Bioaccessibility Research Canada

Discussion Document
Considerations for Bioavailability Experiments

Introduction

Regulations and guidelines used in contaminated site assessment and associated remediation are largely based on minimizing human health risks. In most cases, the total concentration of a target substance in a particular substrate (e.g. soil, sediment or water) has been used to derive the regulations and guidelines. A default assumption is that 100% of a contaminant in soil is bioavailable. Bioavailability is the extent to which a substance can be absorbed by an organism into its systemic circulation. While chemicals that are dissolved in water may be readily available for uptake by various organisms, this may not be the case for soils and sediment where the contaminant may be tightly bound to the soil matrix. It can be argued that only the bioavailable fraction of a contaminant has the ability to exert an adverse effect. Bioavailability can only be determined by using in-vivo (usually animal) studies. Laboratory (in-vitro) tests are being developed to model exposure in the gastro-intestinal environment; the fraction of the contaminant that dissolves in such tests is referred to as the bioaccessible portion.

As stated above, bioavailability is the general term for the absorption of a contaminant into the body of a human or test subject. Absolute bioavailability is the proportion of the contaminant in the administered dose that is absorbed. The test subject, dose administration method, and measurement of absorption can vary, and for this reason, as well as the need to obtain information that is relevant to risk assessment, the relative bioavailability is more commonly reported. The relative bioavailability is the proportion (usually expressed as a percentage) of the contaminant that is bioavailable, relative to the proportion of a standard or control substance that is bioavailable. The standard or control experiments usually involve the use of a soluble form of the contaminant of interest, and sometimes also a mode of administration that ensures complete absorption (e.g., intravenous). The applicability of relative bioavailability measurements in risk assessment is enhanced when control experiments represent experiments from which toxicity reference values (used in risk assessments) were derived.

The concept of “validation” is introduced as a means to ensure the appropriateness of laboratory bioaccessibility tests. In this context “validation” means a favourable comparison of the bioaccessibility results to relative bioavailability. A favourable comparison could mean agreement between

bioaccessibility and bioavailability results, or a predictable relationship between a series of bioaccessibility and bioavailability results.

This document summarizes some ideas and statements compiled from discussions with Stan Casteel, professor in Research Toxicology, Department of Veterinary Pathobiology, College of Veterinary Medicine, University of Missouri. Dr. Casteel is an expert in bioavailability testing and is one of the major contributors to the known published bioavailability data for inorganic elements. Dr. Casteel wrote most of the text that follows, Iris Koch added some of the discussion comments, and Ken Reimer and Viviane Paquin provided editing.

Topic 1: Appropriate Animal Models

Species of mammals have unique anatomical, physiological, genetic and metabolic characteristics which controls when and how they can be used as surrogate models for humans. Selection of the most appropriate animal model will enhance the science and reduce extrapolation error and uncertainty in human risk assessment. Criteria or factors known to be associated with bioavailability determinations of lead or chemical contaminants in contaminated matrices can be classified into three broad categories:

1. Geochemistry/sampling/*in vitro* solubility/bioaccessibility factors
2. Animal model biological, bio-methodological, and research management concerns
3. Human factors, such as the age and sex of the population selected from comparison.

The juvenile pig model is often the model of choice. Aspects of this model that make it uniquely attractive for determining the bioavailability of lead from contaminated media and targeting childhood include:

1. The lack of coprophagy (feces ingestion) and the anatomical and physiological differences associated with this behavior seen in rats (who consume feces from the cage floor) and rabbits (who consume feces directly from the anus)
2. The absence of complicating factors connected with the relatively high biliary excretion of Pb, such as the case with rats and dogs
3. The absence in swine of rapid postnatal developmental changes in the active transport mechanism for lead across the intestinal barrier; this is a problem observed in juvenile rats (narrow juvenile window/early rodent epiphyseal closure phenomenon at end of adolescence)
4. The similarity of immature swine in physiologic age and body weight to the human childhood population
5. The ease of serial blood sampling without risk of anemia in swine. The size and tractable nature of young pigs facilitates repeated blood sampling of ample volume (5-7 mL) for analysis and archiving, and for implantation of intravenous catheters. Similarities in gastrointestinal

physiology and feeding behavior are additional aspects which makes juvenile swine the model of choice. See Table 1 for other considerations and examples which apply specifically to lead.

Table 1. Comparative data on the swine biomedical research model and three other popular research animal models

Parameter	Swine	Rabbit	Rodent (Rat)	NHP*
Behaviour				
Soil Ingestion	Yes	No	No	No
Coprophagic	No	Yes	Yes	No
Omnivorous	Yes	No	Yes	Yes
Intermittent(I) or continuous feeder (C)	I	C	C	I
Research Tractability	Good	Good	Good	Poor
Gastrointestinal Physiology				
Biliary excretion of Pb	Low	Low	Marked	Low
Stomach pH (fasting)	1.8-4.0	1-2	3.8-5	2.8-4.8
GI transit time total (hr)	39-48	27-39	23-27	
Digestion physiology/species irregularities	None	Yes**	Yes***	None
Cecal fermenter	No	Yes	Yes	No
Colon fermenter	Yes	No	No	No
Skeletal Development				
Fetal ossification commences	35d		17-18d	56d
Main fetal skeletal ossification period	60-70d			
Epiphyseal plate closure (long bones)	20-22 mo		None	2-2.5 yr
Skeleton weight as % of body weight	13	Low (7)		
Physiology and Metabolism				
Resting heart rate (beats/min)	80-125	220-380	250-450	150
Adult metabolism (kcal/kg-d)	17.8-38.6	110	130	158
Growth and Nutrition				
Birth weight	1-1.5 kg	60 g	5-6 g	465-490 g
Weaning weight in kg (age in weeks)	9.1 (3)	1.5 (8)	0.04-0.05 (3)	0.9 (12-27)
Feed usage (g/d)	2270	175-225	25-30	150
Reproduction/ Sexual Maturity				
Female sexual maturity age (weight)	6-8 mo (113.6 kg)	5-6 mo (4 kg)	100 d (200 g)	5 y
Reproductive lifespan of female (y)	2-6	5-6	1	12-15
Average litter size at birth (range)	10 (6-18)	8 (1-18)	10 (8-17)	1
Natural lifespan in years (maximum)	10-15 (27)	6-15 (15)	2-3 (4)	16-30 (30)
Economics of Research Maintenance				
Per diem costs	Low	Low	Low	High
Caging costs	Low	Mod†	Mod	High
Feed and bedding costs	Low	Mod	Mod	High
Hours required for care	Low	Mod	Mod	High
Environment enrichment requirements	Minor	Minor	Minor	High
Predictiveness/Relevance to Pb Effects in Humans				

Parameter	Swine	Rabbit	Rodent (Rat)	NHP*
Bioavailability in children	Best	Poor	Poor	Good
Functional neurotoxicity	Poor	Good	Good	Best
Fetal Bioavailability	Good	Poor	Poor	Best
Other Factors				
Ease of repeated blood collections	Good	Fair	Poor	Poor
Ease of oral dosing (non-gavage)	Good	Fair	Fair	Poor
Viviarium zoonotic diseases (both directions: anthropozoonoses and zooanthropozoonoses)	Few	Few	Some	Many
Animal rights groups interest	Low	Mod	Mod	Strong

*NHP = Non-Human Primate (Rhesus macaque; *Macaca mulatta*). Many of the parameter judgments would also apply to other species of NHP, including New World Primates (NWPs); NWPs are generally smaller in size and consume less monkey chow than OWPs.

**Rabbit is primarily a hindgut fermenter/digester; rabbits practice cecotrophy, i.e., the ingestion of night feces; rabbits produce large volumes of bile (7 times the volume of the dog).

***Rats practice coprophagy and have some degree of hindgut digestion, lack a gall bladder but produce bile, and actively excrete large relative amounts of Pb into bile.

†Moderate, low, or high/strong; and good, fair, or poor; and minor or high; and few, some, many; and poor, good, or best, are all assessed arbitrarily, based on author's experience.

Topic 2: Dosing of Animals

A bioavailability experiment typically involves dosing the animals (e.g., 3 animals) with the substance (usually soil) at one or more doses of the substance. Typically a 2 or 3 dose approach is used to account for any instability of bioavailability estimates over a range of doses. The contaminant is measured in an appropriate tissue or organ (see next topic), and the measured concentration is a response. A dose-response curve is thus obtained. For lead in blood, the area-under-curve (AUC) dose-response is best fitted to an exponential equation. For dose response curves obtained with tissue lead (e.g., bone, liver and kidney), best fit curves are obtained using linear equations. Each data point in the blood lead AUC (and tissue lead concentrations) reflects the group mean exposure and the group mean response, with the variability in dose and response shown by error bars.

Even though the toxicity reference values are often based on acute exposure experiments, exposures in the real world are chronic. Using a multiple dosing regimen helps to mimic this condition. More realistic exposures to contaminated material are also modeled by utilizing several days dosing. This method is acceptable to regulatory agencies who regard the robustness of the data as most useful for making decisions.

Topic 3: Organs or Tissues Measured

Some organs or tissues that are measured for contaminant concentrations are more convenient for certain contaminants, but as discussed in Topic 4, it is important that the control experiments be

conducted in the same way. For example, for arsenic (this element will be discussed in more detail in Topic 4), the urinary excretion fraction is used as a convenient and reliable way to assess the bioavailability of arsenic, since the amount of arsenic excreted approximates the amount ingested after 4-5 days of dosing in which a pseudo-steady state is reached.

Blood, liver, kidney, and bone are used to assess the bioavailability of lead. A linear model gives the best fit for liver, kidney, and bone data. For the blood lead AUC endpoint, the linear model is a poor fit. In general, each of three nonlinear models (exponential, Michaelis-Menton, and power) all tended to give similar results in terms of RBA value (the standard deviation in relative bioavailability for a particular test material averaged across the three models was usually less than 3%) and differences in the Akaike's Information Criteria (AIC) for goodness of fit were usually small. On this basis, it was concluded that any of these three models would be acceptable for lead. The power model was not selected because it does not tend toward a plateau, while data from early blood lead pilot studies using higher doses suggest that the blood lead endpoint does tend to do so. Of the remaining two models (exponential and Michaelis-Menton), the exponential model was selected mainly because it yielded the best fit more often than the Michaelis-Menton model. Thus, the exponential model was selected for application to all dose-response data sets for the blood AUC endpoint.

Topic 4: Control Experiments

Bioavailability assessment experiments should always include both positive and negative control groups. Control experiments serve as a basis for quantitative estimation of the effects of the independent variable (the contaminant of interest in the substance being tested) in excess of those effects produced by non-specific changes in the environment or the experimental system. A test material may have non-specific effects like those of the negative control, but may also have specific effects that can be attributed to the contaminant that is unique.

Careful use of negative controls in an experiment prevents erroneous conclusions about the apparent activity of a test material; use of a positive control such as soluble lead acetate in a lead bioavailability experiment prevents making erroneous conclusions about apparent inactivity of a test soil .

A positive control experiment is also used to obtain relative bioavailability. US-EPA defines the relative bioavailability as: "The ratio of the bioavailability of a metal in one exposure context (i.e., physical chemical matrix or physical chemical form of the metal) to that in another exposure context" (USEPA 2007).[1] US-EPA also assumes that lead is absorbed equally from water and food lead, by children and adults, with 50% absorption in children and 20% absorption in adults. It is accepted, and many studies show, however, that absorption is reduced considerably when food is present for lead already in the food (in adults, Rabinowitz 1980),[2] and also added to the food (in juvenile swine, USEPA 2007).[3]

---

[1] U.S.-EPA, 2007. Guidance for Evaluating the Oral Bioavailability of Metals in Soils for Use in Human Health Risk Assessment, OSWER 9285.7-80.

[2] Michael B. Rabinowitz, Joel D. Kopple, George W. Wetherill, 1980. Effect of food intake and fasting on gastrointestinal lead absorption in humans, The American Journal of Clinical Nutrition 33, 1784-1788.

In swine studies, the absorption of lead acetate in the presence of food is reduced by approximately ½ (Table 2, copied from USEPA 2007b).[3]

Table 2. Reduction in soluble lead absorption in juvenile swine under fed conditions. Table is copied from USEPA 2007b.

Measurement Endpoint	Ratio of PbAc Absorption Given With Food Compared to PbAc Given Without Food	Ratio of PbAc Absorption Given 2 Hours After Feeding Compared to PbAc Given Without Food
Blood Lead AUC	0.39 ± 0.05	0.40 ± 0.06
Liver Lead	0.86 ± 0.24	0.58 ± 0.16
Kidney Lead	0.72 ± 0.26	0.73 ± 0.27
Bone Lead	0.35 ± 0.05	0.33 ± 0.05
Point Estimate	0.58 ± 0.28	0.51 ± 0.22

Bioavailability studies typically use semi-fasted conditions for both the substance being tested and the positive control. Semi-fasted means that animals are given the dose on an empty stomach (at least 4 hours after a meal), in a small amount of feed formed into a "dough ball" (around 5 g, which is around 1% of their daily ration of food). Delivering the dose in feed (where swine are given the dose *ad libitum*) is a less accurate way to assess effects of any kind since the measuring of feed consumption adds to the experimental error. Furthermore, based on empirical evidence, we know that delivering a dose under fed-conditions greatly diminishes the bioavailability of many divalent cations (as discussed above for lead). This may be less critical when dosing with fat soluble compounds in a fat-rich diet. Nevertheless, dosing animals in a semi-fasted or fasted state provides a more accurate way to dose and reduces both experimental and biological variability impact on RfD or bioavailability calculations for many metals, metalloids, and some other compounds.

Topic 5: Calculation of Relative Bioavailability

Several sources of information or guidance are available for the calculation of relative bioavailability, but in short it is calculated most commonly by comparing single doses (soil and control) or by comparing slopes (obtained from different doses of soil and control, as discussed in Topic 2). The multiple dose method is preferred. This method, as mentioned above, allows estimates of bioavailability to be made over a range of doses, and it also means that not all the contaminant needs to be recovered from the animal after dosing (mass balance). An example follows for arsenic:

[3] U.S.-EPA, 2007. Estimation Of Relative Bioavailability Of Lead In Soil And Soil-Like Materials Using In Vivo And In Vitro Methods, OSWER 9285.7-77.

In most animals (including humans), absorbed arsenic is excreted mainly in the urine over the course of a few days. Thus, the urinary excretion fraction (UEF), defined as the amount excreted in the urine divided by the amount given, is usually a reasonable approximation of the AF_o or absolute bioavailability. However, this ratio will underestimate total absorption, because some absorbed arsenic is excreted in the feces, via the bile, and some absorbed arsenic enters tissue compartments (e.g., skin, hair) from which it is cleared very slowly or not at all. Thus, the urinary excretion fraction should not be equated with the absolute absorption fraction.

The RBA of two orally administered materials (i.e., a test material and reference material) can be calculated from the ratio of the urinary excretion fraction of the two materials. This calculation is independent of the extent of tissue binding and of biliary excretion:

$$RBA(test\ vs\ ref) = \frac{AF_o(test)}{AF_o(ref)} = \frac{D \cdot AF_o(test) \cdot K_u}{D \cdot AF_o(ref) \cdot K_u} = \frac{UEF(test)}{UEF(ref)}$$

where:

D = Ingested dose (µg)

K_u = Fraction of absorbed arsenic that is excreted in the urine

Based on the conceptual model above, the basic method used to estimate the RBA of arsenic in a particular test material compared to arsenic in a reference material (sodium arsenate in water) is as follows:

Plot the amount of arsenic excreted in the urine (µg/day) as a function of the administered amount of arsenic (µg/day), both for reference material (sodium arsenate) and for test material. Find the best fit linear regression line through each data set. The slope of each line (µg/day excreted per µg/day ingested) is the best estimate of the urinary excretion fraction (UEF) for each material. Calculate RBA for each test material as the ratio of the UEF for test material compared to UEF for reference material:

$$RBA(test\ vs\ ref) = \frac{UEF(test)}{UEF(ref)}$$

A detailed description of the curve-fitting methods and rationale, along with the methods used to quantify uncertainty in the arsenic RBA estimates for a test material, are summarized in Topic 6.

Topic 6: Replication and Uncertainty

Based on the variance calculated from previous experiments it has been established that 4-5 animals/dose group is sufficient. This allows the differences in response between groups to be distinguished, and reasonable confidence intervals around RBA estimates to be established.

Arsenic bioavailability estimates are subject to uncertainty that arises from several different sources. One source of error is the inherent biological variability between different animals in a dose group,

which in turn causes variability in the amount of arsenic absorbed by the exposed animals. This between-animal variability in response gives statistical uncertainty in the best-fit dose-response curves and, hence, uncertainty in the calculated values of RBA. Such statistical uncertainty is accounted for by the statistical models used and is characterized by the uncertainty range around the endpoint-specific and the point estimate values of RBA.

There is also uncertainty in the extrapolation of RBA values measured in juvenile swine to young children or adults, and this uncertainty is not included in the statistical confidence bounds used in the data reduction. Even though the immature swine is believed to be a useful and meaningful animal model for gastrointestinal absorption in children, it is possible that there are differences in physiological parameters that may influence RBA and that RBA values in swine are not identical to values in children. In addition, RBA may depend on the amount and type of food in the stomach, since the presence of food can influence stomach pH, emptying time, and possibly other factors that may influence solubilization of arsenic. In this regard, it is important to recall that RBA values measured in many studies are based on animals that have little or no food in their stomach at the time of exposure and, hence, are likely to yield better estimates of RBA with less variability caused by dietary influences. Thus, these RBA values may be somewhat conservative for humans who ingest the site soils along with food. The magnitude of this bias is not known.

The other source of error is the experimental error. This is minimized by ensuring SOPs for all procedures and methods exist and are strictly followed. Typically the uncertainty associated with analytical measurement error is not reported as a separate uncertainty estimate.

Uncertainty is usually calculated from the best-fit dose-response curves; carrying out the uncertainty analysis in this way assumes that the dose-response points capture all the sources of uncertainty in the experimental procedure (e.g., sample heterogeneity, subsampling uncertainty, measurement uncertainty, between animal variation etc.). Some of these other uncertainties may be negligible.

The following outlines a method that has been used for the estimation of uncertainty[4].

As previously mentioned, RBA values are calculated as the ratio of the slope term for the test material data set (b_t) and the reference material data set (b_r):

$$RBA = \frac{b_t}{b_r}$$

The uncertainly range about the RBA ratio is calculated using Fieller's Theorem as described by Finney (1978). As noted by Finney (1978), when the data to be analyzed consist of two dose-response curves (the reference material and the test material), it is obvious that both curves must have the same intercept, since there is no difference between the curves when the dose is zero. This requirement is

[4] E.g., in Stan W. Casteel, Christopher P. Weis, Gerry M. Henningsen, and William J. Brattin, 2006. Estimation of Relative Bioavailability of Lead in Soil and Soil-Like Materials Using Young Swine, Environmental Health Perspectives, 114(8), 1162-1171.

achieved by combining the two dose response equations into one and solving for the parameters simultaneously, as follows:

Separate Models:

$$\mu_r(i) = a + b_r \cdot x_r(i)$$

$$\mu_t(i) = a + b_t \cdot x_t(i)$$

Combined Model

$$\mu(i) = a + b_r \cdot x_r(i) + b_t \cdot x_t(i)$$

where *μ(i)* indicates the expected mean response of animals exposed at dose *x(i)*, and the subscripts *r* and *t* refer to reference and test material, respectively. The coefficients of this combined model are derived using multivariate regression, with the understanding that the combined data set is restricted to cases in which one (or both) of x_r and x_t are zero.

Regression analysis based on ordinary least squares assumes that the variance of the responses is independent of the dose and/or the response. It has previously been shown that this assumption is generally not satisfied in swine-based RBA studies, where there is a tendency toward increasing variance in response as a function of increasing dose (heteroscedasticity). One method for dealing with heteroscedasticity is through the use of weighted least squares regression. In this approach, each observation in a group of animals is assigned a weight that is inversely proportional to the variance of the response in that group:

$$w_i = \frac{1}{\sigma_i^2}$$

where:

w_i = weight assigned to all data points in dose group *i*

σ_i^2 = variance of responses in animals in dose group *i*

When the distributions of responses at each dose level are normal, weighted regression is equivalent to the maximum likelihood method.

There are several alternative strategies for assigning weights. One strategy estimates the value of σ_i^2 using an “external” variance model based on an analysis of the relationship between variance and mean response using data consolidated across many different swine-based arsenic RBA studies. Log-variance increases as an approximately linear function of log-mean response:

$$\ln(s_i^2) = k1 + k2 \cdot \ln(\bar{y}_i)$$

where:

s_i^2 = observed variance of responses of animals in dose group *i*

$\bar{y}_i$ = mean observed response of animals in dose group *i*

Based on these data, values of *k*1 and *k*2 are derived using ordinary least squares minimization. In an arsenic experiment, the resulting values were -1.10 for *k*1 and 1.64 for *k*2.

The goodness-of-fit of each dose-response model is assessed using the F test statistic and the adjusted coefficient of multiple determination (Adj R^2) as described by Draper and Smith (1998). A fit is considered acceptable if the p-value is less than 0.05.