SAMPLING MANUAL FOR IEUBK MODEL

Prepared for:

U.S. Environmental Protection Agency
Region VIII
One Denver Place
999 18th Street, Suite 500
Denver, Colorado

by:

ROY F. WESTON, INC.
215 Union Blvd, Suite 600
Lakewood, Colorado

December 31, 1996

DISCLAIMER

NOTICE: This document is a draft and has not undergone Agency review and approval. Therefore, the draft guidance contained in this document should not be construed to reflect final Agency policy or recommendations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SAMPLING MANUAL FOR IEUBK MODEL

1.0   INTRODUCTION

Sources of Lead Exposure

Lead is a naturally-occurring element, and various chemical forms of lead are widely distributed in the environment. In particular, lead can be detected in most samples of soil and water, as well as in many foods and in air. Thus, all humans are exposed to "background" levels of lead from multiple exposure pathways, including ingestion of soil and dust, ingestion of water, inhalation of lead-containing particles of soil or dust in air, and ingestion of foods that have taken lead up from soil or water.

In addition to exposure from natural "background" sources of lead, there are numerous man-made (anthropogenic) sources of lead. Lead has been widely mined, refined and used by humans for hundreds of years, and these activities have resulted in substantial increases in lead levels in some local areas of the environment (e.g., near mining and smelting sites, and near some types of industrial and municipal facilities). In addition, lead has been used in a wide variety of common products that can result in human exposure, including lead-based paints, leaded gasoline, leaded pipes and leaded solder, leaded crystal and ceramics, etc.

Human Health Effects of Excess Lead Exposure

Excess exposure to lead is known to cause a wide variety of adverse effects in humans, including anemia, impaired heme synthesis, renal damage, assorted neurological injuries, hypertension, impaired fetal development and maturation, birth defects, and possibly cancer (EPA 1982, 1986, 1987, 1988). These adverse health effects of lead are of potential concern for any human, but greatest attention is usually focused on young children (e.g., age 0-6 years). This focus on young children is because:

- Children typically have higher intake rates per unit body weight of environmental media (soil, dust, food, water, air, paint) than adults

- Children tend to absorb a higher fraction of ingested lead from the gastrointestinal tract than do adults

- Children tend to be more susceptible to the adverse neurological and developmental effects of lead than adults

Lead Exposure Levels of Health Concern

It is currently difficult to identify what degree of lead exposure, if any, can be considered safe for infants and children. An increasing number of studies report subtle signs of lead-induced neurological and/or behavioral effects in children beginning at around 10 ug/dL or even lower, with population effects becoming clearer and more definite in the range of 30-40 ug/dL (EPA 1988). Of special concern are the claims by some researchers that effects of lead on neurobehavioral performance, heme synthesis, and fetal development may not have a threshold value, and that the effects are long-lasting (EPA 1986). On the other hand, some researchers and clinicians believe the effects that occur in children at low blood lead levels are so minor that they need not be cause for concern.

After a thorough review of all the data, the EPA identified 10 ug/dL as the concentration level at which effects that warrant avoidance begin to occur, and has set as a goal that there should be no more than a 5% chance that a child will have a blood lead value above 10 ug/dL (EPA 1991c, 1994b). Likewise, the Centers for Disease Control and Prevention (CDCP) has established a guideline of 10 ug/dL in preschool children, which is believed to prevent or minimize lead-associated cognitive deficits (CDC 1991).

EPA's IEUBK Model

In order to help evaluate the risks which lead poses to young children, the EPA has developed an integrated exposure, uptake, and biokinetic (IEUBK) model for lead. This model is available as a computer program which can be run on any modern PC. The purpose of the model is to predict the level of lead in the blood of a child or a population of children under a specified set of exposure conditions, taking all sources of lead exposure into account.

Detailed discussions of the structure of the model and how to use the model to assess lead exposures at a site are provided in the "Guidance Manual for the Integrated Exposure Uptake Biokinetic Model for Lead in Children" (EPA 1994a). This Guidance manual should be carefully reviewed and followed when using the IEUBK model.

In brief, the model is composed of two main parts. The first part is the exposure section. In this part, the amount of lead which a child ingests or inhales is calculated from data on a) the concentration of lead in each relevant environmental medium (e.g. soil, dust, food, water and air), and b) information on how much of each of these media is ingested or inhaled by a child each day. The second part of the model is the biokinetic section. This part predicts the blood lead level that will result in the child from the specified exposures. This prediction is based on

data regarding a) how much of the ingested or inhaled lead is actually absorbed into the body, b) how much of the absorbed lead enters each of the different "compartments" of the body (e.g., blood, bone, soft tissue, etc.), and c) how rapidly lead is removed from the body by excretion in urine or feces.

In general, the model is intended to evaluate situations where exposure is on-going, and the exposure levels can be reasonably described in terms of long-term averages. In this case, the predicted blood lead level is the expected long-term average value. This long-term average value is generally considered to be the most appropriate basis for evaluating health risks from lead. The model is not presently intended to allow evaluation of occasional or transitory lead exposures that cause "spikes" in blood lead level (EPA 1994d).

If all of the exposure and biokinetic parameters were accurately known for an individual child, it is expected that the model would predict a reasonable point estimate of the blood lead value for that individual child. If multiple children were exposed, each child would have a different combination of exposure and biokinetic parameters, and each would have a different blood lead level. Thus, if perfect knowledge were available for all children at an exposure location, the model would yield a <u>distribution</u> of blood lead values in the population of exposed children.

In real life, exposure and biokinetic parameters are not known for the individuals at a site, but are only available as group statistics from population studies (e.g., estimated mean soil intake rate, estimated mean gastrointestinal absorption fraction, estimated mean body weight, etc.). Because of this, the model does <u>not</u> seek to accurately predict the blood lead level of any one specific individual, but rather seeks to predict the typical blood lead level that would be expected in an "average" child. Blood lead levels in the entire population of all children, especially those that are at the upper part of the distribution (e.g., the 95th percentile) are then estimated by generating the approximate distribution from the estimated central value. This is achieved by assuming the distribution is approximately lognormal in shape, and by applying an estimate of the degree of variability between different children. This descriptor of variability is the Geometric Standard Deviation (GSD).

In general, the model can be used to evaluate two different kinds of populations. The first is the population of all current and/or hypothetical children exposed at the same location (e.g., at a specific home, daycare center, playground, etc.). That is, the environmental lead levels are the same for all children, but intake rates, absorption factors, etc, differ between children, leading to different values for different children within the population. The second type of population is the population of all children in a large area (e.g., a community). In this case, variability in blood lead levels arise not only because of individual-specific differences in intake and biokinetic factors, but also because of differences in lead concentration levels in different parts of the community. Either application of the model is acceptable, but the two applications should not be confused with each other.

Data Needs

In either application of the model, three types of data are required in order to yield reliable predictions of blood lead values:

1. Lead concentrations or lead intakes from all environmental media or other lead exposure sources (outdoor soil, indoor dust, air, drinking water, food, paint)

2. Human exposure parameters to lead-containing environmental media (e.g., intake rates of soil, dust, water, air, food, paint). These exposure rates are usually considered to be age-dependent.

3. Pharmacokinetic parameters for lead, including absorption rates, and distribution and clearance rates for various internal body compartments (blood, bone, soft tissue, etc.). These parameters are also age-dependent.

Ideally, most of the model input parameters (especially those in group 1 and group 2, above) would be based on site-specific data. However, site-specific data are often lacking for many parameters, and it may not be feasible or practical to collect such data. Therefore, the IEUBK model provides recommended defaults for all of the parameters of the model except for site specific concentrations of lead in soil. These defaults are summarized in Table 1-1. The basis for each of these defaults is provided in EPA (1991b).

Purpose and Organization of This Document

Because of the availability of defaults exposure and biokinetic parameters, the minimum requirement for application of the IEUBK model to an exposure location is a reliable estimate of the mean concentration of lead in soil at that location. However, the accuracy of the model can often be significantly improved if additional site-specific values are obtained for other model parameters, including lead concentrations in indoor dust, drinking water, locally grown foods (e.g., garden vegetables), air and paint. In addition, model accuracy can sometimes be improved by collection of site-specific exposure and demographic data, and/or reliable data on blood lead levels in current populations. The purpose of this document is to provide both general and specific guidance to people who are planning to use the IEUBK model as a tool to evaluate lead risks at a site, specifically with regard to the type and amount of site-specific data to collect to support the modeling effort.

In addition to this Introduction, this document is organized into 3 main parts:

Chapter 2 reviews a number of basic questions which must be answered when planning a site-specific data collection effort to support the IEUBK model, identifying the options available and the principles by which site managers can decide which options are best for their particular sites. Topics covered in this chapter include the following:

- What environmental media should be sampled?
- Where should environmental samples be collected?
- When should environmental samples be collected?
- How many samples of each medium are needed?
- What analytical method should be used?
- Should a demographic survey be performed?
- Should a blood lead survey be performed?
- Should geophysical speciation be performed?
- Should bioavailability tests be performed?

Chapter 3 provides more specific guidance for sample collection and analysis by medium, including:

- Soil
- Dust
- Water
- Paint
- Air
- Home-grown food

Chapter 4 discusses data reduction techniques that should be used to obtain site-specific parameters for use in the IEUBK model.

2.0   GENERAL GUIDANCE

As with any scientific endeavor, the more data and the more information that are available, the better the quality of analysis that is possible. Thus, if cost and schedule were not of practical concern at a site, guidance for a sampling plan to support the IEUBK model would be simple: collect multiple measurements of lead in every environmental medium at every plausible location where children are or might be exposed, collect detailed data on childhood exposures to each of those environmental media, and obtain data on blood lead levels in any currently exposed population. This combination of data offers the best opportunity to correctly evaluate the risk which lead poses to current and future populations of children, to correctly identify the most important sources of lead exposure, and to accurately estimate the likely efficacy of various remedial alternatives. However, cost and schedule <u>are</u> of significant concern in most cases, so considerable judgement must be used in deciding what to sample, how many samples to collect, and whether or not to perform any "special" studies. The following sections discuss the factors which should be considered when making sampling decisions for the IEUBK model.

2.1   What Media Should Be Sampled ?

It is important to recognize from the outset that the IEUBK model is a <u>multi-media</u> model, incorporating information on lead exposures from all sources. This includes those that may be attributable to current or past waste releases or disposal activities at a Superfund site, and those that are related to natural, area, or ubiquitous sources. Thus, at least in concept, the model can be used to evaluate two different questions:

1)   What are the expected distributions of blood lead values at this site, assuming that exposures from non-waste related sources are equal to <u>national average</u> values.

2)   What are the expected blood lead distributions in children at this site, taking into account the <u>site-specific</u> levels of exposure to non-site related sources

Decisions regarding what media to sample depend in part on which question is being asked. In general, the model is usually used to evaluate the first question, and media requiring sampling are only those that are or may have been impacted by site-related waste releases or disposal activities. Typically this would include soil, dust, groundwater, and sometimes local foods and air. However, it must be remembered that if this approach is followed, the calculated probability of elevated blood lead values at specific residences and in the community will tend to be too low if site-specific exposures from other sources (natural, area sources, paint, lead plumbing, etc.) are substantially higher than the national average (default) values.

The second question is usually evaluated only when it is known or suspected that site-specific exposures from non-waste related sources are substantially higher than national average values, or when there are reliable site-specific blood lead data available and a comparison between the observed values and the model predictions is desired. In this case, site-specific data are needed for all potential lead sources at the site.

Soil

As noted above, the single most important medium to sample to support the IEUBK model is soil (specifically, surface soil). This is because this is the only model input parameter for which there is no credible default value, and because this is the medium most likely to be clearly related to site-specific waste release or disposal practices.

Dust

After soil, the next most important medium to sample is indoor house dust (assuming houses or other buildings exist on the site). This is because indoor dust is often as large or an even larger source of exposure for young children than outdoor soil.

If dust measurements are not available, the default assumption employed in the model is that the mass fraction of indoor dust that is derived from soil is 70%. Thus, the concentration of lead in indoor dust can be approximated as:

$$C_{dust} = 0.70 * C_{soil}$$

This default value is based on observations at a number of sites, and is quite reasonable in many cases. However, there are some potential limitations to the use of this default. First, observations at some sites (mainly mining sites) suggest the mass fraction of soil in dust may be lower than 70%, indicating the use of the 70% default value may tend to be conservative for at least some sites. Second, the mass fraction of dust that is derived from soil is likely to be quite variable from house to house (depending on things such as frequency of cleaning, whether the windows are kept open or closed, how frequently children and pets carry soil into the house of the feet or clothing, etc.), but use of a point estimate can not account for this variability. Finally, this default equation does not take into consideration the contribution of non-soil sources of lead to the dust, including indoor sources (e.g., leaded paint) and area sources (airborne emissions from industrial sources, etc.). Because of these potential limitations associated with the default assumption, indoor dust samples should be collected and analyzed for lead whenever an adequate number of homes exist within the study area to provide a meaningful number of values. If measured dust lead concentrations are found to be substantially lower than predicted by the default, this can result in a substantial decrease in the predicted mean blood lead level and in the likelihood of exceeding some health-based target.

There are two basic options for measuring lead in dust: concentration (ug/kg) and loading (ug/m^2). Studies at several sites have found that there is usually a statistical correlation between blood lead levels and either of these measures, so both types of measure can be helpful in assessing risks to children. However, the current version of the IEUBK model is designed to accept data on dust concentration, and there is not yet a convenient way to include the dust loading data. Therefore, based solely on the requirements of the IEUBK model, it is only necessary to measure dust concentration and not dust loading. However, these two options are not mutually exclusive, and it is possible to collect both types of data simultaneously (see Section 3.2), and this approach is generally encouraged. Collection of both types of data has the advantage that when the model is adapted to handle dust loading data, further sampling will not be required. In addition, indoor dust loading data are often used as a means of assessing the efficacy of remedial actions taken at a home by comparing measured loading levels with "clearance levels" established by HUD (EPA 1994c, HUD 1995).

Paint

There is little question that leaded paint on exterior and/or interior surfaces of a home can be an important source of lead exposure. This exposure can be direct (i.e., by ingestion of paint chips or flakes) or indirect (i.e., by ingestion of soil and/or dust that has been contaminated by lead chips or flakes). Thus, measurement of the level of lead in paint, coupled with observations on the condition of the paint (tight, weathered, chipping, peeling, etc.), is often quite valuable in understanding the sources of lead in soil and dust, and in identifying locations where direct paint chip ingestion would be of concern.

However, there are also several reasons why measurement of lead in paint may not always be necessary for supporting the IEUBK model. First, direct exposure to paint (i.e., ingestion of chips) is likely to be intermittent, and is likely to result in lead doses far in excess of those typically received from food, water, soil, dust and air. Thus, when paint chip ingestion occurs, it is likely to result in a large but temporary "spike" in blood lead levels. At present there is insufficient toxicological information to know how to assess the health effects of such exposures (EPA 1994d), and the IEUBK model does not presently allow for quantification of the effects of such exposures. With regard to the indirect exposure pathway, if lead levels have been measured in soil and dust, the contribution of paint is already accounted for, and there is no need to measure lead levels in paint except to assess how much the paint may have influenced lead levels in soil and/or dust. While this paint-to-soil/dust evaluation may be of value, a limitation is that the contribution of leaded paint to soil and dust depends on the condition of the paint, but the condition can easily change over time, either from "good" to "bad" (as a result of weathering and aging) or from "bad" to "good" (as a result of re-painting). Thus, leaded paint will always be a potential source of contamination for both soil an dust, but whether paint has actually served as an important source in the past, or will serve as an important source in the future, cannot be known with certainty.

In summary, collection of data on the concentration of lead in paint and the condition of the paint can be very helpful in providing the risk manager with a complete view of lead sources at the site and in assessing the potential for direct exposures and for contamination of soils or dusts due to paint weathering in the future. However, such data are not always easy to interpret, and are not required to run the IEUBK model.

Water

Humans typically ingest a lot more water than soil or dust, so even low levels of lead in water can be an important contributor to total lead exposure. Therefore, measurement of site-specific lead levels in water is always desirable for improving the accuracy of the IEUBK model.

Lead in water can arise from two distinct sources: 1) lead dissolved or suspended in the source water used by a private well or a municipal system, and 2) lead dissolved from pipes and other plumbing fixtures within an individual house. Thus, a complete understanding the importance of lead in water as a potential source to residents requires knowledge of lead levels both in the source water and at the tap. However, assuming that the main purpose of the risk assessment is to focus on risks from lead released to the environment from a site-related source, it is normally only necessary to analyze lead levels in source water (ground water, surface water), and not in samples collected from household taps. In keeping with standard EPA guidelines, data on lead levels in source water (usually groundwater) should always be investigated if installation of a private well is plausible, even if current and future residents are likely to be served by municipal water systems and installation of a private well is not especially likely.

In the case where a correlation analysis will be performed to investigate the relationship between measured environmental lead levels and observed blood lead levels in children, measurements of lead levels at the tap are needed to account for possible exposures from in-house (plumbing) sources.

Air

Lead is not volatile, but can exist in air adsorbed to small respirable dust particles (PM_{10}). However, in most cases, the dose of lead inhaled is very small compared to the dose ingested from soil, dust and other sources. For example, using the default exposure assumptions in the IEUBK model for a child age 2-3 (see Table 1-1) and the default assumption regarding the amount of soil which exists as respirable particles in air (see EPA 1990), the ratio of the lead dose from inhalation of PM_{10} to the dose from ingestion of soil and dust is as follows:

$$\frac{DI_{air}}{DI_{soil}} = \frac{C_{air} \cdot BR}{C_{soil}\,IR_s + 0.7\,IR_d}$$

where:

DI_{air} = Daily intake of lead from air (ug/day)

DI_{soil} = Daily intake of lead from soil and dust (ug/day)

C_{air} = Concentration of lead in air (ug/m^3). This can be estimated as follows:

$$C_{air} = C_{soil}*PEF$$

where PEF= particulate emission factor (default = 2E-04 mg/m^3)

BR = Breathing rate (5 m^3/day)

IR_s = Ingestion rate for soil (61 mg/day)

IR_d = Ingestion rate for dust (74 mg/day)

Based on these values, the inhaled dose of lead from soil suspended in air is less than 0.001% of the ingested dose. Even when there are other sources of lead in air (e.g., a nearby municipal power plant or industrial facility), the dose from inhalation of air is still usually much less than from oral ingestion pathways. On this basis, there is usually little need to collect site-specific data on lead levels in air for the purpose of supporting the IEUBK model. However, measurements of lead (or other chemicals) in air may still be needed to evaluate the importance of air-borne transport of particulate matter from source areas to exposure locations.

Food

Reliable data are available from FDA "market basket surveys" on lead levels in typical foodstuffs, as are data on the typical amounts of each type of food ingested by children. Thus, the default dietary intake values shown in Table 1-1 are normally adequate for estimating exposure though "store-bought" foods.

Additional sources of dietary exposure that are sometimes considered at Superfund sites include ingestion of homegrown garden vegetables and ingestion of meat or milk from locally-raised cattle. The concentration of lead in these foods can be estimated from the concentration of lead in the soil where the vegetables are grown or the cattle are raised (e.g., Baes et al. 1984), but the calculations involve a lot of assumptions and estimates, and the results are not likely to be highly accurate. Therefore, if exposure through a local food pathway is suspected to be important, collection and analysis of samples is expected to yield data that are far more robust that the modeled values. If samples of locally-raised foods are collected for analysis, data on consumption rates should also be collected at the same time.

With respect to exposure by the home-garden vegetable pathway, there have been too few studies at Superfund sites to allow reliable evaluation of the relative importance of this exposure pathway. However, limited data (e.g., Sverdrup 1995) indicate that the dose of lead from garden vegetables consumed by children may be 5-10% of that from typical soil and dust ingestion (EPA 1995a). Although only a fraction of the total intake, such exposures can increase the predicted geometric mean blood lead value by a similar percentage, and this can have an even

larger effect on the predicted risk of exceeding a target blood lead value (depending on how high the geometric mean value is).

With respect to lead exposure via locally raised beef and milk, exposure of the cow can occur by ingestion of contaminated fodder, ingestion of contaminated soil while grazing, or ingestion of contaminated water. However, lead that is absorbed by a cow from these sources is expected to be retained mainly in bone and not in the milk or flesh of the animal. Thus, except in cases where locally-raised cattle provide a large fraction of the beef and/or milk ingested by a child, exposure from local beef products is expected to be sufficiently minor that the cost of sampling and analysis of such samples is usually not warranted.

2.2   Where Should Samples Be Collected?

The answer to the question "Where should I sample?" depends in turn on the answer to the question "Where will children be exposed?". Locations where exposure is known or suspected to occur, now or in the future, are referred to as "Exposure Units". Once the location of all exposure units is known, then samples of all relevant media (including soil, dust, and groundwater, and possibly paint and/or locally-raised food) should be collected in all exposure units.

It is generally assumed that young children (e.g., age 0-2 years) are likely to spend most of their time at their own house, and the exposure unit most often selected for evaluation in the IEUBK model is an individual home (including both the interior of the home and the immediate yard area). In this case, the parameters needed for the model are the mean concentration of lead in the outdoor soil of the yard (averaged over the entire yard area) and the mean concentration of lead in indoor dust (averaged over all indoor areas where the child could be exposed).

In accord with this objective, outdoor samples should be collected from multiple locations within the yard. If it is known or assumed that exposure is random across the yard, then sampling locations should also be random (or systematic). However, it is often known or supposed that children are more likely to play in some areas of their yard than others, so some sampling schemes use a biased sampling strategy, taking samples preferentially from known or suspected play areas and/or bare areas. This approach may be helpful for evaluating the exposures of current children who live at the house, but is of uncertain merit at locations where no children currently reside, and might be misleading in the assessment of risks to future children who might play in other locations in the yard. Therefore, true randon sampling is generally preferred.

Two basic tactics are available for handling the multiple soil samples from within a yard. In the first case, each yard soil sample is analyzed separately. This allows the recognition of "hot spots" that may exist within a yard, and also helps evaluate the degree of small scale

11

heterogeneity and the attendant problems of measurement error. Alternatively, all of the soil samples from a yard can be composited. The measured value in the composite is then taken to be the best estimate of the mean for that yard. This approach obviously allows a savings in analytical costs, but looses information on the degree of variability between samples and tends to force "all-or-none" remedial decisions at a residence. Obviously, strategies intermediate between these two extremes are also possible (e.g., compositing sample from the front yard and back yard separately, etc.). In general, the best approach is to collect a number of composite samples, combining the benefits of compositing with the advantages of discrete sampling (see Section 3.1).

The same logic regarding soil sample collection applies to indoor dust samples. Dust should be collected from multiple locations within the home that are selected to be representative of locations where young children are likely to be exposed (floors, carpets, and surfaces in the rooms most often occupied by the children). Because it can be difficult to obtain enough dust at each location to permit a reliable chemical analysis, it is common to composite dust samples, sometimes ending with only one sample per house.

Unfortunately, this relatively simple concept of the home and yard as an exposure unit becomes rather uncertain as the child becomes older. That is, children become more mobile as they age, and some children age 4-6 may spend considerable time at daycare centers, neighbors' houses, or at local playgrounds. In addition, lead levels in indoor dust often depend not only on the lead level in the immediate yard area, but also on lead levels in the general area surrounding the home (neighbors yards, nearby waste piles, bare areas, industrial releases, etc.). The exact mechanism by which such sources contribute to indoor dust is not certain, but is likely to involve airborne transport and/or transport on the feet of children, pets, etc. In any event, because of this pathway, lead levels in indoor dust are usually best described by an equation of the form:

$$C_{dust} = D_0 + k*C_{soil}$$

The fact that D_0 (the concentration of lead in dust that is derived from non-yard sources) is usually not zero has important implications regarding the expected efficacy of decreasing exposure at a residence by reducing only the concentration of lead at that residence.

Because of these two problems (the tendency for children to be exposed over an area larger than their own yard as they become older, and the transport of material from "off-yard" locations into most houses), it is sometimes helpful for the risk assessor and the risk manager to consider risks not only on the level of a single residence, but also on the level of a "neighborhood" (e.g., several blocks of homes). It is important to realize that these are not mutually exclusive considerations, and the object should be to recognize both individual residences and groups of residences that are of potential concern. Consequently, the sampling plan for lead in surface soil

should not be restricted to current or future residences, but should also include neighborhood playgrounds, "attractive nuisances" (site features that may tend to attract children to play there), school yards, daycare centers, etc.

Because lead is a naturally occurring metal, it is expected that some of the lead present in soil, dust and water is due to natural ("background") sources. At many sites, background levels in soil and water will be sufficiently low that they will not constitute a significant proportion of the dose to children at the site, and there is little need to precisely quantify what the background contribution is. However, at some sites, especially in mineralized areas, lead levels in soil and/or groundwater may be substantially higher than national averages, and it may be helpful to the risk manager to know what the "background" exposure levels are. In this case, collection of data from a number of carefully selected background locations is recommended.

2.3   When Should Samples Be Collected?

Lead is relatively stable in soil, and concentration values do not fluctuate significantly over the short term. Thus, in most cases, the time (season) of soil sample collection is not critical. However, there are several other media when the timing of sample collection may be important.

The first case is groundwater. There are a number of reasons why the level of lead in groundwater might vary with time, including variations in rain or snowmelt, variations in groundwater depth, etc. Characterization of the temporal variability in groundwater concentration of lead and any other chemical of concern should always be considered as part of the site characterization effort.

Also, as already discussed above, lead levels in tap water may vary as a function of how long the water has been in contact with leaded pipes or leaded solder. Thus, samples taken at different times of day (first flush, post flush) are needed to characterize lead exposure from this source.

Finally, the amount of lead-contaminated material entering indoor house dust from outdoor sources (soil, air emissions, etc.) presumably depends on a number of factors that are influenced by the weather (windows open or shut, mud on shoes, bare areas covered with snow, etc.), so indoor dust lead levels may vary as a function of season. This is supported by data from a survey at a smelter site in Montana, where indoor dust levels of lead and arsenic appeared to show a seasonal pattern (tending to be lower in winter than summer), although this difference did not prove to be statistically significant. Because it is the long-term average dust level that is needed for accurate predictions of the IEUBK model, dust sampling over time is desirable when schedule permits.

2.4   How Many Samples Are Needed?

The IEUBK model requires as input the estimated arithmetic mean concentration of lead in soil, dust, and other environmental media at an exposure unit. Direct measurement of such mean values is complicated by three types of variability:

- Spatial variability
- Temporal variability
- Analytical variability

Spatial variability refers to differences in concentration as a function of sampling location within the exposure unit. It is very common to observe substantial spatial variability in lead levels in soil and other solid media (dust, waste piles, etc.). The scale of the variability can be both quite small (differences between samples separated by only a few meters) and very large (e.g., differences between the center and the edge of a "footprint" from stack fallout). Temporal variability refers to differences in concentration at the same location as a function of time. As noted in the preceding section, this can include groundwater, tapwater and possibly dust, but is usually not of concern for soil, waste piles, etc. Analytical variability refers to the differences between replicate analyses of the same sample. Usually, differences due to analytical variability are rather small compared to spatial and/or temporal variability.

Granted that there is significant spatial and/or temporal variability in lead levels in a medium, the number of samples needed to estimate the true mean concentration of lead in a particular medium within a particular exposure unit depends on two factors: 1) desired accuracy, and 2) sample variability. In general, the greater the desired accuracy and the greater the variability within the unit, the greater the number of samples needed. These principles are illustrated in Figure 2-1, which shows the 95% lower and upper confidence limits about the mean of a lognormal distribution as a function of sample number and variability (GSD). As expected, uncertainty in the estimate of the mean decreases as N increases, and, for any given number of samples, uncertainty is larger for distributions with high variability (high GSD) than for distributions with lower variability (small GSD).

The degree of accuracy needed when estimating the mean concentration in a medium is a matter of judgement, and should be selected using EPA's Data Quality Objectives (DQO) procedure (EPA 1992a, 1993a). In general, the most important factor to consider in selecting DQOs is whether high accuracy makes any difference in the outcome of the model run or in subsequent risk management decisions. For example, high accuracy is not needed for media that are contributing only a small fraction of the total dose of lead, and high accuracy is not essential if the predicted blood lead distribution has a 95th percentile blood lead value that is either clearly above or clearly below a level of concern (typically 10 ug/dL). However, if the predicted 95th percentile blood lead level is close to the decision threshold, then accurate estimates of the mean are needed, at least for the main sources of exposure (typically soil and dust).

For these reasons, it is very helpful if preliminary data can be obtained to estimate 1) the approximate lead levels in environmental media, 2) the variability in those levels, and 3) the likely range of IEUBK model outputs than can be expected at the site. These data can then be used to form a proper basis for selection of the optimum number of samples of each type of medium to collect to meet the DQOs.

2.5   What Analytical Procedure Should Be Used?

There are two basic choices for the analysis of environmental media for lead:

- Wet chemistry methods

- X-ray fluorescence (XRF) techniques

The chief factors which influence selection of one method over another center around sensitivity (detection limits), accuracy, speed, and cost. The advantages and disadvantages of the various options are reviewed below.

2.5.1   Wet Chemistry Methods

Sample Preparation

All "wet chemistry" methods for the analysis of lead in solid samples (soil, dust, slag, etc.) involve acid digestion of the sample prior to analysis. Refluxing in nitric acid followed by hydrogen peroxide (e.g., method ILMO1.0, SW846 method 3050, ASTM method ES 36-94) is generally the method of choice if graphite furnace atomic absorption spectroscopy (GFAA) analysis is to be performed. For inductively-coupled plasma (ICP) and inductively-coupled plasma-mass spectroscopy (ICP-MS) analyses, hydrochloric acid is added to increase the digestion efficiency and overall metals solubility. These two digestion method are actually acid leaching methods and not vigorous total metals digestions. If more extensive digestions are required, perchloric acid and hydrofluoric acid digestions in combination with nitric and hydrochloric acid are available. Hydrofluoric acid digestion may be desirable if silica-containing material (e.g., slag) is present and/or comparison to XRF analysis is desired. Caution should be exercised in requesting these more vigorous methods because perchloric and hydrofluoric acids are more dangerous to handle and require more specialized equipment, and some laboratories are not equipped to handle these acids. In addition, graphite furnace analysis is not recommended as an analysis technique when using these vigorous digestions.

Analytical Methods

There are several methods available for the analysis of sample digestates. Generally these methods involve either graphite furnace AA or some type of inductively couple plasma (ICP) analysis. Graphite furnace has been the traditional method chosen for low level (<50 ug/L) analysis. However, this technique is subject to interference and, as noted above, is not suited to the more vigorous digestion extracts.

Traditional ICP analysis could only reach an instrument detection limit of around 50 ug/L (in the acid digest), but much progress has been made in modifying ICP analytical instrumentation to reach detection limits that are comparable to those achievable by graphite furnace (about 1 ug/L). Two techniques in particular have been developed which reach detection limits comparable to graphite furnace. ICP-Trace analysis is simply a modification of the tradition ICP method. This involves a longer optical viewing path which lowers the detection limit. Reliable detection limits of 1-3 ug/L have been demonstrated. Correctable interferences do exist for this method, and if this analysis method is chosen, the laboratory should demonstrate its' ability to accurately correct for these interferences. In ICP-MS, a mass spectrometer replaces the traditional ICP optical system as the detector. The ICP-MS can reach even lower detection limits than either traditional ICP, ICP-Trace, or graphite furnace, and has the added advantage that several mass isotopes can be monitored if isotopic ratios are important to a study. This method is also subject to interference and the laboratory should be able to demonstrate the ability to correct for this problem. A final advantage of ICP analysis is that it can handle analysis of extracts digested by the more vigorous methods, provided the laboratory has taken the necessary steps to handle these materials.

2.5.2 XRF Techniques

Historically, XRF techniques were generally viewed as being suitable only for screening, but recent advances in technique and instrumentation now permit this method to yield results that are fully comparable with CLP methods for soil and dust. XRF techniques are also convenient for measuring lead in paint, but are not usually used for measuring lead in water.

There are several different types of XRF instruments, differing in the spectrum and intensity of the x-rays used to irradiate the sample, and in the detection system used to measure the intensity of the induced fluorescence in the sample. These differences in instrumentation result in differences in performance (detection limit, susceptibility to interferences from arsenic and other metals, ease of sample preparation, etc.). Table 2-1 summarizes the characteristics of the most important varieties of XRF instruments.

Field Portable XRF

Field portable XRF techniques have the distinct advantage that they are fast (results are available in real time), and they are relatively accurate (about ± 6% compared to CLP) (EPA

1993). In addition, the cost per sample is often substantially lower for field-portable XRF than for CLP, at least if a large number of samples are analyzed. Newer instruments can usually achieve detection limits of 200-500 ppm in soil, depending on the degree of interference from arsenic and other metals. Thus, this method is often helpful for initial site characterization and for initial mapping of soil lead levels above 500 ppm. The field-portable XRF instrument is also the most convenient way to measure lead levels in interior and exterior paint *in situ*, without having to remove a sample of the paint.

Fixed-Base XRF

Three basic types of laboratory-housed ("fixed-base") XRF instruments are available. The wavelength-dispersive system (WDS-XRF) has the lowest detection limit (about 2 ppm), but requires the most complex sample preparation and is the most costly to purchase. Energy-dispersive systems (EDS-XRF$_1$, EDS-XRF$_2$) have detection limits in the range of 5-50 ppm, depending on the degree of interference from arsenic and on the nature of the x-ray source used to irradiate the sample. In all three of these instruments, the level of sensitivity is more than adequate to identify areas where soil or dust leads are of potential health concern. Fixed-based XRF analyses are also relatively inexpensive compared to CLP procedures.

Validation

There have been a number of studies which have shown that there is good correlation between lead levels measured by CLP and by XRF (including both field-portable and fixed-base measurements). However, this type of inter-method comparison remains an important component of a proper QA plan, and all sampling plans that utilize XRF techniques should include CLP analysis of about 5%-10% of the samples to demonstrate that the XRF results are accurate. As noted above, in the special case where CLP results are consistently and significantly lower than XRF, the possibility should be investigated that some of the lead exists in a silica-rich matrix that is not fully dissolved by the normal CLP digestion procedure.

2.5.3   Summary

When lead is the main chemical of potential concern, analysis of samples by field portable and/or fixed-base XRF techniques can offer considerable advantages in cost and time compared to standard CLP methods. However, because EPA has not yet published official standardized XRF protocols, it is important that personnel be properly trained and that care be taken to ensure the instruments are properly calibrated. Analysis of 5%-10% of the samples by CLP as well as by XRF allows validation of the accuracy and comparability of the XRF data.

2.6   Should Demographic Data Be Collected?

If a residential population currently lives on or near an area of potential concern for lead exposure, then the possibility exists that a survey of these residents could yield information that would improve the accuracy of the IEUBK model predictions. However, if a stand-alone demographics survey is being contemplated (i.e., a demographics study not accompanied by a parallel blood lead study), the only data items that are immediately valuable are those that can be directly or indirectly converted to input parameters in the IEUBK model. For example, the survey might seek to obtain site-specific data on variables such as first-flush and post-flush water intake rates, time spent indoors, time spent outdoors, home-grown garden vegetable ingestion rate, etc. The mean values from the survey could then be substituted into the model for the existing national default values.

Ideally, the demographic survey would also include the question "On average, how much soil and how much dust does your child ingest each day?", but of course parents have no way of knowing this. However, parents can provide information on parameters such as the frequency and extent of mouthing and hand-to-mouth activity, which is probably the main pathway for soil and dust ingestion. Thus, questions of this type can help establish an approximate relative rank of an individual child in the overall distribution of soil and dust exposure. However, such data, in and of themselves, are not very useful in adjusting the mean soil and dust intake rates unless it is known whether the site specific surrogates are higher or lower than average. Likewise, collection of other parameters which are often observed to be correlated with blood lead levels (e.g., socioeconomic status, dietary status, house age, etc.) are not very helpful since they do not enter explicitly into the model.

However, if a blood lead study is being planned, concurrent collection of demographic data on a very wide variety of parameters can be valuable in assessing the relative importance of various lead sources and behavioral traits, as discussed below.

2.7   Should Blood Lead Data Be Sought?

While the IEUBK model has many advantages as a predictive tool to evaluate the potential effects of lead sources on a population of children, other types of data can also be valuable. This includes direct measurement of blood lead levels in current child residents in the area(s) of concern (EPA 1994a, 1994b). If such a blood lead survey is properly planned and executed, the resulting data can provide a number of potentially useful information items. For example, site summary statistics (mean blood lead, percent of the population above 10 ug/dL, etc.) can be compared with corresponding national average statistics (Brody et al. 1994, Pirkle et al. 1994) in order to obtain a general sense of how much impact site contamination may have caused in the population. Further, the site statistics can be compared with health-based objectives and guidelines in order to determine if population-based health goals are being exceeded. In addition, blood lead studies which include reliable data on lead levels in various environmental media (soil, dust, paint, water, food) and which obtain reliable demographics

data (age, sex, race, mouthing frequency, dietary status, etc.) can provide valuable insights into the media and exposure pathways that are the primary sources of concern in a population (e.g., see EPA 1994a, Sections 4.3 and 4.4).

However, there are also some important limitations to the use of blood lead measurements as the only index of lead risk. First, care must be taken to ensure that a sufficient number of children are studied, and that these children are representative of the population of concern. Second, blood lead values in an individual may vary as a function of time (day to day, season to season, year to year), so a single measurement may not be representative of the long-term average value in that individual. Third, because of the variability between people in contact rates for various media, it is expected that blood lead values will differ (either lower or higher) between individuals, even when they are exposed under the same environmental conditions. Thus, a blood level that is below a level of concern in one child living at a specific residence does not necessarily mean that some other child who might be exposed at the same location might not have a higher (and possibly unacceptable) blood lead level. Fourth, population-based studies are not well-suited to detecting the occurrence of occasional sub-locations where risk is elevated, even if average risks are not above a level of concern. Finally, blood lead measurements reflect exposures and risks under current site conditions and population characteristics, which may not always be representative of past or future site conditions. For these reasons, results from blood lead studies are not adequate, in and of themselves, as a techniques for assessing current and future lead risks at a site.

Despite the numerous limitations of blood lead studies, the potential advantages and gains in understanding are such that the benefits of seeking and employing blood lead data to be used in conjunction with the IEUBK model should always be at least considered at sites where lead is likely to be the main chemical of concern. However, EPA is not authorized under Superfund to collect or sponsor such studies, so if such data are to be obtained and evaluated, they must come either from community-based health programs or from site-specific studies performed by other agencies or parties.

It is important to recognize that planning and performing a reliable blood lead study requires special expertise, and is not inexpensive. Thus, considerable cost-benefit judgement must be exercised in deciding whether or not to request a blood lead study. The most important factor to evaluate is the size of the current population currently living in the area(s) of potential concern. Because of the wide variability between people, studies that lack sufficient participants from the impacted areas usually lack the statistical power to determine whether blood lead levels are impacted or not, and to allow analysis of the importance of different potential lead sources. There is no fixed minimum number of participants which guarantee success, but the higher the number of participants, the more likely it is that meaningful conclusions can be drawn. If the only objective is to compare geometric mean values, then a diference of about 10% can be detected wth 95% confidence with a set of about 20-40 samples

(depending on the GSD of the data set) (Gilbert 1987). If the objective is to compare the observed fractions of children above 10 ug/dL, or to investigate the correlation betwwen blood lead and environmental lead levels in soil, dust, or other media, then significantly more samples are required.

The second factor to consider is the cost of performing the study compared to the potential cost savings if the data support a higher clean-up level. In general, the larger the site and the greater the number of exposure units (residences, properties) that may need remediation, the greater are the potential benefits of performing a blood lead study, and the more carefully this option should be considered.

2.8   Should Geochemical Characterization and/or Bioavailability Studies be Performed?

Lead, like most other metals, can exist in the environment in a number of different chemical and physical states. This is of potential importance to the risk assessment process because the bioavailability of lead (the extent to which ingested lead is absorbed into the body) may depend on the chemical and/or physical properties of the lead. In addition, knowledge of the chemical and physical nature of the lead may provide valuable information regarding the likely source of the lead and which fate and transport process are likely to be of importance.

At present, the IEUBK model assumes that lead in soil and dust is absorbed about 60% as well as lead in food and water. That is, the relative bioavailability (RBA) is 60%. There are three basic alternatives available for investigating whether the site-specific bioavailability of lead might be different from this default.

1)   Measure the bioavailability of lead in an appropriate animal study.

2)   Measure the solubility of lead in an in vitro test system, and estimate the bioavailability by extrapolation

3)   Characterize the physical and chemical forms of lead present by electron microscopy, and estimate the relative bioavailability by extrapolation from in vivo and/or in vitro studies of other test materials

The first option is clearly the most direct approach to obtaining reliable site-specific data. However, because of the special resources and the time and cost requirements of animal studies of this type, performing such tests is not usually a feasible option unless site-specific samples can be submitted to an on-going test program.

The second option has the advantage that in vitro tests are inexpensive and fast, but has the distinct disadvantage that the measured value (typically percent solubilized under some specified set of conditions) is not a direct measure of bioavailability. However, if an in vitro test indicates that a soil or other sample has very low solubility, this can be used to support the notion that bioavailability may be lower than the default.

The third option has not been feasible until recently. However, the EPA has just completed measurements of lead bioavailability on a number of different test materials from lead contaminated sites across the country. In all cases, the test materials were well characterized by electron microprobe analysis, so extrapolation of results across different samples is now possible, at least conceptually. The chief difficulty with this approach is that it is not always easy to decide if a site sample is sufficiently similar to a sample that has been tested in animals, since each sample contains a variety of different lead forms in a variety of different sizes and matrix associations. Thus, a new site sample might be similar in some regards but different in other regards, and knowing whether to apply the results from one sample to another may become problematical.

In all cases (regardless of the method employed to investigate bioavailability), one of the most difficult problems in the application of the data stems from the variability in composition between different samples from the same site. In this situation (assuming that the estimates of bioavailability for the different samples are not all alike), it may be difficult to judge over how large an area an particular estimate of RBA may be applicable. One way to reduce the uncertainty with this issue is to collect and analyze a larger number of samples, but the cost of doing so may be prohibitive (especially with animal tests). On the other hand, it is almost certainly better to use even limited site-specific data than to rely only on the default value.

3.0   DETAILED PROTOCOLS AND GUIDANCE

This section provides detailed guidance on sample collection, preparation and analysis techniques to be used in site investigations intended to support the use of the IEUBK model. A number of other EPA guidance documents already exist on environmental sampling and analysis methods, and the following protocols summarize and cross-reference these guidance documents, as appropriate.

3.1   SOIL

3.1.1   Exposure Unit

The basic concept of an exposure unit is that it comprises the area over which a person is randomly exposed. As discussed previously, there are two alternative tactics which can be followed in defining an exposure unit for childhood exposure to lead. The first is to assume that, because of the young age of the child, exposure occurs primarily in and about the home. In this case, the exposure unit evaluated by the IEUBK model is the residence of the child, including the dwelling itself and the surrounding yard. In many cases, this area will be defined by the property boundary of the residence. However, in cases where current of future residential properties are large, the property should be divided into units no larger than about 0.5 acre (EPA 1994a), and each subarea should be sampled separately.

The second approach to defining an exposure unit for lead is to recognize that children may not be exposed only in their own homes, but also at daycare centers, neighborhood playgrounds, homes and yards of neighbor children, etc. In this event, the object is to estimate an intake-weighted mean concentration of lead, based on information about the lifestyles and behaviors that are customary for the majority of children within an community or neighborhood.

In general, the second approach (defining the size of the exposure unit on a case by case basis, employing as much site-specific data as possible) is considered to be technically superior and is recommended whenever feasible. However, it is recognized that this approach is more difficult to implement, and may not be warranted in all cases. Therefore, the first approach ("yard-by-yard") is considered an acceptable surrogate exposure unit. Whichever approach is selected, the decision should be made before sampling begins, and should reflect the mutual decision of the site project manager and the risk assessor.

Note that, because exposure units are based on human activity patterns, it is not normally appropriate to select the boundaries of exposure units based on environmental patterns of soil contamination. However, in the special case when environmental contamination is relatively uniform, sampling and analysis costs can be reduced by defining the sampling area as the combination of a number of uniformly contaminated exposure units.

3.1.2   Sampling Location within the Exposure Unit

There are two basic alternatives for sampling within an exposure unit. In the first case, the objective is to estimate the distribution of blood lead values that is expected in both current and future child residents. As noted above, the basic assumption that underlies the definition of an exposure unit is that exposure is <u>random</u> across the entire exposure unit. Based on this, in the general case, the sample locations should be <u>not</u> be biased toward over-representation of areas that are expected to contain higher-than-average lead levels (e.g., drip lines), areas where exposure is suspected to occur more frequently than other areas (e.g., a play area), or areas where contact with soil is thought to be more likely than for other locations (e.g., a bare area). The reason that biased sampling is not appropriate is that there is no method by which such biased data can be used to derive an unbiased estimate of the true mean over the exposure area (EPA 1994a). Rather, the exposure unit should be sampled using a systematic sampling pattern to ensure balanced representation of all areas of the unit. This is achieved by selecting an initial sampling point at random, and then selecting other sampling points using a rectangular or triangular grid pattern (EPA 1994a). The node spacing in the grid system is selected so that an adequate number of samples are collected from each unit (see below). Detailed descriptions of how to implement such "random systematic" sampling schemes are provided in EPA 1989a and EPA 1994c.

If the assumption of random exposure over the exposure unit is not considered to be realistic, then the exposure unit may be divided into sub-areas (within which exposure <u>is</u> assumed to be random), and each of these sub-areas may be sampled using the same principles as apply for any exposure point. Then the average concentration values in these multiple sub-areas should be combined on an intake-weighted basis to yield a single concentration value applicable to the entire yard. This approach is generally difficult to implement, since quantitative data on current exposure patterns within an exposure unit are typically not available, and current exposure patterns, if available, are not necessarily predictive of future exposure patterns.

In the second case, the objective is to employ the IEUBK model to make as accurate a prediction regarding the blood lead levels in <u>current</u> children residents, and not necessarily in any future residents. This objective is often desirable when blood lead data are available from children currently residing at a location, and comparisons between observed and predicted blood lead values are to be performed. In this even, biased sampling of current play areas based on information about the exposure pattern of the current child residents is reasonable and appropriate.

3.1.3   Sample Number

As discussed previously (see Section 1.0), decisions about the risk from lead at a residence are usually based on an estimate of the probability that a random child living at the residence would

have a blood lead level higher than 10 ug/dL. For convenience, this probability is referred to as "P10". The value of P10 is rather sensitive to the mean soil lead value entered, as illustrated in the following table (calculated using all standard default values in the IEUBK model):

Assumed True Concentration in Soil (ppm)	"True" P10 Value (from IEUBK Model)	Confidence in Soil Concentration		Confidence in Value of P10[a] (%)
		Factor	Range (ppm)	Min -- Max
200	1.3%	1.3	154-260	0.6-2.7
		1.5	133-300	0.4-4.2
		2.0	100-400	0.2-8.3
400	8.3%	1.3	308-520	4.4-16
		1.5	267-600	2.9-21
		2.0	200-800	1.3-34
800	34%	1.3	615-1040	22-46
		1.5	533-1200	16-56
		2.0	400-1600	8.3-73

[a] Calculated using the IEUBK model based on the range of the soil concentrationS

As seen, confidence intervals around the value of P10 are substantially larger than the confidence intervals around the value of the mean soil concentration, and can readily span the decision threshold (P10 < 5%) when the soil concentration is in the range of 300-600 ppm. Therefore, <u>if the soil lead level is close to a decision threshold</u>, it is important to obtain as accurate an estimate of the true mean concentration in the exposure unit as is practical. For log-normal distributions with even moderate variability, it is generally not practical to attempt to define the mean more accurately than within a factor of about 1.2-1.3 (see Figure 2-1). Accepting this level of accuracy as the target, the approximate number of samples needed can be estimated from the following table:

Expected Variability in Soil Concentrations Within Exposure Unit	Approximate Number of Samples Needed to Estimate the Mean Within a Factor of About 1.3
Low (GSD = 1.5)	10
Medium (GSD = 2.0)	30
High (GSD = 3.0)	60

As seen, in order to achieve the desired level of confidence in the sample mean (i.e., the sample mean is likely to be with a factor of 1.3 of the true mean), it is necessary to collect a substantial number of samples from each exposure unit, depending on the degree of variability between samples within the unit. If site-specific data are not available to estimate what the variability is, then assume the variability is probably about equal to a GSD of 2. Based on this, the default recommendation is to collect about 30 samples from each exposure unit.

If the mean soil concentration in the exposure unit is either clearly above or clearly below a level of concern, then establishing such a precise estimate of the mean is not required. In this case, the recommended goal is to collect sufficient samples that the sample mean is likely to lie within 1.5-2.0 fold of the true mean. The number of samples need to achieve this goal can be determined from the following table:

Expected Variability Within Exposure Unit	Number of Samples Needed to Estimate the Mean Within a Factor of about 1.5	Number of Samples Needed to Estimate the Mean Within a Factor of about 2
Low (GSD = 1.5)	7	5
Medium (GSD = 2.0)	15	7
High (GSD = 3.0)	30	12

3.1.4 Compositing

In any case, regardless of the number of samples collected, it is <u>not always necessary</u> that each of these samples be analyzed individually. Rather, they may be combined into a number of composite samples (EPA 1994a, 1994b, 1995). This is because the calculated mean of N individual samples is expected to be quite close to the single measured value for a composite of those same N samples.

There is no firm rule for the number of sub-samples that should be combined to form a composite sample, but most guidelines recommend that the number be at least three and no more than about 10 (e.g., EPA 1994a, 1994b). Based on this, if the total number of samples needed to define the mean is about 30, reasonable choices could range from 10 composites of 3 each to 3 composites of 10 each. However, if too few composites are prepared for an exposure unit, it will not be possible to draw any conclusions about the pattern of lead contamination within that unit. Thus, it is recommended that no fewer than 4 composite samples be prepared per unit, with an optimum number of about 4-5 sub-samples per composite. The sub-samples that are composited into a single sample should be contiguous so that the concentration of that composite may be interpreted as representative of a particular sub-location in the yard. This compositing may be done either in the field or in the laboratory.

3.1.5 Sample Collection Method

Collection of surface soil is accomplished either by use of a sampling spoon or a coring device (EPA 1995). Selection of which method is better is partly a function of the characteristics of the soil, and partly a matter of convenience. In general, use of the spoon technique is probably simpler.

If the sampling location is covered with grass or other vegetation, this should be removed before collection of the sample. This is done by cutting a ring in the grass larger than the collection area, lifting up the grass, and carefully shaking as much as possible of the adhering soil in the root zone back onto the surface of the collection site.

3.1.6 Sample Depth

For assessment of current risks to children from lead in soil, it is usually assumed that exposure to surface soil is much more likely to be of concern than exposure to sub-surface soil. There is no standard definition of what constitutes surface soil, but depths from 0.5 to 2 inches are usually accepted as being reasonable (e.g., EPA 1991c, EPA 1994b, 1995b).

In some cases, it may also be important to investigate lead concentrations as a function of depth below the surface to determine whether any buried sources of potential concern exist. However, any IEUBK analysis of potential future risks to children from potential excavation of buried sources should be clearly distinguished from the analysis of risks from current surface soil levels, since predicting the true mean concentration in surface soil following some hypothetical future excavation or earth-moving operation is especially tenuous.

3.1.7 Sample Preparation

All samples of soil to be analyzed for metals are prepared by drying and sieving. In many cases, field samples are sieved using a 10-mesh nylon sieve to isolate particles smaller than about 2 mm. However, the EPA believes that the particles of greatest concern for human exposure are those with mean diameters less than about 250 um, and that "bulk" soil samples (those sieved to a particle size less than 2 mm) may not be representative of the particles to which children are most likely to be exposed. This is because, at least for relatively dry soils, small particles (< 250 um) are considered more likely to adhere to the hand (Kissel et al., 1996) and be ingested. Also, it is expected that small particles are more likely to adhere to clothing or be carried by the wind into the indoors of a home and contribute to dust contamination. Therefore, the bulk soil sample should be air-dried (at a temperature of 40-60°C) and then sieved using a 60-mesh nylon screen to isolate particles smaller than about 250 um.

This distinction in particle size is important because, in at least some cases, the concentration of metals tends to be higher in the fines than in the bulk sample (e.g., Davis et al. 1992). However, in other cases, the concentration of lead in the bulk soil sample may be similar to the concentration in the fine fraction. If sufficient data are collected (by analysis of paired bulk and sieved samples) to establish that there is no important difference in concentration as a function of particle size, then the sieving step may be omitted if cost or schedule constraints dictate. However, this is not encouraged.

3.1.8 Sample Analysis

Sieved soil samples may be analyzed either by XRF or by standard CLP methods. XRF analysis is recommended in order to conserve analytical costs and to reduce laboratory turn-around time. If XRF is used, this should be performed using a fixed-base XRF instrument with sufficient sensitivity and specificity to produce reliable concentration estimates in the range of 50-5000 ppm. As discussed earlier (see Section 2.5), this can be achieved using an energy-dispersive instrument with either direct rhodium excitation or rhodium excitation and a secondary target wheel using molybdenum or silver targets. In order to demonstrate that the results are comparable to those obtained by standard CLP methods, at least 5%-10% of the samples should be analyzed by both XRF and CLP.

Lead measurements obtained using a field portable XRF instrument may be useful in estimating likely soil lead levels and the degree of variability between sample locations (see Section 2.5), but should not be used to collect data for risk assessment purposes.

3.1.9 Geochemical Characterization

As discussed in Section 2.8, there are a number of potentially valuable information items that can be derived by geochemical characterization of soil samples. This includes data on the size distribution of lead-bearing particles within the samples, the mineralogical or chemical form of the lead in different particles (both on a particle frequency basis and on a mass fraction basis), and the "matrix association" of the particles (the extent to which the lead-bearing grains are encased in a rocky or glassy matrix of relatively inert mineral material).

Data of this type is helpful because it can be used to judge whether or not to extrapolate bioavailability data from one soil sample to another. This approach requires data from other studies (either in vivo and/or in vitro), and such data are now being generated by a number of different laboratories. As these data become available, the ability to make predictions about the bioavailability of lead purely on the basis of speciation data will increase. In addition, lead speciation data can sometimes be helpful in drawing inferences about the likely source of lead contamination. In particular, speciation data can help estimate the fraction of soil or dust lead that is attributable to paint chips or flakes, since the issues associated with paint exposure are different (both biologically and legally) from exposure to other types of lead. Note that it is sometimes possible to identify paint chip contamination of soil using only light microscopy rather than full electron microprobe analysis.

Based on these potential benefits, it is recommended that at least limited soil characterization studies be undertaken. These should include a subset of 2 to 6 samples that are believed to be reasonably characteristic of the site. Lead concentration of the samples is generally not critical, assuming the value is at least 200 ppm.

3.2   DUST

3.2.1   Overview

Although exposure to indoor house dust is generally considered to be one of the most important exposure pathways for young children, there is no standard method for collecting or analyzing dust samples. A thorough review and discussion of the many issues and options regarding dust sampling has been presented in EPA (1995d).

The issues regarding dust sampling are in most ways analogous to those surrounding soil sampling. The parameter needed for input into the IEUBK model is the arithmetic mean concentration of lead in indoor dust to which a child is (or may be) exposed. The basic strategy for sampling depends on which of the following scenarios is thought to be most reasonable:

1)   **Dust lead concentrations vary from location to location within the house, and exposure occurs preferentially in some locations**

   In this case, the concentration value required for the model is the <u>intake-weighted</u> average concentration across all locations where exposure occurs.

2)   **Dust lead concentrations vary from location to location within the house, and exposure is assumed to be random**

   In this case, the mean of samples from N random or systematically selected locations is the required input parameter.

3)   **Dust is uniformly contaminated with lead throughout the house**

   In this case, the mean of N samples from any location in the house yield the appropriate input value.

It is expected that Scenario 1 is the most likely to be realistic in most cases. However, calculation of an intake-weighted average concentration requires data (actual long-term average dust intake rates at each sub-location in the house) that are not likely to be available.   Data on time spent in a location might be used as a crude surrogate for intake at that location, but dust intake depends not only on time at a location, but also on dust loading (the amount of dust per unit surface area) and on the type of activity (the frequency of hand contact with dust-covered surfaces and the frequency of hand-to-mouth activity). Thus, this approach is of uncertain merit. Moreover, even if time- or activity-weighted intake could actually be calculated, the value would only apply to the specific child or children whose activity patterns were used to

derive the weighting factors, and not necessarily to any other children (current or future). Thus, it is not currently feasible (or desirable) to attempt to collect samples that are entirely adequate to satisfy the theoretical requirements of Scenario 1.

In contrast, Scenario 2 does not require data or assumptions about the time or activity-based exposure patterns of current or future children living in the house, and any set of randomly or systematically selected samples from reasonable exposure locations will usually yield a reasonable estimate of the required input value. Further, application of this assumption is analogous to and consistent with the assumption of random exposure to outdoor yard soil. Thus, this approach is usually the most appropriate, and is recommended as the default approach for dust sampling to support the IEUBK model.

If dust contamination is uniform (Scenario 3), then any set of dust samples from within the house will be appropriate. However, available data suggest that there can often be substantial variability in dust lead concentrations within a house (especially if there are indoor sources of lead), so it is not recommended that a sampling plan based on Scenario 3 be employed without data to support the validity of the assumption of uniform contamination.

3.2.2   Sampling Locations

Children can be exposed to dust at many different locations within a home, and any plausible exposure location is a candidate for dust sampling. However, because it is suspected that young children are likely to be exposed to dust predominantly while playing or crawling on floors, it is recommended that indoor dust sampling focus on floor dust. This includes any room where it is reasonable that a child might be exposed relatively frequently (e.g., family room, child's bedroom, kitchen, etc.), including both those with carpets and bare floors.

Some sampling studies also focus on surfaces such as counter tops, table tops, etc, and on the friction surfaces on windows and doors where dust from paint is likely to be generated. However, it is not expected that a young child is likely to have repeated and frequent contact with these areas, and dust samples collected from friction surfaces than contain lead-based paint are likely to contain lead concentrations that are much higher than the typical dust concentration to which a child is exposed. Therefore, collection of dust samples from these areas is not necessary for the purposes of running the IEUBK model. However, collection of such samples can be informative about indoor sources of lead in dust.

3.2.3   Sample Number

As discussed previously, there is no simple rule for knowing how many different samples should be collected, and whether these samples should be analyzed separately or composited. However, assuming that indoor dust concentrations are likely to be approximately as variable

as outdoor soil samples, it is recommended that the same general approach be used for indoor dust samples as outdoor soil samples. That is, a set of at least 3-4 samples should be collected from random locations indoors, with each sample being composed of dust from at least 3-4 sub-locations. Note that, depending on the sampling device, these sub-samples need not be obtained separately, but can be composited directly in the sampling device.

3.2.4   Sampling Device

As discussed in EPA (1995d), there are a wide variety of different dust collection techniques available, and there is as yet no single method that has been selected as most appropriate.

In general, collection methods may be divided into two categories: wipe sampling and vacuum sampling. Both approaches have advantages and limitations (see EPA 1995d). However, wipe sampling can only yield an estimate of lead loading (lead mass per unit area), while vacuum sampling can yield estimates on both loading and concentration. Because the IEUBK model currently requires data on dust concentration and does not presently employ data on lead loading in dust, wipe sampling is not currently appropriate for use in generating data to support the IEUBK model, and vacuum sampling must be used to derive a sample which can be weighed before analysis.

Of the numerous different vacuum collection devices that have been developed or adapted for use in dust sampling, the system developed by researchers at the University of Cincinnati has been most widely tested and applied, and this device is recommended for use by EPA (1995c). Based on these considerations, this device is also recommended for used in collecting samples for use in IEUBK-based risk assessments.

It should be noted that this sampling device does not usually serve to collect 100% of the dust from a sampling location, especially from rugs or other rough surfaces. Rather, because of the relative slow air flow and the lack of mechanical agitation, the device tends to obtain only the superficial dust in the carpet or on the surface. Although this is sometimes cited as a limitation of this device, it is actually more likely to be an advantage, since it is the superficial dust that can be easily removed that is likely to be the main source of exposure for children, rather than dust and dirt that is deeply ingrained or ground into carpets or floors. If a measure of total dust is considered to be required, then a sampling device with a high air flow and/or a mechanical agitator should be used. EPA 1995d provides a discussion of a variety of dust-sampling devices which may be employed.

An alternative approach for collection of dust samples is to obtain a sample from the household vacuum cleaner. This approach is simple and allows for easy collection of sufficient mass of dust for analysis. However, the location from which the dust was collected is unknown, and

therefore the dust may or may not be representative of that to which a child is exposed. Because of this uncertainty regarding sample representativeness, this approach is not recommended.

3.2.5 Sampling Procedure

A detailed protocol for dust sample collection using the University of Cincinnati sampling device is provided in EPA (1995c). In brief, this protocol involves applying a template at the area or sub-location to be sampled, and then vacuuming the area inside the template in a standardized pattern of vertical and horizontal passes with the collection nozzle. The size and shape of the template is variable, but a plastic form with a 25 cm x 25 cm opening is common (EPA 1995d).

Because the area of the template is known, the results of the sampling effort can be expressed either as concentration (ug lead per gram of dust) or as loading (ug of lead per m^2). Note that when concentration is to be measured, special steps are needed to obtain reliable weights for the filter before and after dust collection (see EPA 1995c). As discussed earlier, for the purposes of the IEUBK model it is required that an estimate of concentration be obtained, while estimates of loading are not presently used in the model.

Whenever possible, the mass of dust collected should be at least one gram, since this is sufficient to support either an XRF analysis or a wet chemistry analysis. Analyses can be performed on smaller samples, but accuracy and representativeness may both decrease as sample size decreases.

3.2.6 Sample Preparation

Depending on the source area vacuumed, the "dust" that is collected can consist of a wide variety of materials and a wide range of sizes. Because it is believed that small particles are more likely to adhere to the hands and become a source of ingestion exposure, it is desirable to sieve the sample through a 60-mesh nylon screen in order to remove coarse particles. This step can also serve to remove extraneous material such as carpet fibers, pet hairs, etc., from the sample. Note that because sieving cannot be achieved quantitatively, the results of an analysis of a sieved dust sample should only be expressed in terms of concentration and not in terms of loading. If the amount of material collected is small, this sieving step may be omitted.

3.2.7 Sample Analysis

As discussed in EPA (1995d), some researchers have used relatively mild acid solubilization techniques to extract lead from dust samples and have described the results as "bioavailable lead". This approach is not appropriate for use with the IEUBK model, since the model already

incorporates a parameter that accounts for the fraction of lead in dust that is bioavailable. Therefore, the analysis of the dust sample should be directed toward estimation of "total" lead. As discussed previously, this can be achieved using either standard "wet chemistry" methods that involve vigorous acid digestion, or by XRF techniques. If XRF techniques are employed, sieving to remove fibers, hairs and other light materials is strongly recommended.

3.2.8 Summary

The approach recommended for collecting indoor dust samples for use in the IEUBK model consists of the following key steps:

1. Select at least 3-4 locations within the house where it is reasonable that a child might be exposed to dust, now or in the future. Typically, emphasis should be placed on floor areas where a child may be expected to play or crawl on a regular basis.

2. At each sampling location, select at least 3-4 sub-locations at random. At each sub-location, carefully apply a template and collect dust from within the template using the standardized vacuuming technique. The samples from each sub-location should be collected sequentially on the same filter, generating a single composite sample for that location.

3. If sufficient material has been collected, the sample should be sieved to exclude coarse particles, fibers and other non-dust matter. Whether the sample is sieved or not, the mass of the sample must be determined before analysis.

4. The samples should be analyzed for "total" lead using standard wet chemistry or XRF analytical techniques.

5. The arithmetic mean dust lead concentration, expressed in units of ug/g (ppm), averaged across all samples collected from the house, should be used as input into the IEUBK model for that exposure unit.

3.3 WATER

3.3.1 Sample Selection

As discussed earlier (see Section 2.1), lead may enter water either from natural sources, environmental releases, or from pipes and plumbing fixtures. In most cases, investigations at Superfund sites are concerned mainly with the lead levels in water sources (ground water and/or

surface water), but collection of samples from household taps may also be needed if a correlation analysis of environmental lead levels with measured blood lead levels is being planned.

With respect to source water samples, lead can exist either dissolved or suspended. If the water is (or could be) available to residents through private wells, it is plausible the water could be consumed without any sort of treatment, so the total lead content (dissolved plus suspended) is the appropriate value. However, because filtration or settling of sediment is a common treatment technique even in private wells, the level of dissolved lead should also be measured. Standard methods for the collection and preservation of samples for total and dissolved metals are presented in SW-846 (see Method 3005). In brief, samples intended for total lead should be collected directly in a bottle and acidified by addition of nitric acid (5 mL/L). Samples intended for dissolved lead must be filtered through an 0.5 um filter at the time of collection before addition of nitric acid.

If samples of tapwater are needed, the amount of lead which enters the water from pipes and plumbing fixtures generally depends on how long the water has been in contact with the plumbing. Thus, water that is drawn in the morning ("first flush") is often more contaminated than water that is delivered after the pipes have been flushed out. Consequently, understanding the importance of exposure via tap water requires knowledge of the concentrations in both first flush and post-flush water (as well as knowledge on the intake rates for each water type). In accord with the definition established in the regulation that sets the Action Level for Lead (see Federal Register 56:26460, June 7, 1991), first flush water is defined as water that has stood in the supply system for at least six hours. Samples of first flush water should be collected directly from the tap as it is first turned on, without allowing any to go down the drain. The length of time needed to flush the lines before obtaining a "post-flush" sample varies from house to house, but running the water for 2-3 minutes or until the water changes temperature is usually sufficient.

3.3.2   Sample Analysis

There are a variety of standard CLP methods for the analysis of lead in water samples, including inductively coupled argon plasma emission spectroscopy (ICP), flame atomic absorption spectrometry (FAA), and graphite furnace atomic absorption spectrometry (GFAA). Both ICP and FAA have comparable detection limits for lead in water (usually about 20 to 50 ug/L), while GFAA usually has detection limits of 1-2 ug/L. For the purposes of the IEUBK model, detection limits of 5 ug/L or less are needed, since concentrations higher than this begin to contribute quite substantially to the total lead dose in children. Therefore, GFAA will normally be the method of choice for water. If samples are encountered with concentrations lower than the detection limit for GFAA, these can be assessed simply by assuming a concentration of one-half the detection limit, and more sensitive analytical techniques are not needed.

3.4   AIR

As noted earlier, lead levels in air are usually sufficiently low (on the order of 0.1 ug/m^3) that the dose absorbed from inhaled air is small (less than 1%) compared to the absorbed dose received from ingestion of soil, dust, food and water. Therefore, site-specific measurements of lead in air are often not critical for the purposes of running the IEUBK model. However, at sites where a significant airborne source of lead exists and an air sampling program is undertaken, the following guidance is recommended.

3.4.1   Sampling Device

The data item required for input into the IEUBK model is the long-term arithmetic mean concentration of lead in outdoor air. Because outdoor air concentrations are expected to vary as a function of numerous parameters (e.g., meteorological conditions, time of day, time of year, etc.), it is highly desirable to obtain data that support a meaningful estimate of the long-term average rather than only one or two samples taken over a short time span.

The sampling device best suited to collection of this type of data is the high-volume sampler. This device draws ambient air across a glass fiber filter at rate of about 40-60 cfm (1.1-1.7 m^3/min), and particulate matter is trapped on the filter. The device may be used either to collect repeated 24-hour samples, or may be left to run continuously for periods of 1-2 weeks (so long as the amount of particulate material collected on the filter does not cause a reduction in air flow rate). Detailed protocols for the proper use and operation of this sampling device are given in Appendix B of 40 CFR 50.

3.4.2   Analytical Method

Typically the filter is cut into portions (e.g., into quarters) and one portion is extracted with acid to solubilize the lead from particles trapped on the filter. This acid extract may then be analyzed by either GFAA or ICP. The detection limit for lead on the filter portion is usually about 10-20 ug. If the filter represents a 24-hour sample at a flow rate of 1.5 m^3/minute, this corresponds to a detection limit in air of less than 0.01 ug/m^3, which is more than adequate for the purposes of an IEUBK model run. If the filters are used to collect samples over an even longer time (e.g., one week), then the detection limit is corresponding lower.

It should be noted that samples of blank filters should also be extracted and analyzed in the same fashion as filters used in the high-volume samplers. This is because some filters may contain leachable lead, and correction for this contribution from the filter itself may be important in some cases.

3.4.3 PM$_{10}$ vs TSP

The calculations performed in the IEUBK model to estimate inhalation exposure from lead in airborne dust particles are based on the assumption that all of the particles are respirable. Therefore, the high volume sampler used to collect airborne particulates should be equipped with a device that separates out the largest particles and allows only particulate matter of approximately 10 um or less (PM$_{10}$) to reach the filter paper. If the sampling device collects total suspended particulates (TSP) rather than respirable particles (PM$_{10}$), then an adjustment factor to account for this may be appropriate. Data reported by Pace (1983) suggest that a default factor of 0.5 mg PM$_{10}$ per mg TSP is reasonable, although this ratio can be highly variable.

3.5   FOOD

Data from national marketbasket surveys indicate that children ingest about 5-7 ug/day of lead in the diet. Assuming a soil concentration of 400 ppm and a dust concentration of about 300 ppm, the typical diet contributes an average of about 20% of the total absorbed dose of lead for a child. Thus, any locally-raised food material that contains substantially higher-than-average levels of lead can be an important source of exposure. For example, if 20-25% of a child's fruits and vegetables were derived from a local source than contained 0.10 ppm lead, this local food pathway would contribute an extra absorbed dose of about 2.2 ug/day, an increase of about 15% in the total absorbed dose. Thus, analysis of locally-raised foods for lead can yield data that are important for correct exposure and risk analysis using the IEUBK model.

3.5.1 Locally-Raised Fruits and Vegetables

1.   Lead levels are likely to vary between different types of fruit and vegetable, so samples of all common types of produce grown at the site should be sampled. In this regard, different types of produce mature at different times of year, so the sampling program should not be just an opportunistic "grab" of what is available at any one time, but a systematic sampling throughout the harvest period. It is generally helpful to collect co-located samples of soil to help quantify the relationship between lead in soil and lead in vegetation.

2.   Fruits and vegetables should be washed to remove soil and dust adhering to exterior surfaces. This is probably best be achieved in the field using a squirt bottle of distilled water, but can be done in the laboratory if the samples are received promptly after harvesting.

3.   Assuming that ingestion rates of fruits and vegetables are about typical and that about 20-25% of the total intake is from local sources, a detection limit of about 50 ppb (0.050 ppm) fresh weight (wet weight) is required. There are no official EPA protocols for collection and preparation of locally-raised food samples, but the FDA has developed a method (LIB # 2043)

that has been in use for more than 20 years and has proved very successful in the "Total Diet" survey programs. The detection limit of this FDA method (summarized below) is usually about 20 ppb, and is recommended for use in analysis of nearly any food sample.

a. Weigh about 2.5-5 grams of fresh sample into a 100 mL quartz beaker.
b. Add 3 mL of 40% H_2SO_4, mix, and dry overnight at 110°C.
c. Transfer samples to muffle furnace. Heat for 2 hours at each of the following temperatures: 150°, 200°, 250°, 300°, and 350°. Then increase to 470° and heat overnight (12 hours).
d. Remove and cool sample. Add 4 mL of nitric acid and evaporate to dryness on a hotplate. Return to muffle furnace and heat at 470° for at least 2 hours.
e. Add 4 mL of nitric acid and about 20 mL of water. Heat on a hotplate to dissolve the residue. Transfer to a volumetric flask and dilute to 50 mL.
f. Analyze using GFAA. Report results as ug/gram wet weight.

3.5.2   Locally-Raised Livestock, Game Animals, Fish, etc.

If exposure to lead via ingestion of locally-raised beef, other livestock, game animals (e.g., deer) or fish is of potential concern at a site, the analytical approach recommended for sample analysis is similar to that described above for vegetables and fruits. Key points to remember are: 1) report all results as ug/g wet weight, and 2) strive for a detection limit no higher than about 50 ppb (wet weight) in the animal tissue.

3.6   PAINT

As noted previously (see Section 2.1), it is not usually necessary to collect data on lead levels in paint to support the IEUBK model, since indirect exposure to paint lead (i.e., via ingestion of contaminated indoor dust) is assessed by collection of dust lead data, and presently there is not enough information to allow a meaningful quantitative evaluation by the model of exposure and risk from direct exposures (i.e., ingestion of paint chips).

If data on paint lead levels are desired, the most appropriate sampling and analysis strategy depends on the objectives of the sampling effort. Both the EPA and the U.S. Department of Housing and Urban Development have developed detailed guidelines for obtaining data on paint lead levels to support the goals of Section 403 of the Toxic Substances Control Act (TSCA) and Title X of the Housing and Community Development Act (HCDA) (EPA 1994c, HUD 1995). However, the goals of these programs are somewhat different than are usually appropriate to support lead risk assessment and lead risk management decisions at Superfund sites. Therefore, the following guidance incorporates many sections of the existing guidance documents developed previously, with modifications as necessary to meet the goals of this handbook.

3.6.1   Sampling Locations

There are many different locations in a house where lead-based paint could exist, and a program intended to identify all such locations requires collection of a rather large set of measurements (50-200) within each house (EPA 1994c). However, if the purpose of the sampling effort is to identify current sources of concern to children, then attention can be restricted mainly to locations where paint is in poor condition (i.e., peeling, chipping, cracking), or where painted surfaces are subject to mouthing or chewing by children (e.g., window sills, stair steps, etc.). Therefore, the house should be inspected to identify locations with deteriorated paint or chewable surfaces that are accessible to children, and these should be sampled. This includes both the interior and the exterior of the house. Typically, only about five samples per house will be needed (HUD 1995). Note that friction and impact surfaces need not be sampled (HUD 1995), since the contribution from these surfaces is already evaluated through collection of dust samples.

3.6.2   Analysis Method

Field Portable XRF

Both EPA and HUD recommend that data on paint lead levels be collected with a field portable XRF instrument. There are two basic types of field-portable XRF instruments: those that measure L-line and those that measure K-line x-ray fluorescence. In general, instruments that measure L-line x-ray fluorescence only detect lead in the outer-most surface layers of paint and not in deeper layers, while instruments that measure K-line fluorescence detect lead in all of the layers of a paint sample. Therefore, use of instruments that measure K-line fluorescence are recommended to assess risks from direct paint chip ingestion. However, measurements based on L-lines fluorescence may also be helpful, since surface layers are the most likely to be contributing to indirect paint exposure via chipping, peeling and chalking.

As with any analytical procedure, it is important that measurements be obtained by properly trained and experienced personnel. Detailed protocols are given in HUD (1995). In brief, a measurement of lead level in paint at a potential source area is achieved by making three independent measurements in adjacent areas of the suspected source, with each reading being at least 15 seconds in duration. The mean of the three readings in then taken as the value for that area. The instrument must be calibrated and validated using a standard paint sample (1.02 mg/cm^2) available from NIST, and subtraction of the instrument reading from clean substrate (i.e., an area of wall or other substrate with all paint removed) may be needed in some cases.

The result of an XRF measurement of lead in paint is usually expressed as mg/cm^2. The current definition of lead-based paint is 1 mg/cm^2 (EPA 1994c), and most modern instruments have quantitation limits lower than this.

Wet Chemistry

In some cases, measurement of lead levels may not be possible using a field-portable XRF instrument. For example, curved surfaces or surfaces where there is less than 3-4 square inches of intact paint cannot be reliably evaluated using XRF. In these cases, or if confirmation of XRF data is required, samples of paint may be removed and analyzed in the laboratory.

The area selected for sampling should be marked off and all of the paint down to the substrate should be collected. Typically, the area sampled should be between 1 to 4 square inches, depending on the mass of paint that exists per unit area. In any event, the area from which the paint is removed must be measured carefully so that the results can be expressed in units of mg/cm^2. If it is not possible to measure the area accurately, then the result can only be expressed in terms of mg/g or percent by mass in the paint.

Compositing of paint samples to reduce analytical costs is possible and permissible (HUD 1995), but is not generally recommended. This is because lead concentrations may vary substantially from location to location, and mixing samples of paint from areas with high lead and non-lead paint makes it very difficult to identify sources. If compositing is employed, it is recommended that no more than 5 sublocations be combined, and the area sampled at each sublocation should be equal. Note that if several sub-samples are composited, decisions about the possible presence of leaded paint must be based on the assumption that only one of the subsamples is "positive" (i.e. greater than 1 mg/cm^2). For example, if four subsamples were composited, any result above 0.25 mg/cm^2 would be consistent with the hypothesis that at least one of the subsamples contained lead at more than 1 mg/cm^2, and a re-analysis of the separate subsamples would be needed to determine if this was so. However, if the result was less than 0.25 mg/cm^2, then it can be concluded that none of the subsamples could have lead above 1 mg/cm^2. Therefore, compositing is likely to be most useful when it is suspected that the samples to be composited do not contain lead-based paint.

All samples of paint removed for analysis should be sent to a laboratory recognized by EPA's National Lead Accreditation Program. Field kits for measuring lead in paint samples are not recommended at present (EPA 1994c, HUD 1995).

4.0   DATA REDUCTION

4.1   Overview

It is important to stress that the environmental concentration input parameters to the IEUBK model are all intended to be arithmetic means. This is true even if (as is often the case) the data for a medium (e.g., soil) tend to be distributed in an approximately lognormal fashion. This is because the long-term (e.g., one year) average intake of a medium by an individual at an exposure unit is given by the average intake rate multiplied by the arithmetic mean concentration, and not by the intake times the geometric mean, mode or any other estimate of central tendency.

However, just because the target input concentration values are arithmetic means does not imply that an IEUBK model run can be performed only if a highly accurate mean value is available for each input parameter. As is the case in nearly all risk assessments, if some or all of the input estimates are uncertain, this should be discussed and evaluated as part of the risk characterization process.

Also note that input parameters called for in the model for soil, dust, etc, are means and not the 95% upper confidence limits (UCLs) of the mean, as is the case for other chemicals evaluated according to standard EPA Superfund guidance (EPA 1992b). This is because the IEUBK model has been developed and validated at a number of sites using arithmetic mean environmental concentration data, and use of UCL values would presumably tend to cause an overestimation of true exposure and risk. However, it is important to note that the purpose of using the UCL rather than the mean is to account for <u>uncertainty</u> in the estimate of the mean based on a limited number of samples, and this objective does not disappear simply because the chemical is lead rather than some other chemical. Therefore, when data on lead levels in a medium are sparse and screening-level evaluations are needed to determine whether lead exposure could be of concern, use of UCLs rather than means may be warranted. However, when the difference between the UCL and the mean is large and the level of risk depends heavily of which value is used, it is recommended that additional sampling be performed rather than attempting to make a decision based on limited data.

More detailed discussions on data reduction techniques for each medium are provided below.

4.2   Soil

For soil, the input parameter needed by the IEUBK model is the arithmetic mean concentration averaged over an exposure unit. Recall that an exposure unit is defined as a location where exposure of a child is known or assumed to be random, and in the case of childhood exposure

to lead is often taken to be the size of a residential property. Note that the arithmetic mean concentration averaged over larger size areas (e.g., over an entire community) is <u>not</u> an appropriate input to the IEUBK model (unless it is assumed a child is randomly exposed across the entire community).

However, just because the desired input parameter is the arithmetic mean at a residence does not necessarily mean that the estimated mean value for a property must be based only on data obtained from that property. This is because geostatistical techniques such as krieging can be used to extrapolate data obtained at a fixed set of sampling locations to estimate likely concentrations at locations intermediate between the sample locations. This type of approach is most useful when there are clear spatial patterns of lead contamination in soil (e.g., a "footprint" of lead contamination resulting from stack fallout), and is less helpful if past waste disposal practices or earth-moving activities have resulted in an un-predictable pattern of "hot spots" mixed with "cold-spots". It is recommended that whenever the area of potential concern is large and detailed sampling of each residence is undesirable or impossible that an expert geostatistician be consulted to identify the most appropriate geostatistical techniques to derive meaningful estimates of mean concentration as a function of location.

4.3   Dust

For dust, the input parameter needed by the IEUBK model is the arithmetic mean concentration averaged over all indoor locations at a specific residence (exposure unit) where a child might reasonably be exposed to dust. Thus, if data are available on dust lead levels in a number of different homes, it is theoretically inappropriate to combine or average these values across different houses. Rather, the mean dust lead value measured at each residence should, at least in theory, be used to assess the risk at that residence.

However, this approach ignores the fact that the dust lead level at a residence is probably not constant, but varies as a function of parameters such as the number of children and pets presently living at the house, the frequency of dusting and cleaning, the extent to which windows and doors are left open or closed, the degree to which leaded paint surfaces are maintained, etc. Thus, even a properly collected set of dust samples from a residence may not be predictive of what future indoor dust lead levels might be. For this reason, an alternative approach for employing dust lead data in the IEUBK model is to analyze the <u>average</u> relationship between soil lead in the exterior yard and the level of lead in indoor dust.

As discussed in greater detail in EPA 1995a, indoor dust contamination can thought of arising from two basic sources: local yard soils and all other (non-yard) sources, including both indoor sources and area sources. Based on this concept, the concentration of lead in indoor dust at a residence can be expressed in the following mass balance equation:

where:

$$C_d = \frac{m_s \cdot C_s + m_x \cdot C_x}{m_s + m_x}$$

    m   =   mass of material in dust derived from either yard soils (m_s) or non-yard sources (m_x).

    C   =   concentration of contaminant in interior dust (C_d), yard soils (C_s), or non-yard sources (C_x).

This equation can be re-written as follows:

where:

$$C_d = k_s C_s + k_x C_x$$

    k   =   mass fraction in dust of material derived from yard soil (k_s) and non-yard sources (k_x).

Because it is not generally possible to estimate the values of k_x and C_x separately, it is usually simplest to analyze paired soil-dust data using a simple two-parameter model, as follows:

where:

$$C_d = k_0 + k_s \cdot C_s$$

    k_0 =   Contribution to indoor dust from non-yard soil sources (ppm). Note that k_0 is equal to $k_x \cdot C_x$.

    k_s =   mass fraction of yard soil in indoor dust (unitless)

    C_s =   concentration in yard soil (ppm)

Thus, given a reliable set of paired soil/dust measurements at a number of different residences in a community, the average contribution of yard soil to dust lead can be estimated by simple linear regression analysis. An example is provided in Figure 4-1. Given this average relationship, the average concentration of lead in a house can the be calculated from the mean value measured in exterior yard soil at the residence. This approach is most useful for assessing risks at residences where dust lead data were not obtained (including hypothetical future residences), and may also be used to calculate values to replace the resident-specific values measured at current residences.

However, it is important to recognize that simple linear regression techniques are susceptible to measurement error in the data. Measurement error is the difference between the true mean value and the measured value at a location, and can be the result of errors (differences) due to analytical variability and/or to sampling variability. Typically, sampling variability (i.e., the chance that a soil sample drawn from a yard does not have a concentration equal to the true mean for that yard) is likely to be the largest source of measurement error. As the amount of measurement error increases, the apparent slope of the line (k_s) will decrease and the apparent intercept (k_0) will increase. If measurement error is likely to exist in a data set, it is recommended that an expert statistician be contacted to discuss ways that the potential problem can be minimized.

4.4   Water

As discussed previously, lead levels in current or potential sources of drinking water may vary as a function of location (proximity to a source of lead contamination), depth, and time. This variability in both time and space can pose a difficult challenge in the data reduction process for water (EPA 1993c).

Ideally, sufficient data would be available to define the long-term mean concentration at each location (e.g., each well), and exposure could be assessed separately at each location. This well-by-well approach is based on the expectation that residents will drink water from a single well and not at random from multiple wells, so averaging across wells is no more appropriate than is averaging soil lead levels across different yards. However, it is usually the case that there are insufficient data to derive meaningful long-term average for each potential drinking water source, and it is desirable to combine data across both time and location. As discussed in EPA 1993c, the best approach in this case is to identify wells that are most likely to be in or near the center of any groundwater plume of lead contamination, and to combine all available data across these wells. An example of this approach is shown in Figure 4-2. Note that it is not correct to combine data across wells that are located in different parts of the plume, since the result would be meaningful only in the case of a person moving randomly between these dissimilar locations.

If samples of tap water are collected from current residences, and if samples of both first-flush and post-flush are collected, the IEUBK model provides a menu for entering these two concentration values under the heading "Use Alternative Water Values ?" in the water data input screen. The screen also calls for an estimate of the fraction of total tap water intake that is first flush, and the fraction that is from drinking fountains. (The fraction that is post-flush household water is calculated by difference). The default values presently in the model are 50% first-flush 35% post-flush, and 15% fountain water. However, these values may be overly conservative, since 50% first-flush is considered to be an RME rather than an average value. Based on this, the Technical Review Workgroup for Lead (EPA 1992c) recommended that the default water intake assumptions be 30% first flush and 70% post flush (assuming that young children do not have frequent exposure to drinking fountain water). Whenever possible, these assumptions should be replaced with site-specific data derived from demographic studies of area residents.

4.5   Air

Similar to the case discussed above for water, lead levels in air are expected to vary as a function of time as well as a function of location in relation to sources of particulate emissions. In the simplest case (e.g., when the source is not very close to the exposure area), the differences in time will be large compared to differences in space, and all data can be combined to yield a single estimate of the mean concentration in a neighborhood. If the source is close

to the exposure areas of concern, however, then there might be significant differences as a function of location, so only data from locations that are within the same exposure unit should be combined.

4.6   Food

The IEUBK model accepts input on dietary exposure to lead in local foods via the data entry screen for DIET using the menu choice labeled "Use Alternative Diet Values ???". The food categories include fruits, vegetables, fish and game. The latter category can also be used for locally raised beef or any other locally raised meat source. In each category, the input parameters required are 1) the mean concentration of lead in the food material (ug/g fresh weight) and the fraction of the total intake of the food type that is derived from local sources.

In the case of fruits and/or vegetables raised in a home garden, the concentration of lead in the produce is presumably dependent on the concentration of lead in the soil, and so theoretically it is not appropriate to average or combine data for vegetables or fruits grown in different locations. However, because data will usually not be sufficient to calculate a meaningful average concentration for each garden (and because there will not be gardens at all locations), it is recommended that concentration data from vegetables or fruits be stratified according to the lead concentration in the soil (e.g., < 500 ppm, 500-1500 ppm, 1500-4500 ppm, > 4500 ppm) and that an average concentration be calculated for each stratum. Assuming that data are available for a number of different types of produce and that lead levels vary from type to type, it is desirable to calculate an intake-weighted average rather than a simple average across all samples within a stratum. Data on average intakes rates of each vegetable or fruit type should be derived either from site-specific demographic data or else from default national average data provided in the Exposure Factors Handbook (EPA 1996). An example of such an analysis is provided in Figure 4-3. If the data are too sparse to support such a calculation, or if the data suggest that lead levels are roughly similar in different types of produce grown in similar soils, then intake-weighting is not necessary and the simple mean across all vegetable samples within a stratum may be used.

The same basic approach is appropriate for beef and other locally-raised livestock. If data are available for animals raised at locations with significantly different soil lead concentrations, and if the data indicate that lead levels in the livestock vary as a function of the soil lead, level, then stratify the data by soil level and average lead concentrations within each stratum. If the data are not sufficient to support a stratification (or suggest that stratification is not needed), then simply calculate an average across all samples. As above, if data are available for different types of livestock, then the desired statistic is the intake-weighted mean to account for different ingestion rates of the different types of meat, although intake-weighting may not be possible or necessary in all cases. National average intake rates for most types of food are provided in EPA 1996.

In the case of game animals, the concentration of lead to which they have been exposed will usually not be known, and it is likely the animals have ranged over a wide area of varying lead concentrations. Therefore, no stratification is needed in this case.

5.0   REFERENCES

Brody DJ, Pirkle JL, Kramer RA, Flegal KM, Matte TD, Gunter EW, Paschal DC. 1994. Blood Lead Levels in the US Population.   Phase 1 of the Third National Health and Nutrition Examination Survey (NHANES III, 1988 to 1991). JAMA 272:277-283.

CDC. 1991. Centers for Disease Control. Preventing Lead Poisoning in Young Children. A Statement by the Centers of Disease Control - October. U.S. Department of Health and Human Services. Public Health Service.

Davis A, Ruby M, Bergstrom P. 1992. Bioavailability of Arsenic and Lead in Soils from the Butte, Montana, Mining District. Environ. Sci. Technol. 26:461-468.

EPA.   1982.   An Exposure and Risk Assessment for Lead. U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office. EPA 440/4-85/010.

EPA. 1986. Air Quality Criteria for Lead. June, 1986, and Addendum, September, 1986. U.S. Environmental Protection Agency, Office of Health and Environmental Assessment, Research Triangle Park, NC. EPA 600/8-83-028F.

EPA. 1987. Preliminary Review of the Carcinogenic Potential of Lead Associated with Oral Exposure. Internal review Draft. U.S. Environmental Protection Agency, Office of Health and Environmental Assessment. OHEA-C-267.

EPA. 1988. Drinking Water Regulations: Maximum Contaminant Level Goals and National Primary Drinking Water Regulations for Lead and Copper; Proposed Rule. U.S. Environmental Protection Agency. Fed. Regist. 53:31516-31578 (August 18, 1988).

EPA.   1989a. U.S. Environmental Protection Agency, Office of Emergency and Remedial Response. Risk Assessment Guidance for Superfund. Volume I. Human Health Evaluation Manual (Part A). EPA Document EPA/540/1-89/002.

EPA.   1989b. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards.   Review of the National Ambient Air Quality Standards for Lead: Exposure Analysis Methodology and Validation. QAQPS Staff Report. EPA-450/2-89-011.

EPA.   1990. U.S. Environmental Protection Agency, Office of Emergency and Remedial Response. Risk Assessment Guidance for Superfund. Volume I. Human Health Evaluation Manual (Part B, Development of Risk-Based Preliminary Remediation Goals). EPA Document EPA/540/R-92/003. Publication 9285.7-01B.

EPA. 1991a. Three Cities Urban Soil-Lead Demonstration Project. Midterm Project Update. U.S. Environmental Protection Agency (OS-240) 21S-2001. May, 1991.

EPA. 1991b. Technical Support Document for Lead. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response.

EPA. 1991c. Update on OSWER Soil Lead Cleanup Guidance. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response. Memorandum from Don R. Clay, Assistant Administrator. August 29, 1991.

EPA. 1992a. U.S. Environmental Protection Agency, Office of Emergency Response and Remediation. Guidance for Data Usability in Risk Assessment (Part B). PB92-963362. May, 1992.

EPA. 1992b. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response. Supplemental Guidance to RAGS: Calculating the Concentration term. Publication 9285.7-081. May, 1992.

EPA. 1992c. U.S. Environmental Protection Agency, Technical Review Workgroup for Lead. A TRW Report: Review of the EPA Uptake Biokinetic Model for Lead at the Butte NPL Site. October, 1992.

EPA. 1993a. Office of Solid Waste and Emergency Response. Data Quality Objectives Process for Superfund. PB94-963203. September, 1993.

EPA. 1993b. U.S. Environmental Protection Agency, Office of Research and Development. An X-ray Fluorescence Survey of Lead Contaminated Residential Soils in Leadville, Colorado. EPA/600/R-93/073. March, 1993.

EPA. 1993c. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response. Supplemental Guidance to RAGS: Improving Risk calculations for Groundwater Contamination. Draft. August, 1993.

EPA. 1994a. Guidance Manual for the Integrated Exposure Uptake Biokinetic Model for Lead in Children. U.S. Environmental Protection Agency, Office of Emergency and Remedial Response. Publication Number 9285.7-15-1. EPA/540/R-93/081.

EPA. 1994b. Revised Interim Soil Lead Guidance for CERCLA Sites and RCRA Corrective Action Facilities. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response. Memorandum from Elliot P. Laws, Assistant Administrator. July 14, 1994.

EPA. 1994c. Guidance on Residential Lead-Based Paint, Lead Contaminated Dust, and Lead-Contaminated Soil. U.S. Environmental Protection Agency, Office of Prevention, Pesticides and Toxic Substances. Memorandum from Lynn R. Goldman, M.D., Assistant Administrator. July 14, 1994.

EPA. 1994d. Technical Background Document for Soil Screening Guidance. Review Draft. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response. EPA/540/R-94/106.

EPA. 1995a. Standard Operating Procedure. Evaluating Exposure from Indoor Dust. U.S. Environmental Protection Agency. Region 8 Superfund Technical Section.

EPA. 1995b. Urban Soil Lead Abatement Demonstration Project. Review Draft. U.S. Environmental Protection Agency, National Center for Environmental Assessment. EPA/600/R-95/139.

EPA. 1995c. Residential Sampling for Lead: Protocols for Dust and Soil Sampling. Final Report. U.S. Environmental Protection Agency, Office of Pollution Prevention and Toxics. EPA 747-R-95-001.

EPA. 1995d. Sampling House Dust for Lead. Basic Concepts and Literature Review. U.S. Environmental Protection Agency, Office of Prevention, Pesticides and Toxic Substances. EPA 747-R-95-007.

EPA. 1996. Exposure Factors Handbook. Volume II - Food Ingestion Factors. U.S. Environmental Protection Agency, Office of Research and Development. Draft Report, April 1996. NCEA-W-0005.

Gilbert RO. 1987. Statistical Methods for Environmental Pollution Monitoring. Van Nostrand Reinhold., New York. p. 174.

HUD. 1995. Guidelines for the Evaluation and Control of Lead-Based Paint Hazards in Housing. Washington, DC, U.S. Department of Housing and Urban Development. June 1995.

Kissel JC, Richter KY, Fenske RA. 1996. Investigation of Factors Affecting Soil Adherence to Skin Using a hand-Press Technique. Bull. Environ. Contam. & Tox.(submitted for publication).

Pace TG. 1983. The Use of TSP Data to Estimate PM_{10} Concentrations. Memo from Thompson G. Pace, USEPA Air Management Technology Branch to John D. Bachmann, Strategies and Air Standards Division. January 25, 1983.

Pirkle JA, Brody JD, Gunter EW, Kramer RA, Paschal DC, Flegal KM, Matte TD. 1994. The Decline in Blood Lead Levels in the United States. JAMA v. 272 (4): pp. 284-291.

Sverdrup. 1995. Uptake of Lead and Arsenic by Garden Vegetables. Study Report for the Kennecott Site. Report prepared by Sverdrup Corporation (Maryland Heights, MO) and Life Systems Inc (Cleveland, OH) for USEPA Region VIII. Dated February 3, 1995.

TABLE 1-1 DEFAULT EXPOSURE ASSUMPTIONS FOR IEUBK MODEL

Medium	Parameter	Age (years)						
		0-1	1-2	2-3	3-4	4-5	5-6	6-7
Air	Breathing Rate (m^3/hr)	2	3	5	5	5	7	7
	Time outside (hr/day)	1	2	3	4	4	4	4
	C_{in}/C_{out}	0.30	0.30	0.30	0.30	0.30	0.30	0.30
	Absorption Fraction	0.32	0.32	0.32	0.32	0.32	0.32	0.32
	Default concentration (ug/m^3)	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Diet	Daily lead intake (ug/day)	5.53	5.78	6.49	6.24	6.01	6.34	7.00
	Absorption fraction	0.50	0.50	0.50	0.50	0.50	0.50	0.50
Drinking Water	Ingestion Rate (L/day)	0.20	0.50	0.52	0.53	0.55	0.58	0.59
	Absorption fraction	0.50	0.50	0.50	0.50	0.50	0.50	0.50
	Default concentration (ug/L)	4	4	4	4	4	4	4
Soil/Dust	Total daily intake (mg/d)	85	135	135	135	100	90	85
	Fraction of total that is soil	0.45	0.45	0.45	0.45	0.45	0.45	0.45
	Fraction of dust derived from soil	0.70	0.70	0.70	0.70	0.70	0.70	0.70
	Absorption fraction	0.30	0.30	0.30	0.30	0.30	0.30	0.30
All	GSD_i	1.6	1.6	1.6	1.6	1.6	1.6	1.6

TABLE 2-1 CHARACTERISTICS OF XRF INSTRUMENTS FOR LEAD ANALYSIS

Instrument Characteristic	Instrument Type			
	WDS-XRF	EDS-XRF$_1$	EDS-XRF$_2$	FP-XRF
Detector	LiF crystal	Si-Li	Si-Li	Si-Li
Excitation source	Rh tube (direct)	Monochromatic Mo/Ag	Rh tube (direct)	Radioisotope
X-ray line	L alpha	L alpha or L beta	L alpha	L alpha
Environment	Vacuum	Vacuum	Vacuum	Air
Interferences	None	Minor (As, Bi, Se)	Significant (As, Bi, Se) Higher background	Peak overlaps High background Low count rates
Sample preparation	Complex (fused or pressed pellets)	Simple (loose dried powders)	Simple (loose dried powders)	Very simple (in situ or loose soil in plastic bag)
Detection limit (low arsenic) (high arsenic)	2 ppm 2 ppm	5 ppm 15 ppm	15 ppm 50 ppm	250 ppm 500 ppm

WDS-XRF = Wavelength dispersive system
EDS-XRF$_1$ = Energy dispersive system with Rh excitation of secondary target wheel
EDS-XRF$_2$ = Energy dispersive system with direct Rh excitation of sample
FP-XRF = Field portable system