Daily Soil Ingestion Estimates for Children at a Superfund Site

Edward J. Stanek III* and Edward J. Calabrese

Ingestion of contaminated soil by children may result in significant exposure to toxic substances at contaminated sites. Estimates of such exposure are based on extrapolation of short-term-exposure estimates to longer time periods. This article provides daily estimates of soil ingestion on 64 children between the ages of 1 and 4 residing at a Superfund site; these values are employed to estimate the distribution of 7-day average soil ingestion exposures (mean, 31 mg/day; median, 17 mg/day) at a contaminated site over different time periods. Best linear unbiased predictors of the 95th-percentile of soil ingestion over 7 days, 30 days, 90 days, and 365 days are 133 mg/day, 112 mg/day, 108 mg/day and 106 mg/day, respectively. Variance components estimates (excluding titanium and outliers, based on Tukey’s far-out criteria) are given for soil ingestion between subjects (59 mg/day), between days on a subject (95 mg/day), and for uncertainty on a subject-day (132 mg/day). These results expand knowledge of potential exposure to contaminants among young children from soil ingestion at contaminated sites. They also provide basic distributions that serve as a starting point for use in Monte Carlo risk assessments.

KEY WORDS: Soil ingestion; Monte Carlo risk assessment; children; Superfund site; exposure assessment

1. INTRODUCTION

The amount of soil ingested by children at contaminated sites may potentially impact their health. Estimates of these amounts are needed in order to establish health risk–based site-specific cleanup levels. Reliable estimates of soil ingestion by children, however, have been difficult to obtain due to methodological limitations, the small amount of soil ingested, and the variability in the ingestion behavior between children and over time.

Early semiquantitative estimates of daily soil ingestion were based on hand wipes and estimates of hand-to-mouth behavior among young children. Recent estimates of soil ingestion are based on a mass-balance trace element, with three large mass-balance studies in the United States conducted on children in Amherst, Massachusetts, Washington State, and Anaconda, Montana (a Superfund site). Each of these mass-balance studies have included multiple trace elements, subtracted trace element amounts in food from fecal amounts, and produced soil ingestion estimates for each trace element based on totals over the individual study periods.

The studies have differed in some other respects. The study in Washington State included adjustments for absorption of trace elements (through urine measures); the Anaconda, Montana (hereafter referred to as “Anaconda”), study based soil ingestion estimates on soil particle size <250 μm (as opposed to 2 mm). Both the Amherst, Massachusetts, and Anaconda studies included daily estimates of trace elements in food and fecal samples for the study periods (8 days and 7 days, respectively), while the Washin-
ton State study combined 4 days of food and fecal samples to obtain a single average soil ingestion estimate per element per child.

The magnitude of the trace element specific estimates of soil ingestion in the studies have differed. This has prompted development of strategies, such as the best tracer method (BTM),\(^{(11,15–17)}\) for obtaining a common soil ingestion estimate over the trace elements. The BTM selects trace element estimates to combine using elements with low soil equivalent amounts in food.

As of this writing, there has been only one attempt to characterize variability in soil ingestion between days for a given subject, and between subjects.\(^{(18)}\) Estimates of these sources of variability are important to distinguish short-term exposure\(^{(19)}\) from longer term average exposure, and to predict exposure over longer time periods than the typical study design time period.\(^{(20)}\) Estimates of variability in soil ingestion between days and subjects (as opposed to the uncertainty in the amount of soil ingested on a particular subject-day) are particularly relevant for risk assessment based on Monte Carlo modeling.\(^{(21)}\) Such models simulate average soil ingestion over chronic (long-term) exposure periods. Assuming a lognormal distribution for soil ingestion estimates from the Amherst study, data were extrapolated for children over an entire year.\(^{(18)}\) While this procedure is highly speculative due to the number of assumptions required, 50% of children were estimated to ingest 75 mg/day or less on average over 1 year. The same computations predicted that the majority of children (63%) would ingest 1 g or more of soil at least once in a year.

The purpose of this study is to provide estimates of variability between days and subjects in soil ingestion among children living on a Superfund site. Daily soil ingestion estimates are constructed for children in the Anaconda children’s study using a methodology similar to that of the Amherst soil ingestion study.\(^{(16)}\) Daily soil ingestion estimates are used to predict average soil ingestion for children over a longer time period (30 days, 90 days, and 365 days); a discussion of the use of these estimates in Monte Carlo risk assessments follows.

2. MATERIALS AND METHODS

Anaconda, Montana, is the site of a copper ore-processing plant that operated between 1884 and 1980. As a result of the processing operation, arsenic (As) contaminated the soil over a broad area southeast of the plant that included the town of Anaconda. Although the processing plant is no longer in operation, concern remains over the potential adverse effects of As exposure from casual ingestion of the soil, particularly among young children.\(^{(11,21)}\) In brief, a mass-balance study of soil ingestion in a stratified random sample of 64 children aged 1–4 years residing in Anaconda, Montana, was conducted to evaluate possible exposure to soil through ingestion.\(^{(11)}\) Trace element amounts in food, fecal samples, and soil were collected for eight trace elements—aluminum (Al), cerium (Ce), lanthanum (La), neodymium (Nd), silicon (Si), titanium (Ti), yttrium (Y), and zirconium (Zr)—over an 8-day period (Days 0–7). On Days 0–6, duplicate of the ingested food was collected; on Days 1–7, fecal samples were collected. This provided a potential total of seven food and seven fecal samples per subject. Soil ingestion estimates were computed simultaneously using the eight trace elements for up to 7 days for each subject. For each element for a subject-day, soil ingestion was estimated by determining the difference between total food intake and fecal output and dividing the difference by the concentration of the trace element in soil. Soil concentrations were measured in soil of <250 \(\mu\)m particle size for 62 subjects, and <2 mm for two subjects where the smaller particle size soil was unavailable.

A 28-h transit time from food consumption to fecal production was assumed; the trace element amount from food in fecal samples was calculated for 331 fecal sample-days, representing 427 days of intake. Daily estimates of soil ingestion were determined in a similar manner to the Amherst study.\(^{(18)}\) More details are available in a technical report from the authors.\(^{(22)}\) Fecal sample data were suspect on 19 subject-days where fecal samples were reported to have been partially or completely missed. Days with no fecal output for a child was considered evidence of irregular fecal production by the child; soil ingestion was calculated for each day accordingly.

The possibility that fecal samples were missed and not reported as missing by parents was evaluated by comparing the freeze-dried fecal/food weight ratios by children’s age from the Amherst and Anaconda studies. The comparison revealed lower relative fecal weights for children 30 months of age or older among the Anaconda subjects. The sensitivity of such possible underreporting was evaluated by identifying subjects with fecal/food weight ratios less than one standard deviation below the mean ratios among similar Amherst study children (30+ months of age), and then recomputing soil ingestion estimates after excluding these subjects.
2.1. Reducing Uncertainty

Previous attempts to improve the accuracy of soil ingestion estimates have focused on identifying and excluding potentially biased or highly variable trace element estimates (using outlier criteria); selecting more reliable trace element soil ingestion estimates based on low food/soil ratios; and using soil concentrations in appropriate particle-size fractions.\(^{(14,17,18)}\) Outlier criteria used in the Amherst study resulted in exclusion of 37.5% of the data.\(^{(18)}\) In the Anaconda study, additional trace elements (Ce, La, and Nd) were included that were anticipated to occur minimally in food and thus be more reliable. These trace elements, however, occur in very small amounts in soil and are thus prone to error due to small nonsoil, nonfood ingestion amounts (referred to as “source error”). Such apparent source error was evident for these elements and motivated development of alternative strategies for outlier identification than used in previous studies. Ingestion of nonfood, nonsoil trace elements will positively bias estimates of soil ingestion since the trace element will erroneously be attributed to ingested soil. On the other hand, if a high trace element estimate of soil ingestion is mistakenly identified as an outlier, true soil ingestion will be underestimated. For this reason, outlier identification was made with caution. Trace element estimates for a subject-day that exceeded the third quartile (or were below the first quartile) by more than three times the interquartile range were identified as outliers. These outliers correspond to far-out values as described by Tukey\(^{(23)}\) in a box plot (p. 44).

Uncertainty in soil ingestion occurs if the concentration of trace element in the sampled soil does not correspond to the trace element concentration in ingested soil. The Anaconda study evaluated trace element concentrations in soil of different particle sizes. Comparisons of soil ingestion estimates based on concentrations in different particle size soil revealed that uncertainty was minimized when soil concentrations were based on soil <250 μm in diameter.\(^{(14)}\) For this reason, soil concentrations used in the present evaluations were based on particle sizes <250 μm in diameter.

High and variable trace element exposure from dietary ingestion of trace elements from other nonfood/nonsoil sources can produce highly variable soil ingestion estimates. Past soil ingestion studies have eliminated soil ingestion estimates based on the trace elements barium, manganese, and vanadium for such reasons. Trace elements were reviewed to examine whether uncertainty (estimated by the standard deviation of the difference between a single trace element estimate and the average of the other trace element estimates for a subject-day) was unusually large for any trace element. This review identified Ti as uncertain, with other elements having lower levels of uncertainty. Subsequent examination was confined to trace element estimates excluding Ti.

3. RESULTS

Using the methods outlined above, soil ingestion estimates were calculated for each of the eight trace element for 331 days using fecal samples collected on the 64 children in the Anaconda study (Table I). In total, the fecal samples spanned 427 ingestion days, or 95% of the possible days in the 7-day study. For one subject on two days, analytic results were available only for the trace elements Al, Si, and Ti.

Trace element ingestion from food was relatively low when expressed as an equivalent amount in soil. For 75% of the study days, the trace element ingestion from food per day was <100 mg of equivalent soil for six of the eight trace elements, exceptions being Ti and Zr. The three rare earth elements Ce, La, and Nd had the lowest amounts in food, when expressed as soil equivalents.

The uncertainty of each trace element daily soil ingestion estimate was evaluated by computing the standard deviation of the difference between the trace element estimate of soil ingestion, and the average of other trace element estimates for a subject-day. Estimates of the standard deviation for Ti were 20 times larger than those for any other trace element. Due to the large uncertainty for soil ingestion based on Ti, this trace element was eliminated from further consid-

<table>
<thead>
<tr>
<th>Number of days</th>
<th>Total days with fecal samples</th>
<th>Total days with soil ingestion estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. children</td>
<td>Percent</td>
<td>No. children</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>7</td>
<td>16</td>
<td>47</td>
</tr>
</tbody>
</table>

Total 64 100.0 64 100.0

Source: Stanek\(^{(39)}\)
eration. After eliminating trace element estimates for Ti, the Tukey outlier criteria identified 18 element-subject-days (of 2,984, or 0.45%) as outliers (compared with 31.9% that would have been identified as outliers using the Amherst outlier criteria\(^\text{18}\)).

Estimates of the average daily soil ingestion over the 7-day study period were determined for each child in a manner similar to that used in the Amherst study.\(^\text{18}\) First, the median trace element specific estimate of soil ingestion was computed for each subject-day. For each subject, the median, mean, and standard deviation of the estimates were calculated over the 7 study days. These calculations were repeated based on the average estimated soil ingestion per day (Table II) for comparison with the Amherst study results, and for reference with the mixed model analyses.

The results in Table II summarize estimates of soil ingestion over a 7-day period for the 64 subjects. The first two rows in Table II are based on estimates corresponding to the median from among the seven trace element estimates for each subject-day. The third and fourth rows in Table II are based on estimates corresponding to the mean from among the seven trace element estimates for each subject-day. Since the distribution of trace element estimates was, in general, positively skewed on a given subject-day, the estimates in the last two rows of Table II are slightly larger than the comparable estimates in the first two rows.

The estimates in the first row of Table II include those of the median soil ingestion over the 7 days for a given subject. For example, half of the children ingested \(\leq 8\) mg/day on half of their study days. Similarly, 90% of the children ingested \(\leq 82\) mg/day on at least half of their study days. Comparable estimates from the Amherst study (see Table 4\(^\text{18}\)) were 13 mg/day and 126 mg/day, respectively. The third and fourth rows in Table II summarize the distribution of the average (over 7 days) daily soil ingestion for the 64 children. Half the children ingested \(\leq 17\) mg/day, on average, over the 7 days. Ninety percent of the children ingested, on average, \(\leq 111\) mg/day over the 7-day period. Comparable estimates from the Amherst study (see Tables 4 and 5\(^\text{18}\)) were 45 mg/day and 186 mg/day, respectively. Similar interpretations can be applied to the last two rows in Table II.

Figure 1 presents the cumulative soil ingestion distribution using the median subject-day values and the mean over days. Included in Fig. 1 are upper and lower 95% confidence limits based on order statistics for quantiles.\(^\text{24}\) The changing width of the CL band reflects variability due to sampling of children, but does not account for other sources of variability and uncertainty. More robust interval estimates are the focus of current investigations.\(^\text{25–27}\)

Of particular interest are characteristics of children in the upper percentile of the soil ingestion distribution. Detailed estimates are presented for the child (Subject 16) who had the highest average soil ingestion per day based on the median (173 mg/day) of daily trace element estimates (Table III). The high mean soil ingestion for this subject was due to ingesting 600–700 mg of soil on Day 7 of the study.

### 3.1. Estimating Soil Ingestion over Longer Time Periods

The variance components based on restricted maximum likelihood estimates from nested mixed models

---

**Table II.** Distribution of Daily Soil Ingestion (md/day) Over 7 Days in Anaconda, Montana, Based on 7 Trace Elements (Excluding Titanium) Assuming a 28-hour Food Transit Time with Corresponding Amherst, Massachusetts Estimates

<table>
<thead>
<tr>
<th>Study</th>
<th>Over elements within a day</th>
<th>N</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>P25</th>
<th>Median</th>
<th>P75</th>
<th>P90</th>
<th>P95</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anc.</td>
<td>Median</td>
<td>64</td>
<td>13</td>
<td>49</td>
<td>−14</td>
<td>8</td>
<td>30</td>
<td>82</td>
<td>107</td>
<td>136</td>
</tr>
<tr>
<td>Amh.</td>
<td>Median</td>
<td>64</td>
<td>32</td>
<td>0</td>
<td>13</td>
<td>50</td>
<td>126</td>
<td>138</td>
<td>185</td>
<td></td>
</tr>
<tr>
<td>Anc.</td>
<td>Mean</td>
<td>64</td>
<td>31</td>
<td>56</td>
<td>−3</td>
<td>17</td>
<td>53</td>
<td>111</td>
<td>141</td>
<td>219</td>
</tr>
<tr>
<td>Amh.</td>
<td>Mean</td>
<td>64</td>
<td>179</td>
<td></td>
<td>10</td>
<td>45</td>
<td>88</td>
<td>186</td>
<td>208</td>
<td>7,703</td>
</tr>
<tr>
<td>Anc.</td>
<td>Mean</td>
<td>64</td>
<td>14</td>
<td>59</td>
<td>−14</td>
<td>4</td>
<td>26</td>
<td>120</td>
<td>128</td>
<td>151</td>
</tr>
<tr>
<td>Anc.</td>
<td>Mean</td>
<td>64</td>
<td>36</td>
<td>72</td>
<td>−7</td>
<td>16</td>
<td>72</td>
<td>151</td>
<td>160</td>
<td>283</td>
</tr>
</tbody>
</table>

*Note:* Amherst, Massachusetts, estimate data are from Stanek and Calabrese.\(^\text{18}\) P25 = 25th percentile; P75 = 75th percentile; P90 = 90th percentile; P95 = 95th percentile.

*Estimates are from column labeled “Overall” in Table 4, p. 281.\(^\text{18}\)*

*Estimates are from column labeled “Overall” in Table 5, p. 281.\(^\text{18}\)*

Source: Stanek.\(^\text{30}\)
fit using the mixed models procedure in the Statistical Analysis System (SAS)\(^{(28)}\) are summarized in Table IV for the various outlier criteria. The first two rows in Table IV illustrate the impact of Ti and exclusion of outliers using the Tukey\(^{(23)}\) criteria. The variance component for uncertainty dramatically decreases when soil ingestion estimates based on Ti are eliminated, and is moderately reduced when the few additional outliers based on Tukey’s criteria are eliminated.

The last three rows in Table IV are included for comparison with the Amherst study.\(^{(18)}\) All three variance components are impacted by exclusion of trace element estimates identified by the Amherst outlier criteria (row 4 in Table IV). Since over 30% of the data are excluded, however, this strategy may underestimate both variability and uncertainty. The last two rows in Table IV provide a direct comparison of results from the two studies. The variance components estimates based on the outlier criteria “Exclude titanium and Tukey far-out” are considered to be best. These estimates of variance components are used to determine soil ingestion estimations for 30 days, 90 days, and one year (365 days).

When there is uncertainty or day-to-day variability in soil ingestion for subjects, the distribution of the estimated average daily soil ingestion (for subjects) over a short period is more spread out than the actual soil ingestion over a longer period.\(^{(29)}\) Thus, the 95th-percentile soil ingestion based on observed 1-day soil ingestion estimates will be higher than what one would expect for the 95th-percentile of the true 1-day soil ingestion distribution. Similarly, the estimate of the 95th-percentile soil ingestion distribution based on average soil ingestion estimates over 7 days would be expected to be higher than the true 95th percentile of the 7-day average soil ingestion for subjects. This narrowing of the distribution in estimates when daily variability and uncertainty is reduced is called “reversion to the mean.” The phenomena occurs in many settings, for example, with baseball batting averages that are more spread out between players early in the season (more uncertainty) compared to the end of the season (less uncertainty).\(^{(29)}\)

The extent to which estimates are expected to be shrunk toward the mean can be estimated from variance components. The shrinkage is usually expressed in terms of a shrinkage constant, \(k\), which ranges from zero to one. The extent of shrinkage is evaluated by multiplying \(k\) by the difference between an estimate

---

**Fig. 1.** Cumulative distribution of 7-day average soil ingestion estimates (using the median of the estimates for aluminum, cerium, lanthanum, neodymium, yttrium, and zirconium per day) and 95% CI around quantiles for 64 children from the Anaconda study.
and the mean. For example, suppose seven trace elements are used to estimate soil ingestion over a 7-day period for each subject, resulting in a mean soil ingestion estimate of 31 mg/day, with an estimate for the 95th-percentile soil ingestion of 141 mg/day (as in the third row of Table II). The estimate of the 95th-percentile soil ingestion is not the same as the true 95th-percentile soil ingestion, since uncertainty is present. The shrinkage constant attempts to account for the uncertainty, and depends on the relative variance estimates.

The variance components from row 3 of Table IV are used to illustrate how the constant k is determined. Let the estimates of the variance components be represented by $S$ (for subjects), $D$ (for days), and $E$ (for error). The extent of shrinkage of an estimator can be determined by a simple formula. The formula consists of a numerator, representing the variance in true soil ingestion based on averages over the desired time period ($T$), and a denominator, corresponding to the variance of the soil ingestion estimate actually realized. Let us define

$$T = \text{the time period for which average soil ingestion is desired},$$

$$n = \text{the number of days over which estimates of soil ingestion have been made},$$

$$m = \text{the number of replications (trace elements) used in calculating each daily estimate}.$$

Then the shrinkage constant is determined by the formula: $k = (S + D/T)/(S + D/n + E/nm)$. Using these expressions with $n = m = 7$ (corresponding to the Anaconda study protocol), and taking the mean as 31 mg/day, and the observed 95th-percentile soil ingestion as 141 mg/day (as in the third row of Table II), the estimate of the true average 95th-percentile soil ingestion for children over time periods of $T = (7$ days, 30 days, 90 days, and 365 days) are 133 mg/day, 112 mg/day, 108 mg/day, and 106 mg/day, respectively.

The sensitivity of these estimates was evaluated

### Table III. Detailed Soil Ingestion Estimates (mg/day) for Subject 16 with Highest Soil Ingestion in Anaconda, Montana, Assuming a 28-hour Food Transit Time

<table>
<thead>
<tr>
<th>Day</th>
<th>Mean</th>
<th>Median</th>
<th>Aluminum</th>
<th>Cerium</th>
<th>Lanthanum</th>
<th>Neodymium</th>
<th>Silicon</th>
<th>Yttrium</th>
<th>Zirconium</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>46</td>
<td>76</td>
<td>5</td>
<td>76</td>
<td>90</td>
<td>297</td>
<td>21</td>
<td>94</td>
<td>−259</td>
</tr>
<tr>
<td>2</td>
<td>122</td>
<td>96</td>
<td>34</td>
<td>50</td>
<td>96</td>
<td>444</td>
<td>−7</td>
<td>113</td>
<td>126</td>
</tr>
<tr>
<td>3</td>
<td>49</td>
<td>37</td>
<td>5</td>
<td>37</td>
<td>64</td>
<td>150</td>
<td>−4</td>
<td>74</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>49</td>
<td>37</td>
<td>5</td>
<td>37</td>
<td>64</td>
<td>150</td>
<td>−4</td>
<td>74</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>136</td>
<td>162</td>
<td>−215</td>
<td>50</td>
<td>204</td>
<td>453</td>
<td>68</td>
<td>228</td>
<td>162</td>
</tr>
<tr>
<td>6</td>
<td>123</td>
<td>100</td>
<td>50</td>
<td>46</td>
<td>100</td>
<td>270</td>
<td>79</td>
<td>177</td>
<td>136</td>
</tr>
<tr>
<td>7</td>
<td>649</td>
<td>702</td>
<td>722</td>
<td>415</td>
<td>510</td>
<td>928</td>
<td>702</td>
<td>780</td>
<td>484</td>
</tr>
<tr>
<td>Median</td>
<td>122</td>
<td>96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>167</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Stanek.198

### Table IV. Mean and Variance Component Estimates (Standard Error, SE) of Daily Soil Ingestion Assuming Various Outlier Criteria for 64 Anaconda or Amherst Children (mg/day)

<table>
<thead>
<tr>
<th>Outlier criteria</th>
<th>Anaconda, MT</th>
<th>Amherst, MA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SE)</td>
<td>S = Subject variance (SE)</td>
</tr>
<tr>
<td>None</td>
<td>−35 (41)</td>
<td>66,989 (18,661)</td>
</tr>
<tr>
<td>Exclude titanium only</td>
<td>37 (9)</td>
<td>3,380 (935)</td>
</tr>
<tr>
<td>Exclude titanium and Tukey far-out</td>
<td>37 (9)</td>
<td>3,466 (932)</td>
</tr>
<tr>
<td>Exclude titanium and Amherst criteria</td>
<td>28 (7)</td>
<td>1,923 (530)</td>
</tr>
<tr>
<td>Include aluminum, silicon, yttrium, zirconium</td>
<td>5 (9)</td>
<td>3,621 (998)</td>
</tr>
<tr>
<td>Amherst, MAa</td>
<td>Include aluminum, silicon, yttrium, zirconium</td>
<td>32 (11)</td>
</tr>
</tbody>
</table>

*Based on estimates of daily soil ingestion of up to 8 days on 63 subjects, excluding Subject 851 (the pica subject).
to the possibility that some children 30+ months of age may have had underreported fecal samples. A total of 19 children were identified as having freeze-dried fecal/food weight ratios ≤1 standard deviation below comparable ratios reported by the same-age-group children in the Amherst soil ingestion study.\(^{18}\) Excluding these 19 children (mean age, 37.8 months; standard deviation, 6.5) soil ingestion estimates and variance components were recalculated comparable to Tables II and IV among the remaining 45 children (median age, 25.1 months; standard deviation, 10.2). From these results, the comparable mean and 95th-percentile estimates were 40 mg/day and 150 mg/day, respectively, while the variance components estimates were \(S = 4,167, D = 10,138; \) and \(E = 18,670.\) Using these estimates for the 45 children, the estimate of the true average 95th-percentile soil ingestion for children over time periods of \(T = (7 \text{ days}, 30 \text{ days}, 90 \text{ days}, \text{and } 365 \text{ days})\) are 143 mg/day, 123 mg/day, 119 mg/day, and 117 mg/day, respectively.

4. DISCUSSION

Mass-balance soil ingestion studies such as Anaconda offer the best opportunity to understand and quantify ingestion of soil by children. Such studies have a key advantage over other possible indirect approaches in estimating exposure due to soil ingestion because the methodology has been validated among adults.\(^{30,31}\) A few mass-balance studies\(^{6,10,11}\) have been conducted that provide estimates of variability and uncertainty in soil ingestion in addition to typical daily soil ingestion levels. Incorporating daily food and fecal values and the use of multiple trace elements provides insight on the dynamics of ingestion and the mass-balance system. The Anaconda study has expanded the number of possible trace element estimates and focuses attention on soil of smaller particle size. Daily estimation of soil ingestion offers the best opportunity to take advantage of this information by reducing uncertainty, as well as characterizing variability.

Table II summarizes estimates of the average daily amount of soil ingested over a 7- (or 8-) day period for two of these studies. Children in the Anaconda study lived on a recognized Superfund site that was undergoing cleanup, while children in Amherst lived in an academic residential community. Data for both studies were collected over short periods in late summer/early fall. It is important to note that all of the estimates in Table II characterize “typical” soil ingestion amounts (based either on the median daily ingestion or the average daily ingestion over the study period). Such estimates correspond to a measure of short-term average exposure.

Figure 1 summarizes the cumulative distribution of soil ingestion based on average amount of soil ingested per day (over a 7-day period). The estimates used in this figure are based on median trace element specific estimates for a subject-day. The 95th-percentile CLs depicted in Fig. 1 are based on order statistics, and account only for sampling variability due to selection of the study children, not uncertainty in the estimates on a subject-day or variability due to selection of different sets of seven consecutive days. For this reason, the CL estimates are too narrow. While this limitation in the CL estimates is recognized, standard methods for correcting for them are not available. Future research based on simulation studies is planned to obtain more robust CL estimates for the cumulative soil ingestion distribution.\(^{25–27}\)

Estimates of the distribution of longer term average soil ingestion are expected to be narrower, with 95th-percentile estimates being as much as 25% lower than those given in Table II. Such average exposure may be most appropriate when considering the impact of chronic (i.e., 1-year) exposure to lead-contaminated soil on blood-lead levels. In contrast, the distribution of soil ingestion on single days can be expected to be broader than the distribution summarized in Table II, and may be more relevant for acute exposures.\(^{19}\)

Strategies to estimate these longer term averages make use of variance components estimates given in Table IV. The day-to-day variance \((D)\) and uncertainty \((E)\) can be interpreted as the average variance over subjects in the study. The variance estimates from the third row of Table IV are thought to provide the best estimate of variability in soil ingestion and uncertainty. These estimates are relevant for Monte Carlo analyses. It should be noted that in the sensitivity analysis, after excluding children 30+ months of age who may have underreported fecal samples, the variance components were larger by 20%, 10%, and 8% for \(S, D,\) and \(E,\) respectively. These modest increases in variance may be due in part to the younger age of the children remaining after the exclusions. It is also possible that the \(D\) variance is not homogeneous between children, with those that have higher soil ingestion estimates also having higher \(D\) variances. Such considerations are beyond the scope of the present analyses.

The apparent dramatic reduction in uncertainty that occurs when using the Amherst study outlier criteria is at the expense of excluding 31.9% of the Anaconda soil ingestion estimates. In light of the low levels
of trace element ingestion from food in the Anaconda study, the reduction in uncertainty is considered to be artificial.

Estimates of longer term average exposure in the Anaconda study can be contrasted with comparable estimates from the Amherst study (using variance components estimates from the sixth row of Table IV). Using values of the observed mean (57 mg/day) and 95th percentile (210 mg/day) obtained after excluding the pica subject, comparable estimates of the true average 95th-percentile soil ingestion for children over time periods of $T = (7 \text{ days}, 30 \text{ days}, 90 \text{ days}, \text{ and } 365 \text{ days})$ are 177 mg/day, 135 mg/day, 127 mg/day, and 124 mg/day, respectively. These estimates are quite similar to estimates from the Anaconda study, although markedly different from the previously published estimate of the 95th percentile of average annual soil ingestion ($T = 365$) based on the same Amherst data of 1,751 mg/day (see Table 8). The difference in these estimates results from differences in methodology; limitations in the methodology used in the earlier study have been extensively discussed. Although methods used in this article to estimate longer term predictions also have limitations, they are judged to be substantially less, and the resulting estimates are considered to be more accurate. A manuscript that further contrasts these two methodologies is under preparation.

While the Anaconda study has some notable strengths—low food ingestion of trace elements, larger number of trace elements evaluated, more targeted soil particle size—a number of factors may affect the variability and bias of the soil ingestion estimates. Many of these factors are not unique to the Anaconda study and have been discussed in the context of daily soil ingestion estimates for the Amherst study. Such factors are briefly mentioned here and include assumptions about transit time, homogeneity of trace elements in meals, and the adequacy of duplicate food samples. In addition, biases may occur due to the exclusion of toilet paper/wipes, absorption of trace elements, incomplete food or fecal sample reporting, and ingestion of trace elements from nonfood/nonsoil sources. A more extended discussion with application to the Anaconda study is available from the authors.

4.1. Other Factors Affecting Exposure Assessment

Soil ingestion estimates are commonly used in the IEUBK blood-lead model to predict blood-lead values among children. In such models, estimates are required for longer time intervals than the study periods of previous soil ingestion studies. Since uncertainty and day-to-day variability was evident in soil ingestion estimates, the distribution of estimates of longer term soil ingestion is likely to be less spread out than the distribution of average soil ingestion over a 7-day time period. While there are statistical methods that can be used to estimate the reduction in spread, the performance of such methods will depend on the distribution of uncertainty and day-to-day variation as well as the reliability of the variance components estimates. More theoretical work is needed in this area both in statistical methodology as well as in its application to soil ingestion research.

As in other soil ingestion studies, the timing and location of the study may be substantively related to the quantity of soil ingested. All soil ingestion studies to date have been conducted in northern parts of the United States in the summer or early fall. The studies do not provide data to generalize to other time periods and any generalizations made therefore are mere judgment calls. Interpretation of the soil ingestion estimates for Anaconda children must be based on the fact that children resided on an active EPA Superfund site. While anecdotal accounts have suggested that the children’s soil ingestion behavior may not have been affected by residing on that particular site, it is possible that parents influenced the children to discourage soil ingestion.

The results of this study provide additional information about soil ingestion amounts among children. The estimates of average soil ingestion (over a 7-day period) appear to be more certain than estimates from the Amherst study due to an improved study design. While the focus of the present investigation was average soil ingestion over the study time period, the daily estimates permit identification of acute exposure days. Although no soil ingestion days were reported for Anaconda study children where ingestion was as high as was seen in the Amherst study for the pica child in at least one child (Subject 27 on Day 4), the soil ingestion estimates ranged from 719 mg to 2,828 mg for the seven trace elements. Further research is needed to characterize such acute soil exposures in the Anaconda study, while longer term soil ingestion studies are needed to better characterize the distribution of daily soil ingestion among children.

ACKNOWLEDGMENT

This research was supported by funding under Contract No. (EPA 8X-0475-NAEX) from the U.S. Environmental Protection Agency Superfund Program.
REFERENCES


