the average soil ingestion distributions have been based on from 64 to 101 subjects followed from 4 to 7 days. The distribution of these soil ingestion estimates will include both bias and uncertainty. We plan future investigations to characterize this uncertainty.

Finally, this study has focused on factors that cause bias in the estimated soil ingestion distribution formed from simple average daily soil ingestion estimates per subject. Other estimation strategies, such as use of best linear unbiased predictors, have not be explored. It is likely that a combination of corrections for biases, and alternative estimators, will provide the capacity for improved soil ingestion estimates in the future.
ACKNOWLEDGMENT
This research was supported by a contract from the U.S. EPA Superfund program. The manuscript benefits from the extensive comments and suggestions of one of the reviewers.

TABLE SOURCES

Table 1. nkmz2.doc
Table 2. nk99p46.sas
Table 3. nk99p56.sas; nk99p52.sas
Table 4. nk99p52.sas

Row 1: From second row in Table 6 (for Al) of nkmz3.doc.
Row 2: From al81.txt created by nkmz99p81.sas.
Row 3: From al83.txt created by nkmz99p83.sas.
Row 4: From al80.txt created by nkmz99p80.sas.
Row 5: From al82.txt created by nkmz99p82.sas.
Row 6: From seventh row in Table 6 (for Al) of nkmz3.doc.
Row 7: From second row in Table 6 (for Si) of nkmz3.doc.
Row 8: From si81.txt created by nkmz99p81.sas.
Row 9: From si83.txt created by nkmz99p83.sas.
Row 9: From second row of Table 6 (for Si, 30% food absorption) of nkmz3.doc
Row 10: From si82.txt created by nkmz99p82.sas
Row 11: From seventh row in Table 6 (for Si) of nkmz3.doc.

REFERENCES
Biases in Soil Ingestion Studies


StataCorp 1997. Stata Statistical Software: Release 5.0 Stata Corporation, College Station, TX, 64