National Institute
for Public Health
and the Environment

Letter report 711701094/2009

M.I. Bakker | W.I. Hagens

Proceedings of the workshop *In vitro* modeling of humane bioavailability of lead from soils

Application to risk assessment of soil

RIVM Letter report 711701094/2009
SIR report 12197

Proceedings of the workshop '*In vitro* modeling of humane bioavailability of lead from soils'

M.I. Bakker
W.I. Hagens

Center for Substances and Integrated Risk Assessment (SIR)

Contact:
Martine.Bakker@rivm.nl

This investigation has been performed by order and for the account of the Ministry of Housing, Spatial Planning and the Environment (VROM), DGM-DP and DGR-LOK, within the framework of project M/711701/BE, Risico's in relatie tot bodemkwaliteit/Implementatie biobeschikbaarheid mens.

Abstract

In vitro digestion methods may be used to estimate possible health risks of lead in soil to children. In a workshop on *in vitro* methods for oral bioavailability of lead from soils two methods were discussed: the tiny-TIM method of TNO and the *In Vitro* Digestion (IVD) method of the RIVM. It was concluded that the different results of the two methods were mainly caused by the pH of the gastric phase and the different separation techniques (ultra-filtration vs. centrifugation). I*n vivo* validation is required for a responsible application of the *in vitro* models for risk assessment of lead in soils. This should be done with samples of made grounds, preferably using young swine. The model showing the highest correlation with the *in vivo* data (for a large bioavailability range) will be considered the best applicable model. In addition, the selected model should be: 1) simple (feasible to operate in routine application at more than one location), 2) responsive to different lead and soil characteristics,
3) accompanied by rigorous Quality Assurance/Quality Control data requirements (e.g. with regard to recoveries, blanks, reproducibility). For both methods, adaptations to the current methods will be needed, e.g. with respect to the amount of soil used. For the IVD-method simplification of the model (to only the gastric phase) should be considered.

Five international, independent, experts were invited to give input to this workshop. The workshop was an initiative of TNO-Quality of life and RIVM as a follow up of the review of RIVM-report 'Oral bioavailability of lead in Dutch made grounds (Hagens et al. 2009).

Samenvatting

Voor het inschatten van gezondheidsrisico's van lood uit bodem voor kinderen kan de beschikbaarheid van lood geschat worden met laboratoriummodellen. In een workshop over *in vitro* methoden voor orale biobeschikbaarheid van lood uit bodem werden twee modellen besproken: de tiny-TIM methode van TNO en de *In Vitro* Digestie (IVD) methode van het RIVM. Geconcludeerd werd dat de belangrijkste oorzaken van de verschillende resultaten van de twee methoden waarschijnlijk de pH van de maagfase en de verschillende scheidingsmethoden (ultrafiltratie vs. centrifugatie) zijn. Een *in vivo* validatie van de modellen is nodig om ze in de toekomst te kunnen gebruiken bij risicobeoordeling van lood in bodem. Dit zou moeten plaatsvinden met behulp van monsters van stedelijke ophooglagen, bij voorkeur met biggen. Het model waarvan de resultaten het beste correleren met de *in vivo* data (voor een grote variatie aan waarden voor biobeschikbaarheid) is het meest geschikt voor deze toepassing. Daarnaast moet het model 1) simpel zijn (te gebruiken op meer dan één locatie), gevoelig zijn voor verschillen in lood- en bodemkarakteristieken, 3) voldoen aan strikte kwaliteitseisen (bijvoorbeeld ten aanzien van recoveries, blanco's, reproduceerbaarheid). Voor beide modellen zijn nog aanpassingen in de methodiek nodig, bijvoorbeeld ten aanzien van de hoeveelheid toegevoegd bodem. Voor het IVD-model zou een versimpeling (alleen de maagfase) overwogen moeten worden.

Vijf internationale, onafhankelijke, experts waren uitgenodigd om input te leveren voor deze workshop. De workshop was georganiseerd door TNO-Kwaliteit van leven en het RIVM naar aanleiding van een review van het RIVM-rapport 'Orale biobeschikbaarheid van lood in stedelijke ophooglagen' (Hagens et al. 2009).

Table of contents

1. Introduction

The bioavailability of lead from soils to humans needs to be taken into account to accurately and reliably assess the health risk of soils polluted with lead to humans. Therefore it is important to be able to predict the bioavailability of this metal from soil-matrices in the human gastro-intestinal system. Validated *in vitro* digestion models can form a first tier in assessing oral bioavailability of a compound from a specific matrix.
In The Netherlands two *in vitro* models are in use that can simulate the digestion of substances, namely the 'tiny' TNO *In vitro* Model (TIM) and the *In Vitro* digestion (IVD) method of RIVM. Both methods can be applied to determine the humane bioavailability of lead in soil matrices. To investigate if both methods would yield comparable results, the bioaccessibility of lead in sixteen soil samples from Dutch made grounds was determined with the two methods. The results of the two models appeared to be very different: the bioaccessibility of lead according to the tiny TIM model was a factor 4-5 lower than that of the IVD model.

Five international experts have reviewed this investigation of the two methods. In addition they have given their opinion on which of the models could be used in the practice of risk assessment. The international experts were:

1. Dr. Pat Rasmussen, Health Canada
2. Dr. Michael Beringer, US-EPA
3. Dr. Karen Bradham, US-EPA
4. Dr. Sohel Saikat, UK Health Protection Agency
5. Dr. Stan Casteel, University of Missouri

The external reviewers studied the report 'Relative oral bioavailability of lead from Dutch made grounds' (Hagens et al. 2009). They gave (prior to the workshop) general comments on the report and answered the following questions:

1. Can you indicate the strong and weak aspects of both *in vitro* digestion models? How can weak aspects be improved?
2. Do you have ideas about the main reasons for the large differences in bioaccessibility obtained with both models (see discussion in Chapter 9)?
3. How well have both models been compared to human and animal bioavailability data, see Chapter 9 and 10, and especially Table 9.3? Can you indicate the weak points in the comparison to in vivo data that has been performed for each model? What do you think of the manner of comparison for each digestion method?
4. Is it possible to indicate, based on the present information, which model you expect to give the most realistic estimate of relative bioavailability? This includes both the determination of the bioaccessibility, and the translation of bioaccessibility to a relative bioavailability factor for lead in soil, see Chapter 5, 8 and 9. Which model is that and why do you think so?
5. What is your recommendation in order to be able to appoint or develop an *in vitro* digestion model that can be safely applied in human health risk assessment of contaminated soils? Do you think one of the presently discussed models can be used for this. A balance is sought between the additional experimental research, the reliability of the method, the costs, and the time frame before implementation.

The present report summarizes the written input of the external reviewers and the discussions and conclusions from the workshop that took place on May 14th 2009, at the TNO institute in Zeist, organized by TNO Quality of Life and RIVM (See Appendix 1 for the workshop

program). In Chapter 2 the input of the external reviewers into the workshop is described. Firstly, a summary of the external reviews is given. Subsequently, the presentations of the external reviewers during the morning session of the workshop are summarized. Chapter 3 includes the discussions at the workshop. Firstly, the discussion on the summary of the reviewer's comments during the morning session is presented. Secondly, the two round table discussions on how the two models should be adapted and validated in order to use them for regulatory purposes are described. Lastly, the final discussion using some propositions is reported. The conclusions drawn during the workshop are listed in Chapter 4.

2. Input of the reviewers

2.1. Summary of the external reviews

The present section gives a summary of the external reviews that were received prior to the workshop. The full reviews can be found in Appendix 2.

General remarks
The calculation of the relative bioavailability (to the bioavailability of lead from food) is considered as relatively complex. Furthermore, the current calculations make direct comparisons of the two methods difficult.

The report presents the argument that use of the lower relative bioavailability factor yielded by Tiny-Tim would reduce costs of soil management decisions. However, to estimate the cost-benefit associated with each approach, the economic costs of potential human health impacts should also be estimated (ie. the cost to the Netherlands of chronic childhood exposure to lead at whatever target concentration is associated with each method).

Based on the large number of soils containing lead exceeding the established intervention value, remediation via a soil amendment might be a more economical way to lower the bioavailability. Phosphate is successfully applied to soils thereby reducing bioavailability by 30-40%.

Question 1 and 3
A summary of the answers of the reviewers to the first and third question 'Strong and weak aspects of both models' and 'How well have models been compared to *in vivo* data' is given in Table 1. The aspects of the IVD-method are given in the left and those of the Tiny TIM method in the right column of the table. The remarks in the same row are connected.

Question 2: Reasons for the large differences
The large differences between the results of the two models on the bioavailability of lead in made grounds may be primarily explained by differences in the separation methods and the conditions in the gastric phase. Several reviewers have commented on the considered difference in sample preparation (different particle size distribution). However, the sample preparation for the two methods has been the same. Apparently, this was not completely clear in the report.

Ultrafiltration versus centrifugation.
IVD uses centrifugation and tiny TIM uses ultra-filtration as a separation method. The reviewers agree that this difference may be an important cause for the different results. Filtering the samples (as in tiny TIM) removes the large soil particles and lead complexes with proteins ≥10 kDa. However, the reversible binding of lead to proteins will not limit bioavailability of lead that is a substrate for carrier-mediated absorptive transport in the upper small intestine.
One reviewer notes that, if lead concentration in the < 10 kDa fraction is highly correlated with the bioaccessible fraction of the lead concentration in the ≥10 kDa fraction, then the Tiny-TIM assay may provide good predictions of *in vivo* RBA, even though it is actually measuring only a portion of the bioaccessible fraction.

Table 1. Summary of reviewer's responses to Question 1 and 3

IVD	Tiny Tim
• Strong: Relatively cheap and easy to perform • Note: Maybe simpler by reducing the second phase	• Weak: Overly complex and time consuming
• Strong: (Preliminary) in vivo validation present. • Weak: Overall fit is not very robust (only relatively low and high in vivo bioavailability have been evaluated)	• Weak: No useful in vivo validation for lead (Rats are considered a poor model. With respect to human data point: children absorb lead to a much greater extent than adults).
• Weak: Very small aliquots of soil, 0.06 g, may be not representative for the entire soil. Small sample size also contributes to measurement error	• Strong: use of a larger quantity of soil which may lead to better reproducible results • Weak: Soil to liquid ratio is high; may not allow adequate physical contact
• Weak: pH sensitive (rejection of samples with 0.6 of soil). Not robust enough	• Weak: Early exclusion of gastric residue (low pH for a short time): slow release of Pb from soil is not included in the system
• Weak: compromised aspects of physiolocial relevance • Strong: physiologically relevant pH, consistency in temperature and transit times	• Strong: Close match of human gastrointestinal tract
	• Strong: replicates effect of eating a meal on absorption as compared to fasting condition • Weak: calculations needed to express values on fasted basis
• Strong: Attention has been given to QA (duplicates, procedural blanks, measurement of certified reference materials) • Weak: Results for reference soils containing bioavailability data are lacking	• Weak: Results for reference soils containing bioavailability data are lacking
• Weakness of both methods: difficulties with obtaining a reliable total Pb value of soils. Extraction with the more aggressive hydrogen fluoride (HF) is recommended	
• Strong: Measured bioaccessibility correlates with lead characteristics	• Weak: No correlation between measured bioaccessibility and lead characteristics of the soil

Conditions in gastric phase

The following differences in the gastric phase may be important causes for the lower bioaccessibility obtained with Tiny-TIM compared to IVD:

- the early exclusion of residual particles, larger molecules and complexes in Tiny-TIM.
- the shorter incubation time in the gastric compartment of Tiny-TIM, which does not allow time for Pb to be released to solutions from slow-release minerals, anthropogenic compounds, and complexes.
- higher pH conditions for a greater proportion of experimental run time.
- limited contact between particles and gastric fluid in Tiny-TIM, caused by the different soil:liquid ratio.

- the addition of a greater variety of ingredients in Tiny-TIM that may act as complexing agents resulting in losses of Pb from solution at all stages.
- the absence of simulation of a completely fasted state for Tiny-TIM.

Question 4
On Question 4 'Which model is expected to give the most realistic estimate of relative bioavailability' the reviewers agree that it is not possible to determine which model gives the most realistic estimates, since more data are needed on *in vivo- in vitro* validation on both models. A number of reviewers mention that the IVD-method seems more believable, since IVD is more sensitive to mineralogy and particle size than the Tiny TIM method.
In addition, several reviewers mention that the selection of which in vitro model provides the best estimate of the human/animal condition should be made on how well it correlates with the *in vivo* data.

Question 5
With respect to the last question 'What is needed in order to appoint or develop a model that can be safely applied in human health risk assessment of contaminated soils?' the following can be concluded from the reviews:

- *In vivo* validation of the methods is needed with relevant samples (i.e. samples of Dutch made grounds):
  - Validation studies with humans are recommended if possible, however there are ethical difficulties, especially since information on children is required.
  - Instead, in vivo studies with juvenile swine are recommended
- The model should be simple (simplifying the model to one or two phases would yield more reliable results and should be considered)
- The model should be responsive to proven differences in the biological availability of Pb in samples from different made grounds
- More quantitative characterization of representative samples of the Dutch made grounds (better information on particle size distribution, further identification of Pb-bearing compounds) would be a worthwhile investment to improve on the PPS ranking.
- Before acceptance as a regulatory protocol the in vitro methods should be run in triplicate for each sample and should be accompanied by rigorous Quality Assurance/Quality Control data requirements (recovery measured by more Certified Reference Materials, procedural blanks, control of pH conditions, high inter- and intralaboratory reproducibility).

2.2. Summary of reviewer's presentations

2.2.1 Sohel Saikat: Answer to Question 1 and UK perspective

Sohel Saikat presented his answer to Question 1(strong and weak aspects of both models):
IVD:

- *Strong aspects*: Simple, high throughput, economical, appears sensitive to variation in chemical and soil specific properties
- *Weak aspects*: Small sample size of the soil, pH maintenance

TIM:

- *Strong aspects*: Large sample size of the soil, can simulate effect of eating meal in absorption

- *Weak aspects*: Complex, time consuming, appears not sensitive to variation in chemical and soil specific properties

Sohel stressed on three basic questions in the development of in-vitro models for use in human risk assessment:
- Do we want a model to be physiologically close to the human gut situation? Or:
- Do we want a model capable of correctly predicting bioavailability of chemicals from soil (may be) at the compromise of physiological relevance? Or:
- Do we want in-vitro model capable of conservative prediction of bioavailability for soil borne chemicals (always)?

Sohel also summarized the UK perspective on the use of *in vitro* bioaccessibility data in land contamination in relation to the risk assessment framework. Risk assessment is undertaken as a tiered approach going from screening to generic quantitative risk assessment to detailed quantitative risk assessment using site specific data. In site specific risk assessment, information like bioaccessibility/bioavailability data produced by an appropriate method could help refine the risk assessment. The current view of the Environmental Agency on *in vitro* methods for the prediction of bioavailability is that its application is limited at this time, given the uncertainties present. There is an expectation that this view may be reviewed in line with ongoing changes in policy and guidance in land contamination.

2.2.2 Mike Beringer: Development and validation of an *in vitro* bioaccessibility method for lead in the U.S.

Mike Beringer started his presentation with a summary of the lead contaminated sites in Missouri due to land mining and the importance of bioavailability in evaluating these sites. He then sketched the history of the development of the juvenile swine method and the in vitro bioaccessibility method used at the US-EPA. The latter is a physiologically based extraction method (stomach phase only). A key component of method validation was conducting side by side comparisons between juvenile swine studies and *in vitro* bioaccessibility experiments. Temperature, pH, run time and physiological solutions of the experimental method were evaluated as part of a sensitivity analysis and also selected, in part, to improve the correlation with the juvenile swine results. Nineteen test materials were evaluated for lead.

The method in short: 1 g of <250 µm substrate is added to 100 ml extraction fluid (pH 1.5) and mixed for 1 hr via end-over-end agitation. The extract is filtered (0.45 µm), the pH is checked to ensure it is within 0.5 pH units of the starting pH, then the extract is analysed for lead.

An important step in validating the relative bioaccessibility leaching procedure was establishing method validation criteria and regulatory acceptance criteria. These criteria were adopted from the Interagency Coordinating Committee for Validation of Alternative Methods (ICCVAM). A good correlation was obtained between the *in vivo* relative (to lead acetate) bioavailability of lead in soils in the swine and the *in vitro* bioaccessibility results ($n = 19$, $r^2 = 0.92$). The soils tested consisted primarily of mining and smelter slags. The results also show that lead in soil samples from smelter deposition is sometimes more bioavailable than lead acetate.

The method as described above is presently used in site-specific risk assessment in the US. The US-EPA *in vitro* model was optimized, in part, to improve the correlation with the *in vivo* swine results using soils mainly from lead mining and milling sites. With respect to soil and lead characteristics, these soils are different than the Dutch soils. Although US-EPA does consider this method validated for a wide range of soil types; there is uncertainty when using

this method for other soils. Therefore, the Dutch soils may not follow the observed correlation and there is uncertainty associated with using the US-EPA *in vitro* method for Dutch made grounds.

2.2.3 Karen Bradham: *In vitro* bioaccessibility method for lead in the US & Question 3

Karen Bradham started her presentation with some general study design parameters for the development of in vitro methods. The 19 soils evaluated for lead (consisting primarily of mining and smelter slags) varied over a large range of bioaccessibilities: from ca. 0 to 90 %. In addition, she showed some experimental results which were evaluated as part of a sensitivity analysis for comparison with the juvenile swine method. Also, the quality controls for the method (blanks, recovery from spikes, duplicate samples, control soils, and reproducibility across labs) was addressed. This study resulted in vitro results that were directly compared with the in vivo results and resulted in a good correlation between the *in vivo* relative bioavailability of lead in soils in the swine model and the *in vitro* bioaccessibility results[1].
How much soil is needed for a measurement depends on the concentration in soil (about 1 kg).

In addition, Karen presented her and Mike Beringer's response to Question 3: How well have both models been compared to human and animal data?

An important note is that the calculations utilized for the IVD and tiny TIM methods make direct comparison of these two methods difficult. This also creates some difficulty in directly comparing each of the in vitro method results with the in vivo results. In general, it is difficult to compare the methods with the bioavailability data due to limited number of soils available for comparison.

Weaknesses of Tiny TIM with respect to validation:

- Lack of comparison of data to juvenile swine bioavailability results
- Tiny-TIM results were compared with the one human study soil (although adults) and rat data (considered a poor model for measuring Pb bioavailability)
- Difficult to compare results with in vivo results due to a limited QA/QC data available in the report

Weaknesses of IVD with respect to validation:

- Figure 5.1 - comparison only included a small number of data points and most of these materials had relatively low and high in vivo bioavailability
- Small soil sample size (0.06 g) may not be representative for the soil under consideration
- The soils that were used for the comparison between in vitro bioaccessibility and in vivo bioavailability were obtained from the US-EPA, i.e. about ten of the nineteen soils that were used to validate the US-EPA bioaccessibility method were kindly donated to RIVM. This implies that the soils consisted of mining and smelting slags that are not relevant for the Dutch situation, certainly not for Dutch made grounds.

[1] The US-EPA evaluated temperature, pH, run time, and physiological solutions of the in vitro method as part of a sensitivity analysis and also to improve the correlation with the juvenile swine results. The final parameters for the in vitro method were selected for a variety of reasons, including a physiological basis.

The method evaluation for both methods is difficult, based on the current information since there is a lack of in vivo data for made grounds. In the development of an in vitro method for Dutch made grounds, the recommendation in Section 9.6 of the RIVM report to perform an *in vivo* study in swine for the made grounds so that a direct comparison of the *in vitro* and *in vivo* data could be determined for these materials may well be considered.

2.2.4 Pat Rasmussen: Research at Health Canada and Question 5

Pat Rasmussen from Health Canada explains that there is gradual acceptance (by some organizations) of site-specific adjustments for oral bioavailability of lead. This is because of the large variability in total lead concentrations and because more adequate information is needed for human health risk assessment.
The approaches for bioavailability at Health Canada are: animal models (swine for oral bioavailability), the toy Safety Protocol used for consumer products (modified for house dust and soil research in background urban settings) and comparisons with Drexler and Brattin (2007, a validation study) for contaminated sites.

Pat's answer to Question 5 (Recommendations for method development to make it applicable in human risk assessment) for IVD is as follows:

Test method:

- Consider simplifying the IVD method to gastric phase only, building on the experience of the US-EPA. The goal is to correlate a method with in vivo results. The variability in gastric-only methods is lower than in methods that use both gastric and intestinal phase.

- The hydrogen fluoride (HF) method helps to detect the total lead in soil by increasing digestion efficiency. Risk assessment in the Netherlands is conducted with Aqua Regia, not HF. If you start a new protocol, then it might be the best time to switch to HF. However, HF requires careful safety protocols and needs to be performed in a specialized lab.
  On the other hand, the aqua regia method has disadvantages as well. The conditions under which the extractions with aqua regia are performed, may vary according to the standards of NEN. As a consequence the analysis of total lead may vary, which is the denominator in the bioaccessibility equation and therefore it can strongly impact the calculation of the bioaccessibility.

- Establisch QA/QC to improve reproducibility. Goal is a simple, reproducible SOP (use diverse CRMs and representative test samples for made grounds). Then perform round robin testing to refine SOP.

Validation:

- For a representative set of made ground test samples: in vitro-in vivo comparison
- Replicate testing to quantify uncertainty

Additional characterization of made grounds:

- Chemical speciation: inorganic and organic Pb compounds. Techniques to identify fine-grained Pb-bearing compounds: (XANES, micro XRD, micro XRF, FTIR).
- Consider inhalation pathway and ingestion pathway
- Influence of size fraction on concentration and bioaccessibility. It is noted that the Dutch soils are very clay-rich. Lead will be in fine fraction. This should be tested
- Study particle size distribution of the different soils

2.2.5 Stan Casteel: Using the swine model in bioavailability assessment

Stan Casteel presented his juvenile swine model for measuring in vivo relative bioavailability of lead in soil (relative to lead acetate solution given i.v.). Reference material (lead acetate) or lead-contaminated soils were administered orally to juvenile swine twice a day for 15 days. Blood samples were collected from each animal at multiple times during the course of the study, and samples of liver, kidney, and bone were collected on the last day of dosing. All samples were analyzed for lead. The Relative BioAvailability (RBA, i.e. relative to lead acetate) of a test material was estimated by fitting mathematical models to the dose-response curves for each measurement endpoint and finding the ratio of doses that gave equal responses. The final RBA for a test material was the simple average of the four endpoint-specific RBA values. Results from 19 different test materials reveal that there is a wide range of RBA values across different exposure materials, ranging from 6% to 105%. This variability in RBA between different samples highlights the importance of reliable RBA data to help improve risk assessments for lead in soil. Although the RBA value for a sample depends on the relative amounts of the different chemical and physical forms of lead present, data are not yet adequate to allow reliable quantitative predictions of RBA from chemical speciation data alone.

The reproducibility of the study is very good: values of 73 and 75 % were measured intra-lab and 77 % vs. 74 % between labs. The swine are an out-bred commercial crossbred that mimics the human population as much as possible. Swine are dosed twice a day for 15 days to mimic children's exposure (multiple exposures). Feeding occurs two hours after dosing. Doses were up to 675 µg Pb/kg bw/d. If the concentration of lead is low or high, the soil mass per dose is adjusted accordingly.
Total recoveries of the lead fed to the swine are not determined (because of excretion via feces). Recoveries from spiked samples are used for the different organs. The bioavailability of lead from soil is determined by comparison of the concentration of lead in blood, liver, kidney and bone after dosing with soil versus the concentration after dosing with soluble lead acetate.
Note that the absolute bioavailability of lead acetate from an aqueous solution in swine is 35%. So, even if the relative bioavailability of lead in a soil is high, this still means that only an absolute small amount of lead is absorbed.
In the *in vivo* studies toxicity to the swine was not observed.

3 Discussions at the workshop

3.1 Discussion of the reviews

The second part of the morning program of the workshop was dedicated to the summary and discussion of the five reviews. After concluding that there was general consensus among the five reviewers, a number of subjects mentioned in Table 1 were discussed to address Question 1-3.

Question 1-3

Complexity
On the complexity of the methods there was general agreement with the comments in the first row of Table 1: IVD is cheap and simple, tiny TIM is (overly) complex and time-consuming.

In vivo validation:
Mans Minekus explained that tiny TIM aims to be a general method for the assessment of bioavailability in the gut (GI tract). It aims to be "substance-independent". The idea behind it is that by simulating the real conditions in the GI-tract as realistic as possible, it will correctly predict bioavailability for all substances. So, in theory, once you have validated the model for a limited number of substances, you can apply it to other substances (without validation). This might be an argument that is relevant for policy makers. After the lead bioavailability issue has been settled, other metals will follow (like in the USA). It could then be nice to have one method that can assess the bioavailability of all metals, preferably in one run.

Amount of soil
The 0.06 grams used in IVD was considered to be very small. However, using 0.6 g may result in a too high pH with the highly alkaline soils. It was generally recognized that sometimes smaller samples are necessary. There were concerns the small amount of 0.06 g might not be representative, and therefore more replicates would be needed if smaller samples are used

Gastric pH
In tiny TIM the pH decreases from 5 to 2 in 90 minutes, whereas in IVD it is 1.5 for 2 hours. The question regarding this difference asked by Johannes Lijzen was: What is a good representation: fed or fasted conditions? Mans Minekus answered that the stomach of small children is only completely empty for a short time, so often the pH is higher than 1-2.

Soil to liquid ratio
The soil to liquid ratio used in tiny TIM is relatively high. It was chosen because it mimics pica rather than hand-to mouth behavior.

Physiological relevance of method
There was consensus on the comment that there is no in vivo data to support the claim that TIM closely matches the gastro-intestinal tract regarding its behaviour with respect to lead.

Question 4

Everyone agreed that it is not possible to determine which model gives the most realistic estimates and that *in vivo* validation studies are needed.

Question 5

Mans Minekus did not agree with the comment on the slide: 'It was noted that the investigation of the differences between IVD and Tiny-TIM will not help to determine which model is better for lead risk assessment purposes'. According to his opinion the investigation will help understand the differences, but not resolve the decision on which model should be used. Sohel Saikat mentioned that we need agreement on what we want from in vitro methods i.e. physiological relevance as close as possible or ability to predict the bioavailability at the compromise of physiological relevance if necessary. Mans: We need to do it as simple as possible, with the parameters needed to predict the in vivo situation. Is a correlation as found in the EPA-studies enough to predict? Mike Beringer added that we cannot replicate absorption and active transport, therefore we use some model that can predict well enough.

3.2 Additional discussion points

Werner Hagens addressed a number of discussion points in his presentation. These were:

1. Main differences between Tiny Tim and IVD
2. PPS (Primary lead phases, Particle size and Secondary lead phases) analysis
2. *In vivo* validation of both methods
2. IVD: stomach phase only or intestinal phase
2. Particle size

Ad 1. Main differences between Tiny Tim and IVD
The differences between Tiny-TIM and IVD are summarized in Table 2. It was agreed by the participants that the ones most likely causing the difference in the results are gastric pH and separation method.
In IVD, the gastric pH is 1.5, which is maintained for 2 hours, whereas in Tiny Tim the pH decreases from 5 to 2 in the same time period. In the first 1.5 hours, 65 % of the material in the gastric phase is emptied in the intestinal phase. This means that less than 50 % of the stomach fraction is 'incubated' at a pH of 2. The reviewers commented that early exclusion of residual particles from the gastric phase, larger molecules and complexes does not allow the slow release of Pb from soil particles, including organo-Pb complexes, which occur in the stomach. These complexes are soluble in a child's stomach.

The ultrafiltration method used in Tiny TIM, leads to lower bioaccessibilities than the centrifugation as used in IVD. This may well be the most important reason for the large difference between results of the two methods.
The reviewers commented that with filtration, lead complexes > 10kD are excluded. This may be incorrect as Pb can desorb from these complexes and become available. In addition, lead reversibly bound to proteins and carrier mediated absorptive transport in the upper small intestine are excluded. Note that children absorb lead to a much greater extent than adults (possibly the part that is actively transported is larger for children because lead uses the calcium-transporters and children have more calcium-transporters because calcium is needed for growth).

Mans Minekus did not agree with the reviewers' point of view. Rob Havenaar added that it is not logical that particles of 0.45 µm would pass the intestine. Even 10 kD may be too large to be absorbed. Pat Rasmussen points at the idea that if large particles would be left in the stomach, they would release ions that would be absorbed later. Rob and Mans noted that only the water phase is removed/emptied from the gastric phase of tiny TIM. The heavier 'stuff' remains in the stomach. Johannes Lijzen added that it would be informative to have the retention time of the particles in the stomach (at present this is not measured).

The British Geological Survey also determined the bioaccessibility of lead in the Dutch made ground samples Werner showed how the IVD-results or the tiny TIM results correlated with the BGS-method.. The three methods appear not to correlate with each other.

Note that for a number of American soil samples the results of the IVD correlated strongly with the US-EPA model (in this comparison there were no soils with intermediate bioaccessibility included).

Table 2 Main causes of the different results of Tiny TIM and IVD for relative bioavailability of lead from made grounds.

Tiny Tim	**IVD**	**Important difference for:**
Dynamic model	Static model	pH range (gastric)[1]
5 gram of soil	0.06 gram of soil	• Accuracy • Representative for hand-to-mouth behavior vs. pica
Ultrafiltration	Centrifugation	Inclusion of lead complexes in the determination of the bioaccessibility
Fed condition without food[2]	Fasted and/or fed condition	• pH • Complexation of lead to food particles
Relatively complex	Relatively simple	Cost

[1] Gastric phase: IVD: pH 1.5 for 2 hours, Tiny TIM: pH decreases from 5 to 2 in two hours.

[2] The influence of the presence of food was tested in singular experiments with lead acetate and a meal. The presence of a meal decreased the bioaccessibility from ~60 % to ~10 %.

Ad 2. PPS analysis

In the PPS analysis a ranking of the soils is performed according to their lead characteristics. The type of primary lead, the particle size and the presence of secondary lead are the three important factors in this ranking procedure. Based on this ranking system, there was a correlation between the lead characteristics of the soils and the relative bioavailability as measured by the IVD-method, whereas this was not the case for the results of the Tiny TIM method. This was unexpected, since the same soils were used in Tiny TIM and IVD. The fact that Tiny TIM uses ultrafiltration as a separation method does not change this expectation.

The lack of correlation with lead characteristics was considered a weak aspect of the TIM-method by the reviewers. Nevertheless, although Pat Rasmussen, strongly supported this approach of ranking made grounds on the basis of Pb speciation (especially solubility and particle size), the choices made in the ranking system may be debatable to some extent. For this reason, the sensitivity of the ranking system was tested by changing the relative weight of the three factors. It appears that the positive correlation of IVD with the PPS-ranking is mainly due to the particle size of the soil samples (See Figure 1, lower left panel). The primary and secondary lead phases also show a positive correlation for IVD, although this is

mainly due to the presence of one soil, ranked as '3' (Figure 1, upper left panel). For tiny TIM, the particle size of the soils also gives a positive correlation (Figure 1, lower right panel), but the primary and secondary lead give a negative correlation, since the one soil ranked as '3' has a low bioavailability as measured in TIM (Figure 1, upper right panel).
It may be concluded that the information from the PPS-ranking is not very solid and additional information on the lead characteristics (e.g. micro X-ray diffraction or micro X-ray fluoresence analysis) of the made grounds is needed to be able to draw firm conclusions about this.

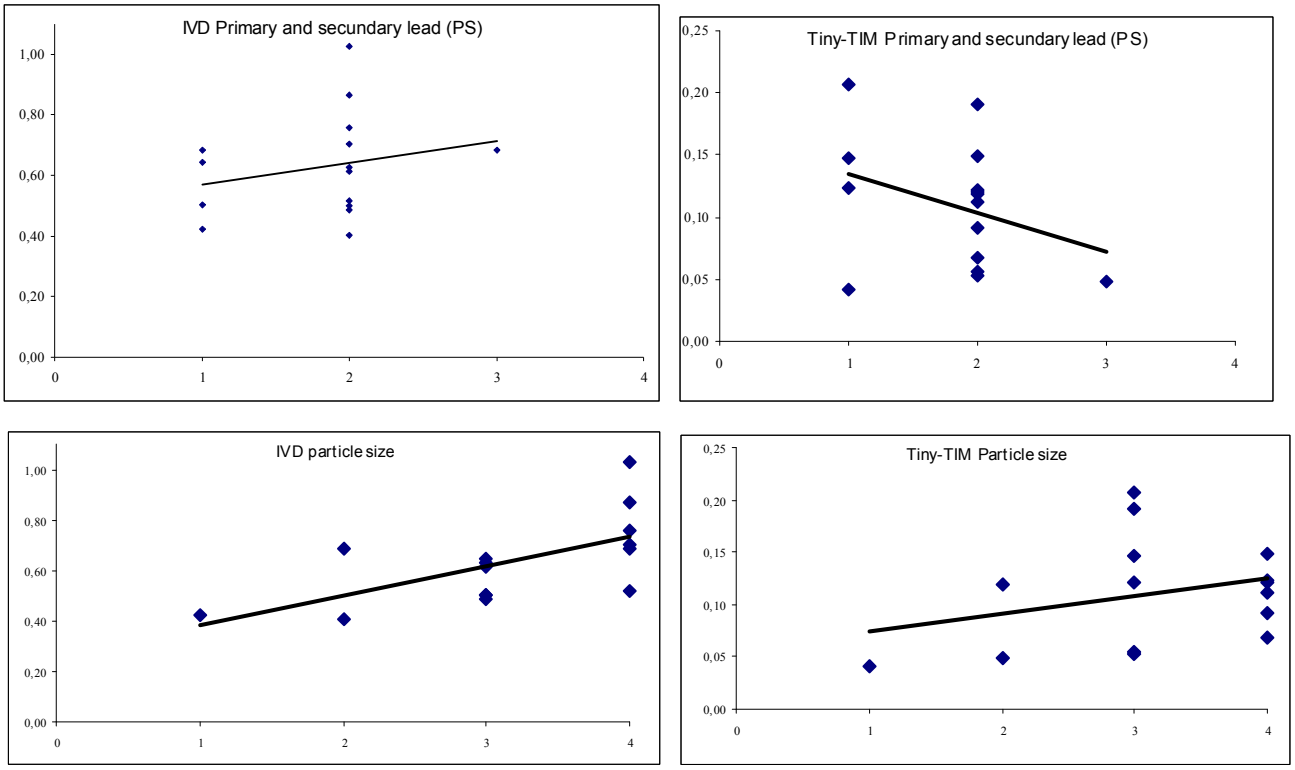


Figure 1. Relation of relative bioavailability of lead in made grounds measured with IVD (left panels) and tiny TIM (Right panels) with lead and soil characteristics. The relation with primary and secondary lead phases are shown in the upper panels, whereas the relation with particle size is shown in the lower panels.

Ad 3. In vivo validation
The reviewers agreed that although some useful *in vivo* validation has been carried out for IVD, this is not sufficient. The soil samples used for the validation have a low or high bioavailability, but there are no samples validated with an 'in between' bioavailability. Moreover, the tested samples are not representative for the Dutch made grounds.
For Tiny TIM, validation has been performed with a human study (one data point) and with a rat study. However, rats are not considered a good model for the bioavailability of lead in children.
In conclusion, more validation is needed, for both models.

Ad 4. IVD: stomach phase only or intestinal phase
The validation studies with IVD also comprised the validation of the results of the stomach phase only. The correlation results for the stomach phase were similar to those of the

intestinal phase. So, in principle, the IVD method could be simplified to the stomach phase. This would be a similar method as is used by the US-EPA.

Ad 5. Particle size
After sieving and removal of large parts, the made ground samples were "smashed" to 2 mm. This means that lumps/particles larger than 2 mm are broken down to 2 mm, and particles smaller than 2 mm are not changed. The < 2 mm fraction is standard practice in soil and agricultural science, whereas a fraction < 250 µm is normally used in human health studies. In general, the smaller particles stick to children's hands. These smaller particles are included in the present study. Nevertheless, to render comparable results, it may be considered if in the future the < 250 µm fraction should be used (also for the determination of total lead).

3.3 Round table discussions

For both models it was discussed which adaptations and additional research are needed to make the model feasible for use in risk assessment. After half an hour of discussion, people got the opportunity to switch discussion groups. In the sections below a short summary of the discussion is given.

3.3.1 Round table discussion IVD

Adaptations to IVD and additional research needed:

- There is confusion about the way bioavailability is expressed: relative bioavailability (relative to food or lead acetate) and absolute bioavailability. Everyone should be very clear about the way bioavailability is used in documentation. In addition, further discussion about the best practices of expressing bioavailability is recommended.

- The amount of 0.06 g soil may lead to irreproducible results due to inhomogeneous soils A reproducibility study is needed. Another possibility is increasing the amount of soil, but keeping the solid to liquid ratio the same. This may be done by increasing the flask volumes.

- Regarding the particle size distribution: It may be considered to use < 250 µm fraction instead of < 2 mm.

- More measurements with Certified Reference Materials are needed.

- To see the influence of the separation methods, different methods should be tried: Centrifugation, filtration with 0.45 µm and ultrafiltration

- There is a need to better understand the made ground: Additional lead characteristics are needed, especially the chemical speciation: inorganic and organic Pb compounds. Techniques that may be used to identify fine-grained Pb-bearing compounds: (XANES, micro XRD, micro XRF, FTIR). In addition the sequential extraction method of BGS may be used.

- Various aspects of both digestion models may be changed and compared to vivo bioavailability data; for the IVD model these are

    - Gastric-phase only? In general, a model should show good correlation with in vivo data; a model can be physiologically based, but is not necessary to achieve validation
    - Separation by ultrafiltration
    - Comparison with results of US-EPA in vitro model

- How conservative should the model be? It should be sensitive enough to reflect difference in mineralogy.
- The normal variation in physiology (fed, fasted etc) results in a considerable variation in bioaccessibility as pH is a determining factor for the bioaccessibility of lead. It should be very clear how conservative the in vitro model should be for risk assessment purposes and before further in vivo validation steps are performed.

- What are the criteria for in vitro methods? Validation is necessary; look for example at the criteria of Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM, http://iccvam.niehs.nih.gov/)
    - When a round robin is performed, be aware of the differences in apparatus (end-over-end shaker, centrifugation system). These differences may cause differences in results.

3.3.2 Round table discussion Tiny TIM

Adaptations to Tiny TIM and additional research needed:

Analyses, irrespective of tiny TIM/IVD

- A robust, standardized method is needed for total Pb content of soil
- Reference materials (soil) are needed for a whole range of bioaccessibilities to validate model

Validation

Swine studies are needed for validation of the in vitro models.
Conditions to simulate in the validation studies:

- Worst case (low pH)
- Children
- Fed/fasted or mixture

(In)sensitivity of tinyTIM to mineralogy

- Tiny TIM could be kinetically limited by liquid/solid ratio and therefore be insensitive to the mineralogy of the soil.
- It was recommended to test a range of different amounts of soil per reaction (0.05 gram versus 0.5 versus 5 gram).
- Unknown retention time in stomach and pH-range applied.
- Validation studies with pigs are recommended:
    - Use soils with low up to high bioavailability (whole range)
    - The soils should be relevant to Dutch soils
    - The samples of the made grounds may be to complex (too many different PPS factors?) You may choose a simple soil instead to start with.

Other newly proposed experiments

Test NL soil samples in pigs and in vitro and see if there is a correlation.
Test US soil samples in in vitro methods, compare with the outcome of the swine studies.
Use both of the above mentioned to study the bioaccessibility of lead from Dutch made grounds.

3.4 Final discussion

A final discussion was held regarding several propositions.

Proposition 1

- The costs of in vitro methods for measuring bioavailability are low compared to those of remediation

- To develop a regulatory framework for the bioavailability of substances in soil it is not necessary to have a simple, cheap model with high sample throughput

Discussion: It is important that the results of the model can be used by policy makers. A simple model makes it more acceptable since responsible parties do not want to spend a lot of money. In addition, a generic factor is helpful.
For a generic factor in the Netherlands, *in vivo* data of made grounds is needed
Mans Minekus: There are approximately 11,000 sites in the Netherlands that are contaminated. The soil is very heterogenic. You cannot always perform *in vivo* tests and a generic factor might not be valid for these sites. Therefore, get as much info as needed, also with *in vitro* (tiny TIM) studies: For 100,000 dollar, only one swine experiment[2] can be performed while more than 100 *in vitro* Tiny-TIM bioavailability data points can be gathered.

The general agreement was that in general, for regulatory framework, an expensive test can be run like *in vivo* studies, for a wide range of different sample types. For site specific studies, a cheap and simple method with a high throughput should be used.

Proposition 2

- Local authorities may decide for themselves which model they use to measure bioavailability: IVD, TIM or US-EPA model

Everybody disagreed (no discussion)

Proposition 3

- (A generic, conservative, default bioavailability of lead in soil (e.g. 0.6) should be used in the regulations instead of values generated with a model) was skipped, because the subject was already covered in Proposition 1.

Proposition 4

- The selection of the best in vitro model should be done on how it correlates with in vivo data

Everybody agreed (no discussion)

Proposition 5

- For both IVD and TIM more validation is needed

[2] $100,000 will pay for the testing of 4 to 6 soils depending on the experimental design and each of the 8-10 experimental replicates will generate 9-10 data points from which a measure of biological variability can be provided

- This should be done by measuring the bioavailability of lead in made grounds in young swine

Everybody agreed on more validation (no discussion)
Everybody agreed to go to in vivo studies with swine (no discussion)

Proposition 6

- IVD likely overestimates bioaccessibility
- TIM likely underestimates bioaccessibility
- But that is not a problem as long as it correlates with in vivo data

It is important that for a range of bioavailabilities the model correlates linearly. Otherwise, your model is a sort of black box. You need to know the soil/lead characteristics to also know the limitations of the model. There is a difference between the two models with respect to physiology. Although both models are physiology driven, in Tiny-Tim the physiology is even more important than in IVD. As long as you know what parameters are involved, you are not working with a black box. For lead, the chemistry in leaching (low pH over time) is important: PPS chemistry versus dynamic physiology based tiny-Tim.

Proposition 7

- IVD could be used but
  -a higher solid/liquid ratio should be applied
  -The pH should be better maintained

There was general agreement that more soil per experiment should be added. This will increase the reproducibility and takes the heterogeneity of the soil better into account.
The opinion of Tom van der Wiele is that also a different separation method is needed.
Therefore, IVD could be used, only with modifications.

Proposition 8

- TIM could be used but a lower solid/liquid ratio in the gastric phase is needed

There was agreement that longer 'exposure' to low pH in gastric phase is needed.
But how much soil should be tested? 1 gram? Less?
Also, additional studies with lead-acetate are needed to compare the results with the in vivo studies.

Proposition 9

- The appointed model should be evaluated for other substances than lead

There was general agreement that the model should be validated for every different substance that is to be tested

Proposition 10

- For risk assessment of metals in soil, other exposure routes (ingestion and inhalation of (house) dust) should be included

Inhalation could be included, however, for lead it is known that although inhaled, the lead will be transported from lungs to mouth. Also, the dose of lead absorbed via the oral route is typically much greater than via inhalation unless air concentrations are elevated. Therefore,, ingestion is more important for lead from soil.

There is also direct ingestion of house dust. Soil is a major constituent of house dust.

Proposition 11

- All the different ‘bioaccessibility terms’ (absolute, relative to PbAc, relative to food) make the discussion of this subject overly complex

- In the future we’ll use the term absolute bioaccessibility as much as possible

This was agreed upon by the participants.

4 Conclusions

To set up a regulatory framework, it may be worthwile to run expensive tests like *in vivo* studies. For site specific studies, a cheap and simple method with a high throughput should be used.

The selection of the best in vitro model should be predominantly based on how well its results correlate with *in vivo* data. For a large bioavailability range a robust relationship between in vitro and in vivo data should be demonstrated. Therefore, in vivo studies with juvenile swine with Dutch made grounds are recommended.

In addition, the selected model should

- be simple (feasible to be run at more than one location)
- be responsive to different lead and soil characteristics
- be accompanied by rigorous Quality Assurance/Quality Control data requirements

For both IVD and TIM methods more validation with *in vivo* data is needed. This should be done by measuring the bioavailability of lead in made grounds in young swine.
Information on the soil and lead characteristics in relation to their influence on bioavailability indicates the suitability of an in-vitro method for different soil types. In other words: The limitations of the in vitro model should be known. At present this is limited to the soil and lead characteristics the model was validated for.

For both models adjustments are needed to obtain a more reliable method. Attention should be paid to the reproducibility of the IVD method (e.g. more replicates or adding more soil but keeping the soil to liquid ratio the same). Simplifying the IVD method to only the gastric phase should be considered. In the tiny TIM method a lower solid/liquid ratio in the gastric phase should be applied (try 0.05, 0.1 and 5 g of soil to see influence of soil mass).

All the different ‘bioaccessibility terms’ (absolute, relative to lead acetate, relative to food) make the discussion of this subject overly complex. Therefore using the term absolute bioaccessibility where possible is preferable.

More quantitative characterization of representative samples of the Dutch made grounds (better information on particle size distribution, further identification of lead-bearing compounds) would be a worthwhile investment to improve on the PPS ranking.

The total extraction method for lead from soils should be evaluated and the use of hydrogen fluoride should be considered.

Appendix 1 Workshop program

Workshop *In vitro* modelling of humane bioavailability of lead from soils
Application to risk assessment of soil quality

On: May 14th, 9.00-17.00
At: TNO Quality of Life
Utrechtseweg 48
3704 HE Zeist
Chair: Ad Ragas, Radboud University Nijmegen

Morning Programme

9.00	Reception	
9.15	Welcome	
9.30	Sohel Saikat:	IVD and TNO models: strong and weak points & UK perspective on in vitro bioaccessibility in land contamination
9.45	Mike Beringer:	Development and Validation of an *In Vitro* Bioaccessibility Method for Lead in the U.S.
10.00	Karen Bradham:	In Vitro Bioaccessibility Method for Lead in the U.S. and Workshop Question 3
10.15	Coffee break	
10.35	Pat Rasmussen:	Recommendations for appointment/development Dutch models
10.50	Stan Casteel:	*In vivo* experiments with young swine
11.05	Summary reviews & discussion (Werner Hagens)	
11.50	Summary morning discussions (Ad Ragas)	
12.05	Lunch	

Afternoon programme

13.00 Welcome

13.15 Introduction
- In vitro models
  - Model TNO (Mans Minekus, TNO)
  - Model RIVM (Werner Hagens, RIVM),
- Bioavailability of lead in Dutch made grounds (ophooglagen) (Werner Hagens & Mans Minekus)

14.00 Summary external reviews (Werner Hagens)

14.20 Round table discussion groups
1. Study design to explain differences between results of the two models
2. Adaptation of RIVM model and additional research needed to make model feasible for use in risk assessment

15.20 Break

15.45 Feedback discussion groups

16.15 Final discussion (Martine Bakker, RIVM)

17.00 Closing

Drinks

Appendix 2. External reviews

1. Review Mike Beringer and Karen Bradham, US-EPA

General Comments (Mike Beringer)

1. Overall, the document and discussion is laid out in a logical format. In addition, the level of detail is appropriate to fully understand the issues.

2. It is apparent that the focus of this research was to study the relationship between bioaccessibility and soil characteristics. As a result, this discussion tends to overwhelm the discussion of the *in vitro* results.

3. In comparing *in vitro* bioaccessibility (IVBA) to *in vivo* relative bioavailability (RBA) results for lead, the U.S. EPA has taken a different approach than the Dutch National Institute of Public Health and the Environment (RIVM). Our approach has entailed comparing *in vivo* (RBA) results directly to IVBA results to determine whether IVBA can be used as a reliable predictor of RBA for human health risk assessment purpose. This evaluation has been used to derive an equation for estimating lead RBA in young children based on IVBA (U.S. EPA, 2007b). This approach is in contrast to that of RIVM which considers the bioaccessibility results comparable to absolute bioavailability. The bioaccessibility results are then converted to relative biovailability using Equations 12 and 13 on p. 91, which account for the bioaccessible fraction that is absorbed and the bioavailability of dietary lead. These equations complicate comparisons between IVBA and *in vivo* RBA because these parameters are associated with additional uncertainty that may obscure the relationship between IVBA and RBA. While this is not a criticism, the RIVM approach makes our review difficult because U.S. EPA uses a different metric when evaluating IVBA results.

Specific Questions (Mike Beringer and Karen Bradham)

1. Can you indicate the strong and weak aspects of both in vitro digestion models? How can weak aspects be improved?

U.S. EPA has evaluated the usefulness of *in vitro* bioaccessibility assays (IVBA) strictly against performance criteria: (1) Does the assay predict *in vivo* relative bioavailability (RBA, test material/soluble lead) with adequate confidence to be useful in risk assessment? (2) Is this confidence robust when the assay is implemented in multiple laboratories (i.e., inter-laboratory reproducibility)? These criteria have been assessed by evaluating empirically-based prediction models (e.g., regression) relating IVBA and RBA, and assessments of intra-laboratory and inter-laboratory reproducibility (U.S. EPA, 2007a, b).

Central to evaluating the first performance criterion has been the expectation that a useful assay would support the application of a function, *f,* for which:

$$in\ vivo\ RBA = f(IVBA)$$

and which yields estimates of RBA from IVBA that are sufficiently reliable to support risk assessment. U.S. EPA never adopted the expectation or requirement that IVBA must equal RBA, only that the relationship between the two (*f*) is known and robust. This approach appears to be fundamentally different from the approach taken in RIVM (2008). RIVM (2008) has evaluated the IVD assay against reported observations of RBA in swine and

concluded, based on the slope and R^2 of the linear regression model, that the bioaccessible fraction (F_B) is given by the results of the IVD assay (i.e., F_B = IVD). This leads to an apparent conflict between the IVD and Tiny-TIM assays, because Tiny-TIM results are less than IVD and, since IVD = F_B, Tiny-TIM must also be less than F_B. What appears to be missing is support for functions, f_1 and f_2, which convert IVD to F_B and Tiny-TIM to F_B:

$$F_B = f_1(IVD)$$

$$F_B = f_2(tiny - TIM)$$

where $f_1 \neq f_2$. From this perspective, in the absence of derivations of f_1 and f_2, we find it difficult to evaluate the strengths and weaknesses of each model, except in very general terms.

The IVD method appears to produce bioaccessibility results that match our general understanding of lead speciation data. In comparison to the Tiny-TIM model, this method is relatively easy to perform. However, we recommend evaluating whether three separate phases are absolutely necessary. Reducing the number of phases will increase intra- and inter-laboratory reproducibility, which based on the available data, is significantly greater than we see with U.S. EPA's validated *in vitro* method for lead. While Oomen et al. (2006) have determined the IVD method is "validated," the results in Figure 5.1 indicate the overall fit is not very robust. Furthermore, it appears that only materials with relatively low and high *in vivo* biovailability have been evaluated.

For the IVD method, the small aliquots of soil (0.06 g fasted and 0.04 g fed) may not be representative of the entire soil due to the heterogeneity of the soil. This small sample size also will contribute to measurement error and decreased reproducibility of bioaccessibility results, as discussed in Section 5.3. The report referred to the need to utilize 0.06 g of soil so there is no rejection of data caused by suboptimal pH incubation. Changes in pH values can easily be accounted for during an *in vitro* extraction by: 1) determining that the pH is out of acceptable range at the end of the extraction and re-extracting the sample, monitoring the pH in 15 minute intervals, and adjusting the pH, if necessary; or 2) monitoring the pH in 15 minute intervals, and adjusting the pH, if necessary.

An apparent strength of the Tiny-TIM method is that it most closely matches the human gastrointestinal tract. A larger soil quantity is used which is responsible, in part, for generating highly reproducible results. Another strength is that Tiny-TIM also replicates the effect of eating a meal on absorption in humans as compared to a fasted condition.

However, while the Tiny-TIM method is based on physiological conditions, the method appears overly complex and is time-consuming. Thus, it is not amenable to a high throughput of soil samples. Another weakness of this model is that calculations are necessary to express the values on a fasted basis. From the standpoint of a regulatory agency, erring on the side of ensuring protection of human health is most often desirable. Therefore, the fasted conditions may be considered most desirable. Also, centrifugation is not superior to filtering. Centrifugation does not remove larger soil particles, allowing the continuation of the extraction process.

The text in Section 8.3.2 states that the Tiny-TIM bioaccessibility result for NIST SRM 2710 is comparable to results obtained in rats. However, rats are considered a poor model for measuring lead bioavailability because of high rates of biliary excretion of lead (Weiss and LaVelle, 1991). The Tiny-TIM and IVD bioaccessibility values for the NIST SRM 2710 are also very different from a bioaccessibility of 75% reported by Drexler and Brattin (2007). We note that U.S. EPA has used this SRM as a control sample for quality control purposes (US EPA, 2007c).

In addition, the calculations for these methods appear overly complex for calculating the estimated relative bioavailability of lead. Furthermore, they make direct comparison of these methods difficult. Most calculations for *in vitro* bioaccessibility are based on the following equation:

% IVBA = (concentration in *in vitro*)(solution volume)/(concentration in solid)(sample weight)

Using this equation or a similar equation would provide a basis for direct comparison. Additional equations may then be employed for calculating potential exposures, correction factors, etc.

2. Do you have ideas about the main reasons for the large differences in bioaccessibility obtained with both models (see discussion in Chapter 9)?

Before discussing the potential reasons for the disparate bioaccessibility results, we note that Figure 5.1 of the RIVM report (RIVM, 2008) shows comparisons of IVD results to highly censored swine data. The censored regression slopes (intercepts forced to zero) were 0.69 (R^2 0.81, 0.6 g soil) and 1.16 (R^2 0.66, 0.06 g soil). However, an analogous comparison of Tiny-TIM assay results does not appear to have been reported. In the absence of this type of evaluation of the Tiny-TIM assay, it is not possible to derive a prediction algorithm for translating Tiny-TIM assay results into *in vivo* RBA. The observation that the IVD assay and Tiny-TIM assays yield dissimilar results when applied to the same test materials is not, by itself, sufficient for concluding which *in vitro* assay yields better predictions of *in vivo* RBA. For example, the Tiny-TIM assay could yield lower estimates of *bioaccessible* lead than the IVD assay, yet still have a high R^2 for predicting *in vivo* RBA.

There are a number of possible reasons for the large differences in bioaccessibility obtained with both models and Section 9.4 addresses the most likely sources. First of all, in the Tiny-TIM assay, efflux from the system is dialyzed against a 10 kDa exclusion membrane (dialysate) and bioaccessibility is calculated as the ratio of the amount of lead in dialysate to the total amount of lead in the efflux. Lead complexes with proteins ≥10 kDa would be excluded from the estimate of the bioaccessible fraction. However, reversible binding of lead to proteins will not limit bioavailability of lead that is a substrate for carrier-mediated absorptive transport in the upper small intestine. The use of a dialysate in the Tiny-TIM assay rather than an acid extraction (which would release lead associated with acidic moieties in protein) in the IVD assay may contribute to the differences in the estimates of bioaccessibility and relative bioavailability (*Rel F*) obtained from the two assays. Note, if lead concentration in the < 10 kDa fraction is highly correlated with the bioaccessible fraction of the lead concentration in the ≥10 kDa fraction, then the Tiny-TIM assay may provide good predictions of *in vivo* RBA, even though it is actually measuring only a portion of the bioaccessible fraction. Based on percentiles for relative bioavailability estimated from the two assays, reported in Table 9.1, $F_{c,Tiny\text{-}TIM} / F_{c,IVD} \approx 0.2$ would yield similar estimates of *Rel F*.

A second possible explanation for the differences between the methods may be due to the sample preparation (see Section 2.3.1). According to the information on page 21, Part A used for the IVD extraction method, was ground to < 500 µm, while Part C that was sent to TNO for the Tiny-TIM extraction method was only sieved and milled to the < 2 mm size fraction. Fracturing or grinding soil samples decreases the particle size, resulting in increased bioavailability/bioaccessibility of metals. Lead extracted by the IVD method in the < 500 µm fraction was significantly higher than Pb extracted by the Tiny-TIM method in the < 2 mm size fraction (Table 9.1). The < 250 µm size fraction is the standard for conducting *in vivo* studies of relative lead bioavailability in the United States (U.S. EPA, 2007b). The particle size is also important because the IVD and Tiny-TIM results are being compared to animal

and human results based on the < 250 µm size fraction. This may skew the direct comparison of the *in vitro* and *in vivo* results.

As indicated in the report, the large differences may be due, in part, to variation among the methods (e.g., centrifugation versus filtering). Filtering the samples removes the large soil particles. Centrifuging the samples continues the extraction process and does not remove the large soil particles, thus allowing the extraction to continue. This may result in significantly higher lead *in vitro* values for the IVD versus the Tiny-TIM, which was evident in Table 9.1.

3. How well have both models been compared to human and animal bioavailability data, see Chapter 9 and 10, and especially Table 9.3? Can you indicate the weak points in the comparison to in vivo data that has been performed for each model? What do you think of the manner of comparison for each digestion method?

As discussed in our response to the previous question, the Tiny-TIM model results have not been compared to *in vivo* models, except for one human study. Thus, one cannot necessarily conclude that the IVD model yields better estimates of *in vivo* relative bioavailability. We agree with the conclusion that differences in total concentration are not responsible for the diverging bioaccessibility results.

In terms of comparing the results to human data, it is important to note that adult volunteers were used in Maddaloni et al. (1998), while the juvenile swine model is intended to represent absorption in young children. This distinction is important because children absorb lead to a much greater extent than adults. If the *in vitro* models are intended to represent absorption in children, then would expect the bioaccessible fraction to be greater than the values reported in Maddaloni et al. (1998). In addition, while the comparison of a soil used in a human dosing trial is important (see Table 9.2), there is only one soil for comparison. It is difficult to conclude which method is more appropriate based on one soil. We also believe too much emphasis is placed on the comparison of the very limited human data to the bioaccessibility results.

Table 9.3 shows relatively large differences between the results of the IVD and Tiny-TIM assays for the *"fasted"* and *"fed"* assay conditions. It is interesting to note that the Tiny-TIM assay shows a pronounced effect of fasting assay conditions on bioaccessibility that may correspond to higher *in vivo* RBA than has been observed in fasted conditions in humans (e.g., Maddaloni et al., 1998). The swine data used to evaluate predictions from the IVD assay were from swine dosed 2 hours before morning and evening meals and, therefore, represent more of a fasted condition than a fed condition. Comparisons of *in vivo* absolute bioavailability (ABA) of lead acetate in swine with or without food showed a pronounced effect of food on ABA (ABA with food ≈0.5x without food; U.S. EPA, 2007b). However, the effect of food on soil lead RBA was not evaluated.

As mentioned earlier, the Tiny-TIM and IVD bioaccessibility values reported for both NIST SRM 2710 and 2711 are significantly less than the values of 75% and 84%, respectively, reported by Drexler and Brattin (2007). The Tiny-TIM and IVD results do not seem plausible because the NIST SRMs are ground to a homogeneous particle size of 74 microns. Thus, one would expect relatively high bioaccessibility results for both SRMs, based on particle size alone.

One of the weak points in the comparison is that the IVD and Tiny-TIM results were compared to *in vivo* results based on the < 250 µm size fraction. This makes a direct comparison difficult to interpret because particle size can profoundly impact bioaccessibility and bioavailability.

Another weakness of this study is that there is no direct comparison of the *in vitro* bioaccessibility and *in vivo* bioavailability for made grounds. We suggest performing a swine feeding assay with a small number of made grounds for direct comparison. As discussed in Section 9.4.5, the quantity of made ground used in the *in vitro* methods should be representative of the site. The Tiny-TIM method uses 5 grams, which does not mimic realistic exposures from "hand-to-mouth behavior," while the IVD method uses too little soil (0.06 and 0.04 g) to be representative of the site. We agree with the recommendation in Section 9.6 to perform an *in vivo* study in swine for the made grounds so that a direct comparison of the *in vitro* and *in vivo* data could be determined.

4. Is it possible to indicate, based on the present information, which model you expect to give the most realistic estimate of relative bioavailability? This includes both the determination of the bioaccessibility, and the translation of bioaccessibility to a relative bioavailability factor for lead in soil, see Chapter 5, 8 and 9. Which model is that and why do you think so?

No, it is not possible to determine which model gives the most realistic estimate of relative bioavailability. With that being said, the IVD results seem more believable for a couple of reasons. Based on the speciation results, one would expect elevated bioavailability for most soil samples. The IVD results generally correspond with the expected results based on mineralogy and particle size. In contrast, the Tiny-TIM model yielded low bioaccessibility results for all soil samples and results did not vary based on mineralogy or particle size. This result does not seem plausible because these factors are known to significantly affect oral absorption of lead. We recommend performing an *in vivo* study in swine for the made grounds so that a direct comparison of the *in vitro* and *in vivo* data could be determined.

5. What is your recommendation in order to be able to appoint or develop an in vitro digestion model that can be safely applied in human health risk assessment of contaminated soils? Do you think one of the presently discussed models can be used for this? A balance is sought between the additional experimental research, the reliability of the method, the costs, and the time frame before implementation.

U.S. EPA has evaluated and validated the single (gastric) phase IVBA assay reported in U.S. EPA (2007) for use in risk assessment. Simplicity, implementability and inter-laboratory reproducibility are important factors for selection of a model to be used for regulatory purposes in the U.S. We agree that investigating the differences between both models will not help determine which model is better for lead risk assessment purposes. This can only be accomplished by comparing results to an *in vivo* model which adequately predicts absorption in humans. This comparison must be made for a robust set of soil samples that represent a large range of soil mineralogy and bioavailability values.

Based on this report, the Tiny-TIM model does not appear to be a viable option for estimating the bioavailability of lead due to complexity, cost, and bioaccessibility results that are consistently lower than alternative *in vitro* methods. Thus, the IVD method holds more promise as a viable method for estimating the bioavailability of lead. We also recommend that RIVM evaluate whether simplifying the model to one or two phases would yield more reliable results. Last of all, a direct comparison of the *in vitro* and *in vivo* results for the made grounds should be conducted to determine which model to implement for risk assessment purposes.

REFERENCES

Drexler, J.W. and Brattin, W.J. 2007. *In vitro* procedure for estimate of lead relative bioavailability. *Human and Ecological Risk Assessment*. 13: 383-401.

Maddaloni, M., Lolacano, N., Manton, W., Blum, C., Drexler, J., Graziano, J. 1998. Bioavailability of soilborne lead in adults, by stable isotope dilution. *Environ. Health Perspect.* 106: 1589-1594.

Oomen, A.G., Brandon, E.F.A., Swartjes, F.A., Sips, A.J.A.M. 2006. How can information on oral bioavailability improve human health risk assessment for lead-contaminated soils? RIVM Report 711701042.

U.S. EPA. 2007a. Guidance for Evaluating the Bioavailability of Metals in Soils for Use in Human Health Risk Assessment. Available online at: http://www.epa.gov/superfund/bioavailability/guidance.htm

U.S. EPA. 2007b. Estimation of Relative Bioavailability of Lead in Soil and Soil-like Materials Using In Vivo and In Vitro Methods. Available online at: http://www.epa.gov/superfund/bioavailability/guidance.htm

U.S. EPA. 2007c. Standard Operating Procedure for an In Vitro Bioaccessibility Assay for Lead in Soil. Available online at: http://www.epa.gov/superfund/bioavailability/guidance.htm

Weis, C.P. and LaVelle, J.M. 1991. Characteristics to consider when choosing an animal model for the study of lead bioavailability. *Chem. Speciation and Bioavailability*. 3: 113-119.

2 Review Stan Casteel

Recommend putting in a list of acronyms in the front of the report.

2.3 Soil Preparation

Was the soil split prior to processing – that is, prior to drying, milling, sieving, and homogenization? Homogenization should be the step performed last just prior to splitting. What was the method of homogenization?

This could explain much of the variability in analytical results of soil samples.

On Mixing and Demixing

Julio M. Ottino[1] and Richard M. Lueptow[2] *Science* 15 February 2008: Vol. 319. no. 5865, pp. 912 - 913

Trying to mix two dissimilar granular materials--such as light and heavy or small and large particles--may lead to counterintuitive results: Putting more and more energy into mixing may actually result in more and more demixing. The robust and varied patterns resulting from demixing have long puzzled practitioners and researchers alike. How can one mix something that does not want to mix?

4.2.2 **Methods** (Lead Characterization (analysis in made grounds)

Table 4.1 needs to include a measure of variability in order to evaluate and compare the different analytical methods. Also, quality control results are lacking for the different methods—references soils (preferably 2 or 3 NIST reference materials) should be assessed by each of the methods. Given the existing information in the table, there is no way to assess the true differences in the analytical methods.

9.3 Results section

The basis for selection of which *in vitro* model provides the best estimate of the human/animal condition should be made on how well it correlates with the *in vivo* data. Little is gained by comparing one *in vitro* method to another. In vitro correlation with the *in vivo* should be the driver.

9.4 Physiological Conditions

Too much attention is focused on concern for physiological conditions. This adds another layer of complication and variability to the results making the validation of what should be a simple and reliable method of *in vitro* assessment of bioaccessibility much more difficult. Suggest you go with the fasted state which is most often used in animal models to mimic the between-meal exposure of children to contaminated soil. This adds a layer of safety for the target human population.

New approaches must continue to be validated via intensive peer review, comparison against actual observations in contaminated settings, and pilot testing before they can be applied as unambiguous regulatory tools. Strict validation criteria are necessary to avoid premature or inappropriate application of methods. *In vitro* methods may vary widely in terms of their specific estimates of bioaccessibility. The important issues are that the methods must correlate consistently with *in vivo* methods, regardless of their complexity, and this correlation must exist across laboratories and for a variety of test materials. The simple method with the best correlation to *in vivo* should be the goal.

Based on current technology there will not be an *in vitro* method that works well to assess all contaminants and media types. Clearly, we know from experience that a new and tailored *in vitro* method must be developed for every contaminant of concern. Lead methods do not work for arsenic, cadmium, vanadium, or organic compounds. Only a well developed animal model has broader application and even the best animal models are not suitable for all organic compounds. The US Environmental Protection Agency has selected the immature swine model as the best surrogate for human children especially when it comes to lead bioavailability. There is good correlation with the *in vitro* testing over a broad range of soil types, lead forms and concentrations. The usefulness of this model is currently being extended and modified to evaluate arsenic bioavailability in other soils and test materials. The immature swine model also looks quite useful for this application. To date, *in vitro* testing has not shown good correlation with the swine model results for arsenic.

Based on the large number of soils containing lead exceeding the established intervention value, remediation via a soil amendment might be a more economical way to lower the bioavailability. We have successfully applied phosphate to soils thereby reducing bioavailability by 30-40%.

3. Review Pat Rasmussen

Background:

The 2008 RIVM document "Relative oral bioavailability of lead from Dutch made grounds" reports on the application of two different *in vitro* test methods to estimate oral bioaccessibility of Pb, IVD and Tiny-TIM, to a set of samples which were collected from Dutch made grounds and characterized using a variety of methods including SEM, multi-element ICP-OES analysis, and Pb isotopic ICP-MS analysis.

The report concludes that the IVD method and the Tiny-TIM method yielded contrasting estimates for the oral bioaccessibility and the relative bioavailability factor for Pb from the Dutch made grounds: 0.67 relative bioavailability for the IVD method (50^{th} %ile) compared to 0.12 for the Tiny-TIM method (50^{th} %ile).

The report also observes that the above present findings for Dutch made grounds are in contradiction with a previous conclusion that both models estimated the relative bioavailability of Pb "in the correct range". As this contradiction has important implications for risk assessment and risk management, RIVM and TNO are seeking comments from external reviewers on the reasons for the differences, the relative merits of the two methods, and recommendations for next steps to determine whether these or other *in vitro* test methods can be safely applied in risk assessment and risk management decision-making.

The following responses address specific questions posed by RIVM and TNO, followed by some general comments.

Responses to Questions:

1. **Can you indicate the strong and weak aspects of both *in vitro* digestion models? How can weak aspects be improved?**

Strong aspects of the IVD method include:
– the attention to physiologically relevant pH conditions (as pH is a key control on release of metals from mineral phases) and consistency in temperature and incubation times.
- the relative simplicity of the technique. A method that is to be used in regulations should be simple and easily reproducible.
- attention to analytical QA (duplicates, procedural blanks and measurement of certified reference materials). More details could be added such as determination of the detection limit.
-preliminary validation of *in vitro* test results using *in vivo* bioavailability data collected for US-EPA soils using a relevant animal model (swine).

Weak points of the IVD method are:
- the pH problems leading to elimination of some data points in Figure 5.1. Difficulties appear to arise with the high pH encountered in some of the Dutch samples (pH 6.0 to 9.5; Table 3.1). One suggestion could be to simplify the extraction to a gastric simulation only, with dropwise addition of acid to adjust/maintain pH.

Weakness of both methods:
- Difficulties with obtaining a reliable total Pb value (for the denominator) need to be resolved (as discussed on page 31). More CRMs are recommended to represent (match) a wide range of matrices and a wide range of total Pb concentrations (currently it appears only NIST Montana Soil 2710 is routinely included) to ensure that all forms of Pb are put into solution, from all types of matrices.
- To obtain accurate and reproducible total values, the more aggressive acid digestion including HF (such as the Deltares method described on page 21) is recommended because the use of diluted aqua regia does not release Pb equally efficiently from all matrices (as discussed on page 30). An inefficient extraction resulting in poor total Pb recovery for some made grounds and not others could introduce a significant bias in the bioaccessibility estimates for the different locations/anthropogenic sources.
- For the total Pb extraction, microwave digestion extraction efficiency is influenced by the soil:liquid ratio (125 mg in 5 mL of acid) – the higher the soil: liquid ratio the lower the recovery of Pb. Changing this ratio to 1:500 or 1:1000 will assist in improving the recovery.

For the Tiny-TIM method a significant weakness is revealed in Figure 7.2 which shows that Tiny Tim yields similar relative oral bioavailability results for high PPS ranked soils (e.g. #43 and #53) compared to lower PPS ranked soils (e.g. #17 and #63). This lack of sensitivity to physical/chemical differences in Pb occurrences in Dutch made lands is a key concern, as these differences are important controls on biological availability. Some weaknesses (and possible causes for this lack of response) are suggested:

- pH is uncontrolled – maintaining the low pH of the gastric stage is important as this is the stage most likely to release Pb. The elevated pH of some Dutch soil samples may interfere with the release of Pb under uncontrolled conditions.
- Incubation time in the simulated stomach conditions (length of time exposed to gastric acid) is too short, and the pH is not low enough, to mimic fasted conditions in a child's stomach. In contradiction with the statement in the report that Tiny-TIM is characterized by "realistic transit times", methods which have longer (2-hr) incubation at pH 1.5 show acceptable correlations with swine bioavailability data.
- Agitation in the stomach compartment appears to be absent – a sedentary model does not appropriately simulate normal activity of a child, and will cause inadequate physical contact between the Pb minerals and the gastric solution.

- soil:liquid ratio (5 g soil in 120 g liquid) does not allow adequate physical contact of the Pb minerals with the gastric solution – it also does not mimic fasted conditions (literature indicates soil-to-stomach acid ratios in children may be as high as 1:10,000)
- early exclusion of gastric residue and early exclusion of all large molecules including organo-Pb complexes, due to the separation after only a few minutes of incubation, does not allow for the slow release of Pb from soil particles, including organo-Pb complexes, which occurs in the stomach.

2. Do you have ideas about the main reasons for the large differences in bioaccessibility obtained with both models (see discussion in Chapter 9)?

The lower bioaccessibility obtained with Tiny-TIM compared to IVD would be caused by a combination of the following differences:

- the early exclusion of residual particles, larger molecules and complexes in Tiny-TIM
- the shorter incubation time in the gastric compartment of Tiny-TIM, which does not allow time for Pb to be released to solutions from slow-release minerals, anthropogenic compounds, and complexes
- higher pH conditions for a greater proportion of experimental run time. It appears that pH is >1.5 for >90% of the 6 hour experimental run time for Tiny-TIM.
- limited contact between particles and gastric fluid in Tiny-TIM, caused by the different soil:liquid ratio, in combination with lack of agitation.
- the addition of a greater variety of ingredients in Tiny-TIM that may act as complexing agents resulting in losses of Pb from solution at all stages.
- the absence of simulation of a completely fasted state for Tiny-TIM.

3. How well have both models been compared to human and animal bioavailability data, see Chapter 9 and 10, and especially Table 9.3? (a) Can you indicate the weak points in the comparison to in vivo data that has been performed for each model?

The IVD method results have been applied to US-EPA soils and plotted against *in vivo* swine data for these soils, with an acceptable correlation and slope using 0.06 g soil per digestion tube (Pg 53 Figure 5.1). The weakness in the comparison lies in the small number of data points on the graph, especially in the intermediate range (bioaccessibility between 20 % and 60 %). Some data points were removed from this graph due to problems with suboptimal pH incubations, so improving the pH incubation aspect of the IVD method is recommended.

For Tiny-TIM, a fundamental weakness is the lack of comparison to bioavailabilities determined using juvenile swine. Swine have been accepted for simulations of oral bioavailability of Pb in humans, whereas rodents have not.

With respect to human bioavailability data, a single data point (Bunker Hill) is insufficient for comment (applies to both methods).

Table 9.3 Spiked Solution Data For the IVD method: More information is needed on the experimental conditions for the IVD Pb acetate experiment used to obtain the information in Table 9.3. (e.g. was 0.06 g Pb acetate used as in the tests for soil?) Section 5.2.1 p. 54 refers to Chapter 9 for more information, but experimental conditions are not given in either chapter.

Table 9.3 Spiked Solution Data For Tiny-TIM: The experimental conditions described in 8.2.2.1 do not reflect the test conditions used for real samples (5 grams of soil in 120 g saliva/water). This suggests two possible experiments to improve the translation to relative bioavailability (1) select a mass of Pb acetate that better represents the Pb content of the

Dutch samples, with the same solid:liquid ratio as used for the real samples, and (2) lower the solid:liquid ratio for real samples to match that of the previous Pb acetate experiments (1 mg in 125 g (low) and 8 mg (high) in 125 g saliva/water.

(b) What do you think of the manner of comparison for each digestion method?

An important strength of this study is the decision to determine the bioavailability of lead from a variety of soils representative of the Dutch made lands, and representative of different Pb species and particle size distribution, based on the previously observed correlation between bioavailability of lead and lead speciation (Oomen et al., 2006; Casteel et al., 2006; US-EPA guidance, and others). The results that are yielded by this approach underscore the importance of site specific information, and allow more accurate risk assessments, as discussed in Section 10.2 page 99.

In general the use of the <2 mm particle size fraction may underestimate the Pb concentration of particles that stick to a child's hand (generally considered to be <150 micron). The use of the <2 mm fraction is standard practice in soil/agricultural science, but in not human exposure studies. Due to the higher mass-to-volume ratio of fine particles, metals concentrations are often higher in finer size fractions. For samples in which Pb is more concentrated in the fine fraction, the use of smaller size fraction can result in a higher bioaccessibility result than the use of a larger size fraction. Consideration of smaller size fractions in future risk assessments is recommended.

There is a lack of analytical quality assessment and quality control (QA/QC) data to demonstrate the reliability of the Tiny-TIM results. Information on QA/QC (including procedural blanks) was provided for the IVD method and for the characterization of the soils, but was missing for the more complex Tiny-TIM method. Accurate measurement of Pb in the various Tiny-TIM extracts would be expected to present an analytical challenge. No information on between-run cleaning procedures for the Tiny-TIM has been provided. Data from procedural blanks would be needed to quantify the contribution of cross-contamination (caused by carry-over of Pb between runs). Cross-contamination would be expected if a high-concentration sample immediately preceded a low-concentration sample. Without QAQC data for Tiny-TIM, reviewers cannot know if the Pb results from the Tiny-TIM are meaningful.

On page 96, the report states that the IVD centrifugation step to separate the released lead from the remaining soil could result in the presence of insoluble large lead-complexes in the liquid fraction, leading to an over-estimate of bioaccessibility. This may be true, but it is also expected that there are large lead-complexes in the liquid fraction which are soluble in a child's stomach. Not including such soluble large lead-complexes would lead to an underestimate of bioaccessibility.

Similarly, on the same page (96) it is stated that it is unknown whether the Tiny-TIM model results in an under- or overestimation of the total lead bioavailability for a child. Based on the results provided in this report, the Tiny-TIM experimental conditions appear unlikely to result in an overestimation, but this will be confirmed after comparison with an appropriate animal model using an appropriate range of sample matrices. In the meantime, it should be added that a method such as Tiny-TIM which omits or excludes soluble large complexes would likely result in an underestimation of the estimated bioaccessibility of lead.

4 Is it possible to indicate, based on the present information, which model you expect to give the most realistic estimate of relative bioavailability? This includes both the determination of the bioaccessibility, and the translation of bioaccessibility to a relative bioavailability factor for lead in soil, see Chapter 5, 8 and 9. Which model is that and why do you think so?

The preliminary testing of both methods using real samples containing Pb compounds with contrasting solubilities (release rates of Pb) and contrasting particle size distributions, as summarized in Figure 7.2 on page 68, is central to the evaluation of the relative merits of the two methods. This preliminary work suggests that more realistic estimates of relative bioavailability are likely to be obtained using the IVD model, based on the general positive relationship between Rel F and PPS ranking observed using the IVD model.

In contrast, the lack of a relationship between Tiny-TIM and PPS rankings in Figure 7.2 indicates that Tiny-Tim is insensitive to differences in Pb speciation and other soil characteristics that control Pb bioavailability. These results (as presented) cast doubt on the validity of Tiny-Tim for the purpose of risk assessment of Pb from Dutch made grounds. To be considered realistic, Tiny-Tim should demonstrate responsiveness to proven chemical differences and proven differences in the biological availability of Pb in samples from different made grounds.

Section 5.1.3 The argument that bioaccessibility should be measured under fasted conditions (accompanied by calculations of average physiological conditions where required) is well-supported in this report (page 51-52). An additional benefit of using fasted conditions is the overall simplification of the experimental conditions and greater likelihood of consistent and reproducible analytical results.

Figure 5.1 page 53. This graph makes a compelling argument that the soil:liquid ratio should be 1:1000 rather than 1:100, under the other IVD experimental conditions. It would be helpful to include vertical error bars associated with the y-axis - ie. the sd of the *in vitro* bioaccessibility measurements – in the present figure error bars are presented only for the *in vivo* measurements.

Section 8.2.2.1 For Tiny-TIM, Pb acetate is introduced as 1 mg in 125 g (low) and 8 mg (high) in 125 g saliva/water solution in the fasted state (pg 77), and the corresponding bioaccessibility is 66% and 56% respectively (page 80). This contrasts with conditions used for real samples where 5 grams of soil are mixed with 120 g saliva/water. Thus the Pb acetate experimental conditions do not reflect the test conditions used for real samples. In the case of Nijmegen for example, 5 grams of soil might contain about 15 mg of Pb. This suggests two possible experiments to improve the translation to relative bioavailability (1) use the same mass of Pb acetate to be representative of the Pb range in the Dutch samples, and the same solid:liquid ratio as used for the real samples, and (2) match the solid:liquid ratio for real samples to the existing Pb acetate experiments (1 mg soil: 125 g saliva/water (low) and 8 mg soil: 125 g saliva/water (high).

Table 8.2 If these Tiny-TIM bioaccessibility (%DE) values are duplicates it would be more correct to express variability as relative percent error – calculation of standard deviation requires minimum of n = 3.

Section 9.4.2 It is argued that the Tiny-TIM relative bioavailability value could be multiplied either by a factor of 2.5 (pg 93) or by a factor of 8.3 (page 94). This points to large uncertainty with Tiny-TIM in the method of calculating relative bioavailability.

Section 8.1.1 The claim on page 74 that Tiny-TIM provides a realistic transit time than is unsubstantiated for Pb-containing particles. One of the most important controls on the release

of Pb from minerals and compounds is pH. It is under the low pH conditions of the stomach compartment where Pb is most likely to be released, yet the DE represents only a brief exposure of gastric fluid – less than 10 % of the 6 hours is spent at pH 1.5. This is contrary to the IVD method and others such as the Drexler method which have longer incubation times and yield results that compare well with swine bioavailability data.

5. What is your recommendation in order to be able to appoint or develop an *in vitro* digestion model that can be safely applied in human health risk assessment of contaminated soils? Do you think one of the presently discussed models can be used for this. A balance is sought between the additional experimental research, the reliability of the method, the costs, and the time frame before implementation.

To be safely applied in human health risk assessment, the *in vitro* test method should be responsive (i.e. show a positive correlation) to proven differences in the biological availability of Pb in samples from different made grounds (using appropriate animal models) and to physical and chemical factors proven to influence Pb release rates from soil constituents (particle size and solublilities, respectively). Early indications from Fig. 7.2 suggest that Tiny-TIM is not suitably responsive, but with a few improvements on the IVD (or substitution of a simplified gastric version of IVD), the IVD test may prove to be useful predictive tool.

As a first step, more quantitative characterization of representative samples of the Dutch made grounds would be a worthwhile investment to improve on the PPS ranking. This would include better information on particle size distribution, and further identification of Pb-bearing compounds, including both inorganic and organic compounds. Use of the SEM is a good start, but more detailed understanding of the composition and morphology of the fine-grained particles can be obtained using micro-XRD and micro-XRF analyses. The XANES approach for bulk analysis using a carefully selected series of standards representing all of the important organic and inorganic Pb compounds will also assist in providing a more quantitative PPS ranking.

In general for all samples, understanding the proportion of total Pb that is made up of organic-Pb compounds arising from anthropogenic sources will be helpful. The loss-on-ignition method used in this study does not distinguish between natural and anthropogenic organic compounds, but a carefully chosen set of organic Pb standards and XANES analysis may assist in providing this information. Table 4.3 shows that the Pb content of organic matter in #63 Niijmegen sample (0.335 wt%) is higher than the total Pb concentration in the bulk sample (appendix, page 132) underscoring the importance of identifying the nature and source of the organic Pb compounds. As this sample is characterized as “city waste” it is likely that anthropogenic organic Pb compounds are present. Moreover, knowledge of the anthropogenic versus natural organic compounds would shed light on the influence of organic matter on bioaccessibility – for example it might be hypothesized that natural organic matter may decrease Pb bioaccessibility, while anthropogenic Pb compounds may have very high bioaccessibilities.

After the PPS ranking is improved for a complete representative set of Dutch made ground samples, then animal bioavailability data should be obtained in triplicate for each of the ranked samples (using juvenile swine). This will provide a valuable and relevant set of control standards on which to evaluate *in vitro* test methods against the “gold standard” *in vivo* bioavailability data.

Before acceptance as a regulatory protocol the *in vitro* test methods should be run in triplicate for each sample and should be accompanied by rigorous QA/QC data requirements (recovery measured by more CRMs; procedural blanks, control of pH conditions).

Finally, data for the inhalation pathway should be included in the risk assessment for each of the Dutch made lands. For this, sampling and analysis of the inhalable and respirable fractions of windblown dust particles containing Pb and other metals will be required.

General Comments:

The report presents the argument that use of the lower relative bioavailability factor yielded by Tiny-Tim would reduce costs of soil management decisions. However, to estimate the cost-benefit associated with each approach, the economic costs of potential human health impacts should also be estimated (ie the cost to the Netherlands of chronic childhood exposure to Pb at whatever target concentration is associated with each method).

Internationally many organizations, including Health Canada, are re-evaluating the TRV for Pb due to growing evidence of negative health effects (neurological and other impacts) at blood lead concentrations lower than 10 µg/dL. Therefore it is recommended that Section 10.1 be updated to reflect the discussion in current international literature (within the last 5 years) related to the re-evaluation of the TRV for Pb.

The information provided on the Pb content of the "rural and background soils" is useful and could be added to Table 3.2 and Figure 3.2. This information helps place the current Dutch intervention value (530 mg/kg; page 96) into context. In Canada, the Canadian Council of Ministers of the Environment (CCME) guideline for residential soil Pb is 140 mg/kg.

The gradual acceptance (by some organizations) of adjustments for oral bioavailability of Pb is based on the body of evidence that different Pb compounds in soil have different chemical solubilities, and that these chemical differences are reflected by corresponding differences in bioavailability determined using swine models. Other important sample characteristics affecting Pb bioavailability include differences in grain size, texture, and morphology. Whichever *in vitro* test method is ultimately chosen (if any), the *in vitro* test results should be responsive to these differences in physical, biological, and chemical characteristics of the Pb occurrences in the Dutch made ground samples.

In this regard, the IVD method shows more promise than the Tiny-Tim method based on the preliminary findings of the present study. The apparent contradiction between current and previous results arises because the relative merits of the two bioaccessibility test methods could not be fully evaluated until both methods were applied to a wide range of actual samples from Dutch made grounds. The new information gained in the current study eclipses the previous conclusion that both models estimate relative bioavailability "in the correct range". The necessary next step will be to compare the *in vitro* test results of both methods with bioavailability data assessed for an appropriate number of representative samples using swine models.

4. Review Sohel Saikat

- **Can you indicate the strong and weak aspects of both *in vitro* digestion models? How can weak aspects be improved?**

Of the two in-vitro models, from an operational point of view, IVD is comparatively simple, cheap and easy to assemble and handle for a routine application during risk assessment. The TIM model has been used for a wide range of compounds in relatively simple matrices (e.g. food). The application of the TIM model for wide range of soil types and lead minerals are limited compared to the IVD model. If validation with human bioavailability or appropriate in-vivo data as a surrogate of human volunteers is the key criterion, one could question the size and power of the data-base of such evaluations for both models. Based on the observations in Denmark, UK and

USA, it may be that one model and particular set-conditions are not always robust enough to adequately predict bioavailability. If that is the case, IVD offers more flexibility in adapting to various conditions. In terms of strength and weakness, TIM is more physiologically based whereas IVD is designed to be adaptable for routine application and therefore has compromised aspects of physiological relevance. As it is not feasible or preferable to undertake human volunteer or animal studies in the risk assessment of land contamination, therefore a consensus is required whether to channel our trusts/confidence in having a more physiologically based method or a method with a high degree of predictive capability of bioavailability (may be at the compromise of some/various aspects of physiological relevance). This aspect is crucial and would need a consensus from science and policy in land contamination.

- Do you have ideas about the main reasons for the large differences in bioaccessibility obtained with both models (see discussion in Chapter 9)?

As the report described, both models are aimed at simulating human gut (gastrointestinal tract) conditions, although, they are different to each other in terms of design, chemistry in digestion, and operating procedure. The way data are generated and calculated in bioavailability determination are also different. These have collectively contributed to large differences in the end results. The individual contribution of the procedural differences can be evaluated and quantified through undertaking a sensitivity analysis, which could allow the ranking of various parameters based on their influences. The effect of centrifugation vs ultrafiltration has been indicated in the study. The effect of pH on the in-vitro and in-vivo correlation has also been well established in the Danish EPA sponsored study undertaken by the DHI Denmark. Model comparison or sensitivity analysis should centre around certain standard(s).

- **How well have both models been compared to human and animal bioavailability data, see Chapter 9 and 10, and especially Table 9.3? Can you indicate the weak points in the comparison to in vivo data that has been performed for each model? What do you think of the manner of comparison for each digestion method?**

Given the size of the soil bioavailability data, it may be considered premature to make a meaningful comparison for the two models. Of the data provided (Table 9.3), Flanderes and Montana 2711 soils have no in-vivo data; Oker 11 result is considered not valid; Montana 2710 is with rat data (not an appropriate human surrogate, rat data were also excluded in the DHI Denmark study undertaken for the Danish EPA). For the Bunker Hill soil, although, two models produce comparable data in fasted test condition, such comparability does not exist in fed conditions indicating considerably higher bioaccessibility data by the IVD model. A similar trend was also observed for food matrices. It may therefore be inadequate to form a conclusion based on just one soil sample. One way, this can be addressed is by expanding the power of the data by using more in-vivo tested samples (*e.g.* soils used under the US EPA lead in-vitro method validation study).

- **Is it possible to indicate, based on the present information, which model you expect to give the most realistic estimate of relative bioavailability? This includes both the determination of the bioaccessibility, and the translation of bioaccessibility to a relative bioavailability factor for lead in soil, see Chapter 5, 8 and 9. Which model is that and why do you think so?**

Both models produce data which are indicative of lead solubility in simulated gut conditions. However there has been significant variability in their bioavailability measurement. TIM has been accepted as physiologically more relevant than the IVD model. With PPS ranking (lead characteristics in soil), IVD model predicted bioavailability correlates well whereas TIM doesn't. Given the inadequacy of the data on in-vitro and in-vivo correlation, it would be premature to form a conclusion on the bioavailability estimates of the models. However, in-vitro data produced by both or one of the models could be of value to characterise a site and inform the detailed quantitative risk assessment when they are complemented with other lines of evidence.

If it is scientifically established that lead absorption primarily occurs in the intestine (Oomen et al. 2006) this means that for lead, TIM model is physiologically limited to indicate actual bioavailability as it does not calculate the test material that has not been emptied from the gastric compartment during the experiment.

- **What is your recommendation in order to be able to appoint or develop an *in vitro* digestion model that can be safely applied in human health risk assessment of contaminated**

soils? Do you think one of the presently discussed models can be used for this. A balance is sought between the additional experimental research, the reliability of the method, the costs, and the time frame before implementation.

The issues that have been identified in the report are not uncommon and could be reiterated with other models in similar studies. From a regulatory side, there needs to be an acceptance that in-vitro data will not be as good as in-vivo data and that in-vitro model predicted bioavailability measurement will have various degrees of uncertainty. As scientists, efforts should be directed at identifying and quantifying these uncertainties, and what other lines of evidence could be used to evaluate/complement the in-vitro measurement.

Large scale animal studies are cost prohibitive and will generate ethical reservation. Therefore any of the following options can be explored to validate the model:

i) Use of soil samples from other studies tested with pig (swine) for bioavailability determination to evaluate/validate in-vitro models (TIM/IVD). A geochemical comparison would need to be made between in-vivo tested samples and representative Dutch made ground soils. If considerable variability is identified, careful evaluation of the extent to which these methods can be applied need to be determined.

ii) A lead biomarker study by measuring lead uptake in people living in representative made ground areas. This study could have the following components:

- Determination of blood lead concentrations
- Determination of soil lead and geochemical characterisation of sites
- Determination of bioaccessibility in soil
- Questionnaire survey to address the effect of confounders

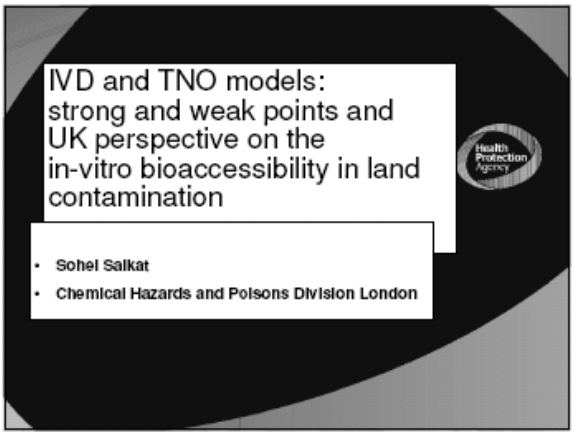
Using this information, assessments could then be made of the correlations between uptake, lead content and bioaccessibility determination. This would provide reasonable information about the validity and appropriateness of the in-vitro methods

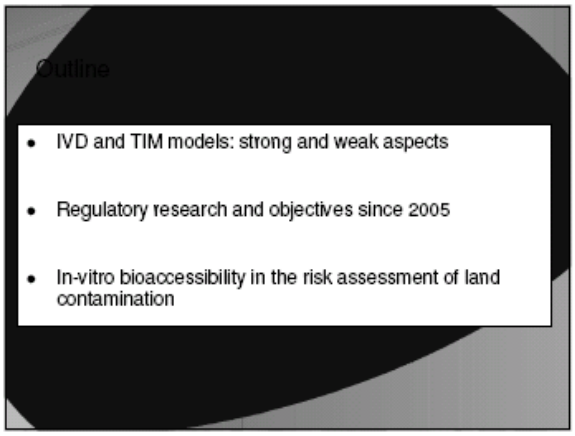
iii) Human volunteer study. This requires volunteers willing to ingest 'contaminated' soils from made ground. This type of study is not completely new and was performed in the US for lead in soils (Maddaloni *et al*. 1998). Finding volunteers for this type of study may not be easy and could also generate ethical reservations.

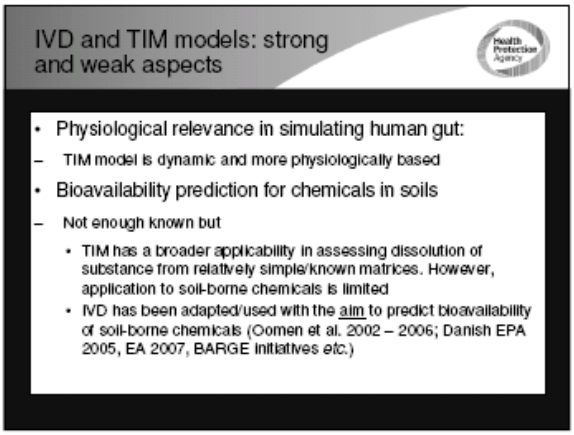
In terms of the choice of methods, IVD has more data on the lead bioaccessibility in soils. This model has showed reasonably good bioavailability prediction for arsenic in soils in the study undertaken by the Environment Agency (unpublished report has been undergoing sign off procedure). In contrast, TIM is technically complex and expensive to assemble and operate and therefore would need adaptation for promoting wider routine application.

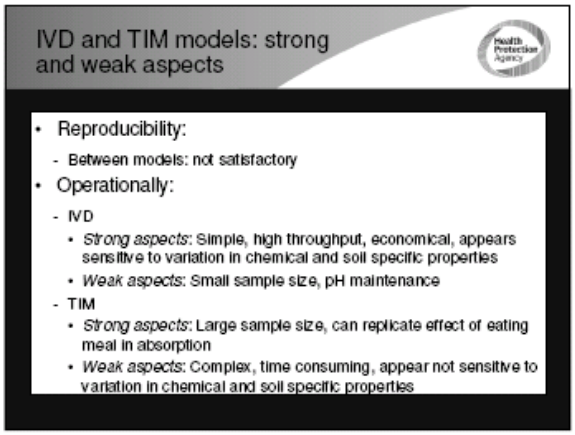
Appendix 3 Presentations

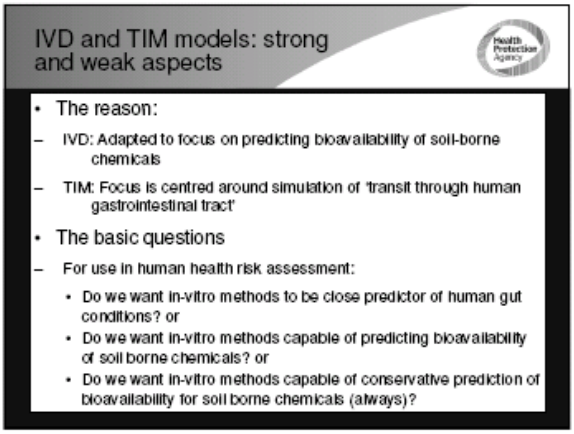
1. Presentation Sohel Saikat

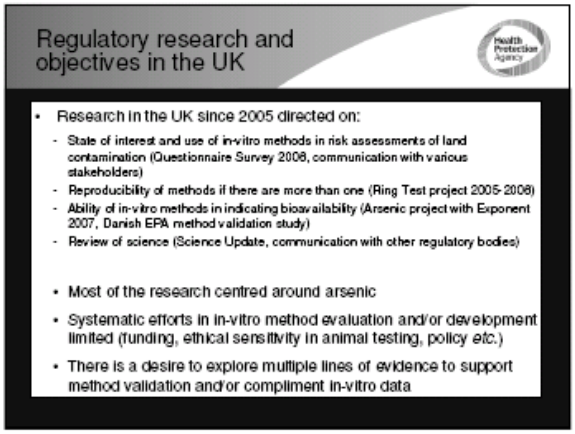












Regulation in dealing with land contamination

Contaminated Land Regime

Part IIA of the Environmental Protection Act 1990

Planning Regime

PPS3 PPS23

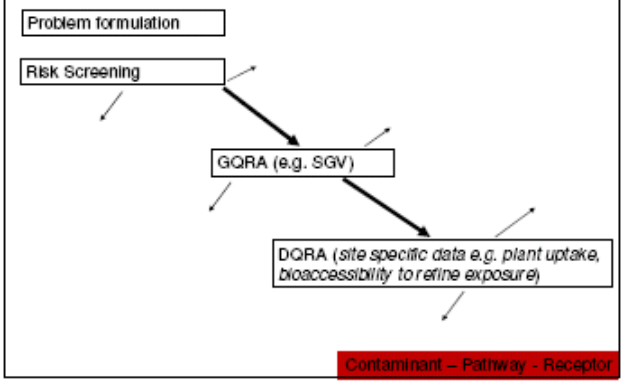
Other Mechanisms

Voluntary agreements

Land contamination

- Regulators
  - Local Authorities
  - Environment Agency
- Public Health
  - Health Protection Agency

Risk assessment approach



Current situation in the UK

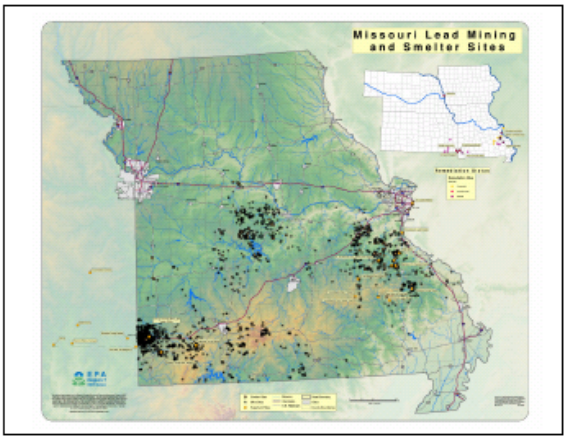
- Reliability of UK *in-vitro* methods in predicting bioavailability is not adequately known
- Inconsistency in the production and use of *in-vitro* data
- Submission of *in-vitro* data as part of DQRA to LAs is on rise
- Environment Agency's view:
  - » 'Given the current uncertainties associated with bioaccessibility testing, we consider its application to be limited at this time'
- The current view is likely to be reviewed in line with ongoing changes in policy and guidance in land contamination

2. Presentation Mike Beringer

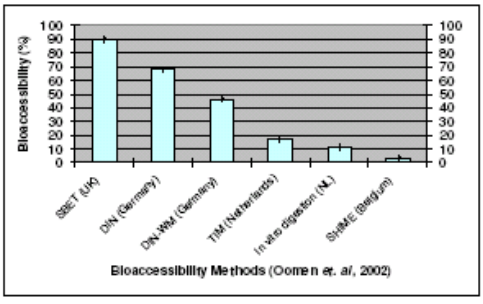
Development and Validation of an *In Vitro* Bioaccessibility Method for Lead in the U.S.

Workshop on Bioavailability of Lead from Dutch Made Grounds
May 14, 2009

Mike Beringer and Karen Bradham
U.S. Environmental Protection Agency



Bioaccessibility of Lead in Montana SRM 2711



Juvenile Swine Method Development

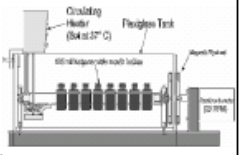
- Leadville, Colorado Superfund Site
  - Lawsuit brought by responsible party
  - Asserted lead bioavailability in soil was very limited
- U.S. EPA Region 8 Bioavailability Studies
  - Research program began in late 1980's/early 1990's
  - Juvenile swine selected as animal model based on similar anatomy and physiology of gastrointestinal tract to human child
  - Phase I: range-finding and pilot studies
  - Phase II: soil testing program of 19 soil and soil-like materials
- Soil Testing Program Objectives
  - Measure RBA to support site-specific adjustments
  - Accumulate knowledge on how RBA depends on speciation

In Vitro Bioaccessibility Method Development

- Several researchers began investigating *in vitro* methods for lead (e.g., PBET) in the late 1980s/early 1990s
- Solubility/Bioavailability Research Consortium (industry, academia, consulting, state/federal government) developed the SOPs and validation exercises
- Side by side comparisons were run with juvenile swine studies
- Temperature, pH, run time, and physiological solutions were adjusted to achieve optimum correlation with juvenile swine results
- 19 test materials were evaluated for lead

Relative Bioaccessibility Leaching Procedure (RBALP – Drexler and Brattin, 2007)

- 1.0 g of <250 µm substrate to 100 mL of extraction fluid
- Extraction fluid
  - 0.4 M glycine in deionized water
  - pH adjusted to 1.5 using HCl



- Mix end-over-end at 37 °C for 1 hour
- Extract supernatant and filter (0.45 microns)
- Measure pH: must be within 1.5 +/- 0.5 pH units

In Vitro Bioaccessibility Method Validation

- OSWER National Bioavailability Work Group - 2002
- Bioavailability Workshop - 2003
- Relying on ICCVAM Criteria (Interagency Coordinating Committee for Validation of Alternative Methods)
  - http://iccvam.niehs.nih.gov/
- Method Validation Criteria
  - Demonstrate method is reliable and relevant for its proposed use
- Regulatory Acceptance Criteria
  - Method fulfills a specific regulatory need

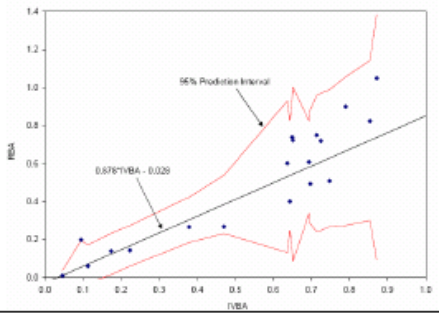
Method Validation Criteria (ICCVAM, 1997)

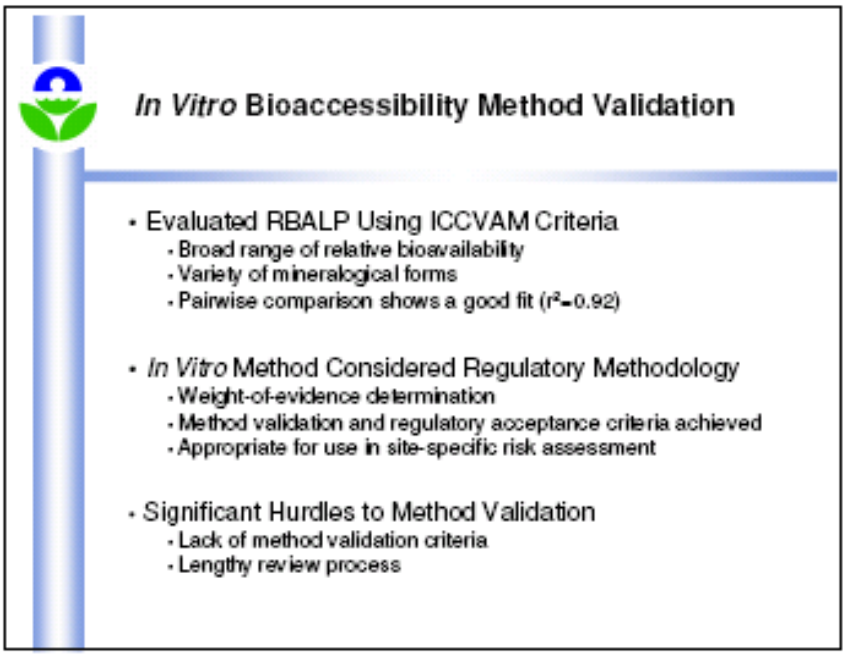
- Scientific and Regulatory Rationale
- Relationship Between Test Method Endpoint and Biological Effect
- Detailed Protocol and Known Limitations
- Within-Test Variability and Reproducibility Among Labs
- Test Method Performance with Representative Agents
- Comparison to Existing Test Method
- Data in Accordance with Good Laboratory Practices (GLP)
- Validity Assessment Data Available for Review
- Independent Scientific Review

Regulatory Acceptance Criteria (ICCVAM, 1997)

- Independent Scientific Peer Review
- Detailed Protocol with SOPs
- Adequately Predicts Bioavailability and Demonstrates a Linkage
- Representative Chemicals Tested
- Generates Data Useful for Risk Assessment Purposes
- Documentation of Strengths and Limitations
- Robust and Transferable
- Time and Cost Effective
- Can Be Harmonized
- Suitable for International Use
- Reduction of Animal Use

Lead – Correlation Between *In Vivo* RBA and *In Vitro* Bioaccessibility (IVBA)




In Vitro Bioaccessibility Method Validation
• Evaluated RBALP Using ICCVAM Criteria
- Broad range of relative bioavailability
- Variety of mineralogical forms
- Pairwise comparison shows a good fit (r^2=0.92)
• *In Vitro* Method Considered Regulatory Methodology
- Weight-of-evidence determination
- Method validation and regulatory acceptance criteria achieved
- Appropriate for use in site-specific risk assessment
• Significant Hurdles to Method Validation
- Lack of method validation criteria
- Lengthy review process

3. Presentation Karen Bradham

EPA

In Vitro Bioaccessibility Method for Lead in the U.S. and Workshop Question 3

Workshop on Bioavailability of Lead from Dutch Made Grounds
May 14, 2009

Karen Bradham and Mike Beringer
U.S. Environmental Protection Agency

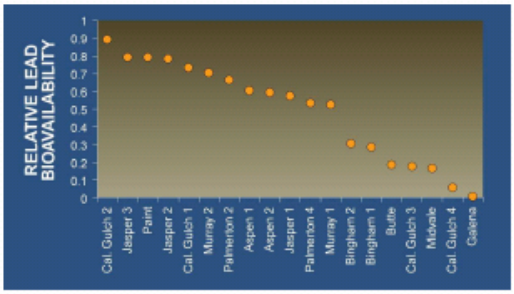
EPA

General Study Design Considerations

- Robust data set
  - concentration range
  - range of RBAs
- Sample preparation
  - Processing (drying, homogenizing, etc.)
  - Particle size (<250 µm)
- Bioaccessibility
  - pH
  - quantity of soil
  - particle size
  - speciation and mineralogy
  - simplicity
- Calculating results for comparison with in vivo results

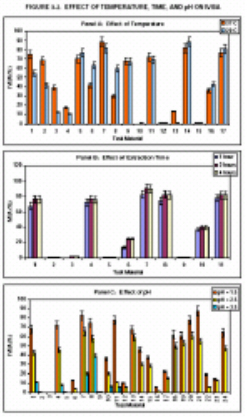
EPA

Relative Oral Lead Bioavailability in Juvenile Swine – Range of RBAs



EPA

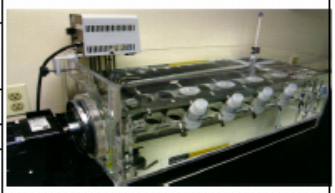
Evaluation of in vitro method parameters for optimum correlation with swine results



EPA

Final In Vitro Method Parameters

Parameter	RBALAP
Soil:Solution Ratio	1:100
Stomach Fluid Composition	0.4 M glycine, HCl added to adjust pH 1.5
Stomach Residence Time	1 hour
Intestine pH	n/a
Intestine Fluid Composition	n/a
Intestine Residence Time	n/a



EPA

In vitro QA/QC

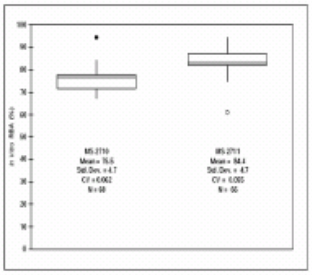
Control limits for these quality control samples were as follows:

Analysis	Frequency	Control Limits
Reagent blank	once per batch	<25 µg/L lead
Bottle blank	5%*	<50 µg/L lead
Blank spike (10 mg/L)	5%*	85-115% recovery
Matrix spike (10 mg/L)	10%*	75-125% recovery
Duplicate sample	10%*	±20% RPD
Control soil (NIST 2711)	5%*	±10% RPD

RPD = Relative percent difference
*Minimum of once per batch

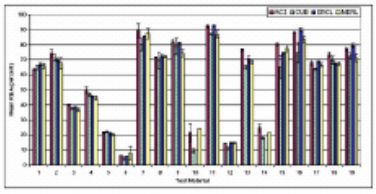
SRMs used to evaluate in vitro method variability

FIGURE 3-3. PRECISION OF IN VITRO BIOACCESSIBILITY MEASUREMENTS



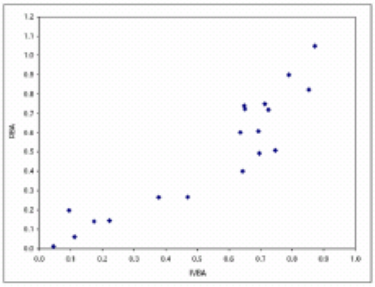
In vitro reproducibility across labs

FIGURE 3-4. REPRODUCIBILITY OF IN VITRO BIOACCESSIBILITY MEASUREMENTS



Direct comparison of IVBA and RBA

FIGURE 3-5. RBA vs. IVBA



Question 3: How well have both models been compared to human and animal data?

- IVD method results applied to US-EPA soils and results were compared with in vivo (swine)
- Tiny-TIM method results for NIST SRM 2710 is comparable to results from rats
- Difficult to determine meaningful comparison of methods with bioavailability data due to the limited number of soils available for comparison

Question 3: How well have both models been compared to human and animal data?

- Comparison of soil used in human dosing trial is important
- Difficult to conclude which method is more appropriate based on one soil
- Human dosing study included adults
  - US-EPA selected juvenile swine as appropriate model to represent absorption in young children
  - Swine model represents more of a fasted condition than fed state (worst case scenario)

Question 3: How well have both models been compared to human and animal data?

- Calculations for these methods may make direct comparison of these methods difficult
  - direct comparison of results with animal and human studies
- Differences in sample particle size complicates comparison of the two methods
  - comparison of in vitro results with animal and human studies
- Method evaluation difficult due to lack of in vivo data for made grounds

Tiny-TIM model comparison - *Can you indicate the weak points in the comparison to in vivo data that has been performed for each model?*

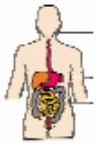
- Lack of comparison of data to juvenile swine bioavailability results
- Tiny-TIM results were compared with the one human study soil and rat data (considered a poor model for measuring Pb bioavailability)
- Large particle size (< 2 mm) used may complicate a direct comparison of results with in vivo (< 250 µm) results
- Difficult to compare results with in vivo results due to a limited QA/QC data available in the report

IVD model comparison - *Can you indicate the weak points in the comparison to in vivo data that has been performed for each model?*

- Using 0.06 grams of soil, the IVD method has been applied to US-EPA soils and resulted in an "acceptable correlation "
- Figure 5.1 - comparison only included a small number of data points and most of these materials had relatively low and high in vivo bioavailability
  - Example: bioavailability data in the mid-range (20-60%) were lacking and some data were removed due to the suboptimal pH incubations
- Small sample size may not be representative of sample

Development of in vitro method for Dutch Made Grounds

- Decision to determine BA of Pb from a variety of Dutch made lands
- Considerations:
  - Sample preparation
  - Reproducibility
  - Worst case scenario
  - Is physiological relevance important?
  - Highly predictive method?
  - Section 9.6: Conduct in vivo study in swine for made grounds for direct comparison with in vitro results



For additional information on US-EPA methods, visit the following websites:

EPA Bioavailability Guidance
http://www.epa.gov/superfund/bioavailability/guidance.htm

Methods Development Research
http://www.epa.gov/heasd/mdab/mdab.htm

Metals Framework, March 2007 http://www.epa.gov/osa/metalsframework/

EPA's TRW BA Committee:
http://www.epa.gov/superfund/health/contaminants/bioavailability/trw.htm

4 Presentation Pat Rasmussen

Health Canada Santé Canada

Healthy Environments and Consumer Safety Branch

RIVM-TNO Workshop
May 14, 2009

Pat Rasmussen
Exposures and Biomonitoring Division,
Environmental Health Science & Research Bureau
Health Canada

Canada

Research at Health Canada

Federal Contaminated Sites
Accelerated Action Plan

Under FCSAAP, Health Canada is responsible for providing guidance, training and advice on the methods to be used to assess risks at federal contaminated sites. See 2006, 2007, and 2008 Bioaccessibility Workshop Proceedings available at
http://www.cntc.ca

and

NSERC Metals in the Human Environment Research Network

http://www.mithe-rn.org

Research at Health Canada

Gradual acceptance (by some organizations) of site-specific adjustments for oral bioavailability of Pb

Why
- Variability in total Pb: biologically available Pb from site to site
- Better information needed for human health risk assessment

Approaches
- Appropriate use of animal models (swine for oral bioavailability)
- EN-71 Toy Safety Protocol used for consumer products
  - modified for house dust and soil research in background urban settings
- Compares with Drexler and Brattin (2007) for contaminated sites.

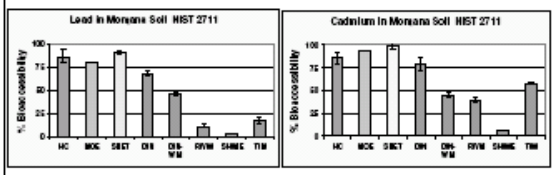
Question #5

- What is your recommendation in order to be able to appoint and/or develop an in vitro method that can be safely applied in human risk assessment of contaminated soils (including Dutch made grounds)?"

Consider simplifying IVD method

- Build on US-EPA experience: single-phase assay (gastric) correlates with *in vivo* results (swine)
- Resolve difficulties maintaining pH
- Establish QA/QC to improve reproducibility
- Goal is a simple, reproducible SOP
- Round robin testing to refine SOP

Similarity of Gastric-Only Assays
(NIST 2711)



Extractions using Gastric Fluid Alone

Combination Gastric Fluid & Intestinal Fluid Extractions

Canadian (HC and MOE) methods plus lab comparison results from Oomen et al (2002)

$$\% \text{ bioaccessible} = \frac{\text{gastric acid-extracted Pb*}}{\text{total Pb in matrix}} \times 100$$

*Variability in the Numerator

- pH
- Method for physical mixing
- Liquid to solid ratio
- Complexing agents and other constituents
- Filtration *vs* centrifugation

$$\% \text{ bioaccessible} = \frac{\text{gastric acid-extracted Pb}}{\text{total Pb in matrix*}} \times 100$$

*Variability in the Denominator

- **Aggressive acid digestion (pg 21) recommended**
- **Otherwise difficulties in obtaining quantitative recoveries**
- **Can result in > 100% bioaccessibility**
- **Can result in bias when comparing different sites**

$$\% \text{ bioaccessible} = \frac{\text{gastric acid-extracted Pb}}{\text{total Pb in matrix}} \times 100$$

Development of SOP (1)

Use diverse Certified Reference Materials (CRMs)

- representing different matrices (include anthropogenic components such as paint and combustion products), and

- wide range of Pb concentrations.

-% recovery monitoring of denominator

-reproducibility of both numerator and denominator*

* enhanced reproducibility due to finely ground, homogeneous CRMs

$$\% \text{ bioaccessible} = \frac{\text{gastric acid-extracted Pb}}{\text{total Pb in matrix}} \times 100$$

Development of SOP (2)

Use representative test samples of Dutch made grounds

- Challenges with alkaline samples

-maintaining gastric pH is key

-Challenges with heterogeneous samples

- particle size fraction(s) needed for risk assessment

- diverse Pb compounds/species in made grounds

-Other parameters (equipment for physical mixing, solid-liquid ratio, filtration or centrifugation)

Sample Preparation: Particle Size

RIVM-TNO Report

- **Milled and sieved to <2 mm**
- **Used directly for Tiny-Tim method**
- **Ground to <500 micron for IVD**

Drexler and Brattin (2007), US-EPA

- **Sieved to 250 micron**

Canadian practice

- **Risk assessors require metal concentrations to be reported for multiple size fractions**
- **Finer fractions more relevant for human exposure assessment (ingestion/inhalation)**
- **Tendency for metals to concentrate in finer fractions**
- **In Canada < 2 mm fraction used in agricultural protocol, not human health risk assessment**

Validation Testing

Goal of validation:

- Show significant positive correlation between *in vitro* and *in vivo* (swine model) test results
- *in vitro* results should parallel *in vivo* results for a set of samples that represent the Dutch made lands
- Test samples selected to represent full range of relevant Pb species/compounds and host matrices

Characterization of made grounds

Selection of test samples for *In Vivo* validation

Sample selection, characterization & PPS ranking well-described in RIVM report.

- Heterogeneity is a big challenge
- Distribution of Pb species in different particle size fractions.
- Variation of particle size distribution amongst samples

Suggestions

- Techniques to identify fine-grained Pb-bearing compounds (XANES, micro XRD, micro XRF, FTIR)
- In "city waste" anthropogenic organic Pb compounds are likely to be present (drying agents, stabilizers, plasticizers) and may be more bioaccessible than Pb associated with soil organic matter.
- Determination of total and bioaccessible Pb in multiple size fractions

Heterogeneity of Made Grounds

Advantages of a simple fast reproducible cheap method for handling the heterogeneity problem:

- Higher throughput of samples is possible
  - representative sampling of each site
  - adequate replicates of each sample
- Run heterogenous samples in triplicate - important for small sample mass (60 mg).
- Assess Pb content of multiple particle size ranges
- Useful for screening - to select a representative set of made ground samples for more expensive *in vivo* tests

Recommendations - Summary

Selection of Test Method

- Simplified gastric method to estimate "bioaccessible metal"
- Aggressive analytical method to determine "total metal"
- QA/QC; round robin testing to fine-tune SOP

Test Method Validation

- Build on US-EPA experience
- Select representative set of made-ground test samples for direct comparison of *in vitro* and *in vivo* results obtained using swine model
- Replicate testing to quantify uncertainty

Additional Characterization to Ensure Test Samples Represent Made Grounds

- Chemical speciation – inorganic and organic Pb compounds
- Consider inhalation pathway (windblown dust) as well as ingestion pathway
- Influence of size fraction on Pb concentration and Pb bioaccessibility
- Particle size distribution

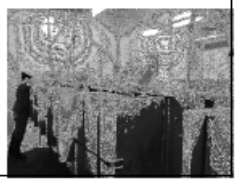
Examples – Current Bioaccessibility Research at HC

Rasmussen & colleagues

5. Presentation Stan Casteel

Using the Swine Model in Bioavailability Assessment

Stan W. Casteel, DVM, PhD
College of Veterinary Medicine
University of Missouri
casteels@missouri.edu



Reducing Uncertainty in Risk Assessment

- Which model minimizes the uncertainty?
  - Proximity to the human condition.
- Which model addresses the issue of biological variability?
- Are studies reproducible within and between labs?
  - **73 vs 75% within lab, 77% vs 74% for inter-laboratory.**

What is Bioavailability?

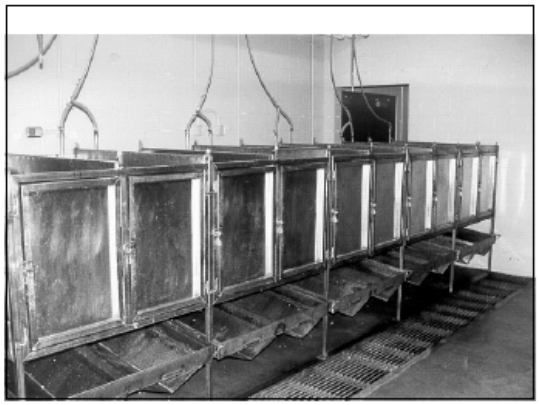
- A term that describes the rate and extent to which the contaminant/nutrient/drug is absorbed and becomes available at the site of action/accumulation.
- Most studies involve measuring contaminant in blood, urine, and tissues to establish estimates of bioavailability.

Why Measure Bioavailability?

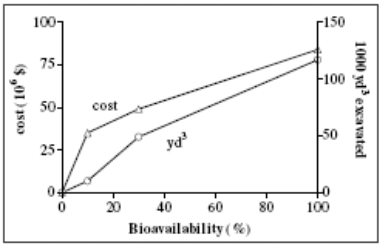
- **Best available science**
  - **Bioavailability estimated in an appropriate animal model closest to the human condition.**
  - **Ethical considerations of using best available science to protect human health.**
- **Site-specific adjustments of default assumptions**
- **Incentives to minimize clean-up costs and disruptions**

Swine Model QA/QC Features

- **Q.A.P.P. (protocol and S.O.P.s)**
- **GLP compliant**
- **Chain of Custody**
- **Double-blind random samples**
- **Peer-Reviewed**



Effect of Bioavailability on Cleanup Costs



Relative Bioavailability (RBA)

- **BASIC CONCEPT:** *Equal absorbed doses cause equal biological responses*
- Ingested Dose(test) • ABA(test) = Ingested Dose(ref) • ABA(ref)

  By definition: RBA = ABA(test)/ABA(reference)

  Thus: RBA = Ingested Dose(ref) / Ingested Dose(test) = ratio of ingested doses that cause equal responses

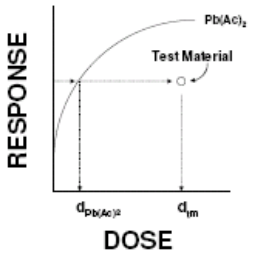
  ≠ ratio of responses at equal doses

Pb Relative Bioavailability (Blood)

Ratio of doses of Pb from $Pb(Ac)_2$ to Pb from soil (soil=tm=test material) producing an equivalent response—

. RBA = $\frac{d_{Pb\,(Ac)_2}}{d_{tm}}$

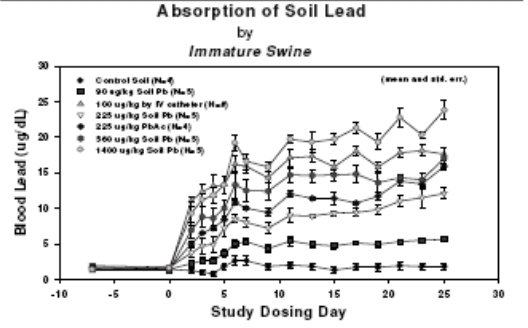
RBA EXAMPLE



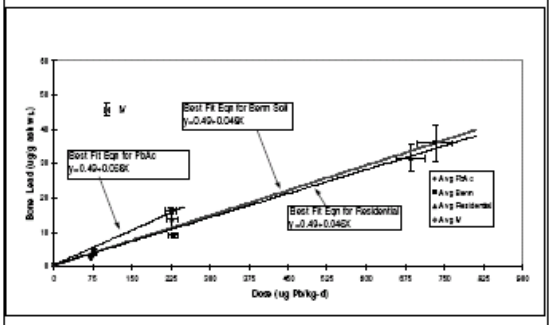
General Study Design

Group	N	Treatment	$Pb(Ac)_2$ or Soil Pb (mg/day)	Pb intake ug/kg/day

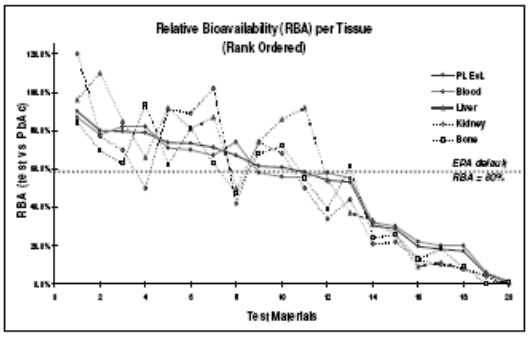
Blood Response As An Index of Bioavailability



Bone As A Measure of Bioavailability



RBA ESTIMATES: Soil-Lead at 20 Sites

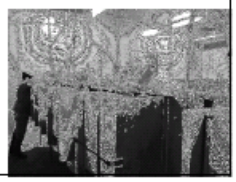


Model Utility

- Versatility--assess metals (As, Cd, Cr, Pb, V), organic compounds and Au-, Pd- and Ag-nanoparticles.
- Uniform age and size of animals—inbred versus outbred.
- Multiple doses reflecting real-world exposure.

EXPOSURE

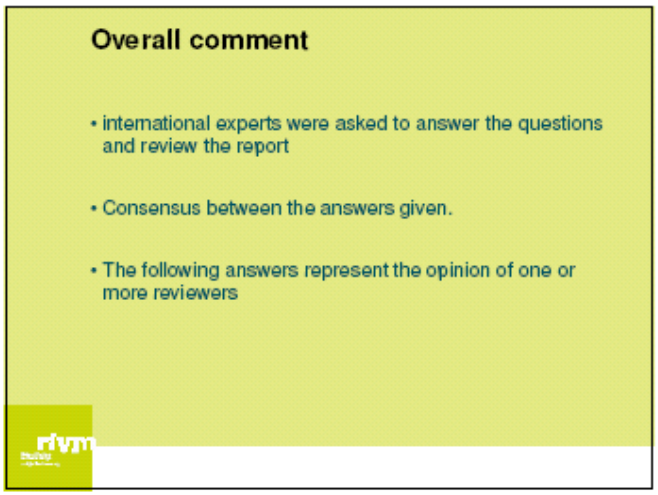
- Exposure methods—single dose or multiple?? Which reflects real-world exposure.
- Reference material (PbAc) provided for comparison (RBA calculations)
- Dosing twice per day (9:00 AM, 3:00 PM)
- Feeding twice per day (11:00 AM, 5:00 PM)

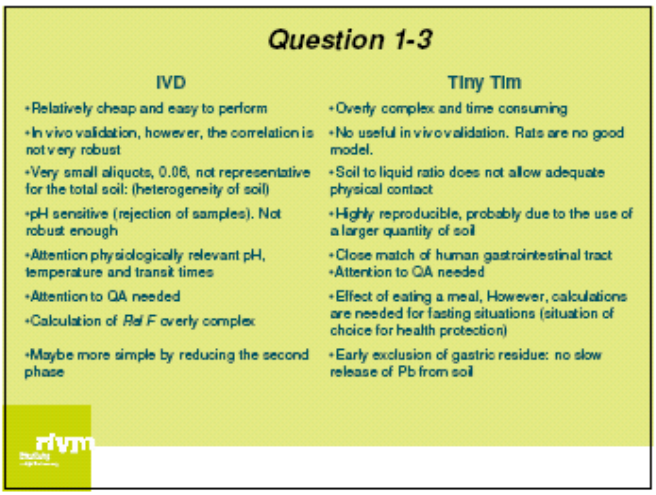


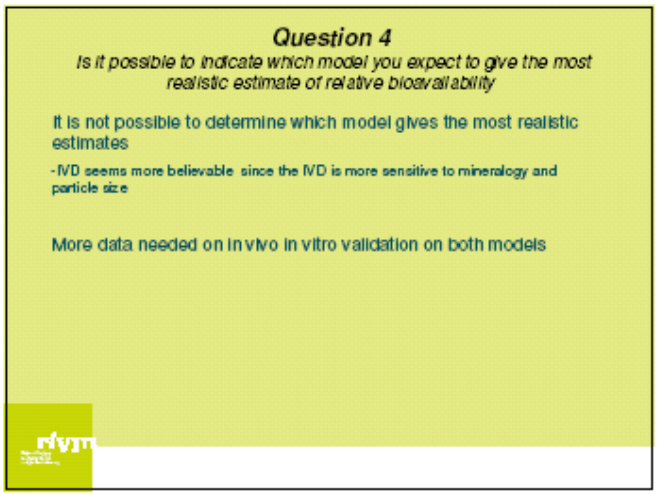
CONCLUSIONS

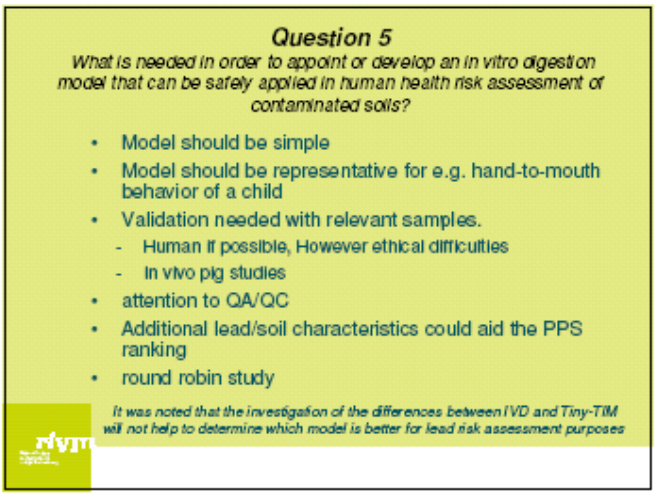
- 20 Superfund soils had Pb RBAs ranging from 86% to 3%
- Bioavailability using animal models are required to validate in vitro and other measurements that estimate bioavailability.
- The swine model provides an estimate of biological response variability that no in vitro system can mimic.
- High profile site-specific bioavailability estimates will continue to demand data from the most appropriate animal model.

6. Werner Hagens: Summary reviews and discussion









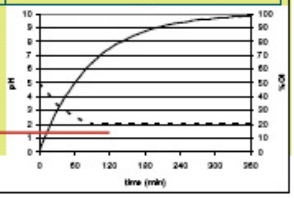
Discussion points

1. Differences Tiny-Tim / IVD
2. PPS analysis
3. In vivo validation
4. IVD: stomach phase only or intestinal phase
5. Particle size

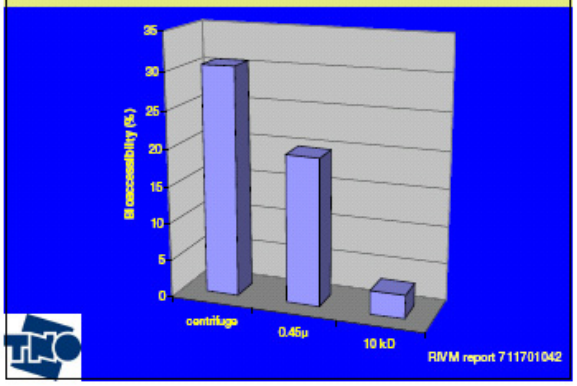
Differences Tiny-TIM and IVD *in vitro* digestion models

Tiny Tim	IVD	Important difference for:
Dynamic model	Static model	pH range (gastric)
5 gram of soil	0.06 gram of soil	+Accuracy +Realistic for hand-to-mouth
Ultrafiltration	Centrifugation	Lead complexes
Fed condition without food	Fasted and/or fed condition	+pH +Complexation of lead to food particles
Relatively complex	Relatively simple	Cost

Gastric phase IVD:
pH 1.5 for 2 hours



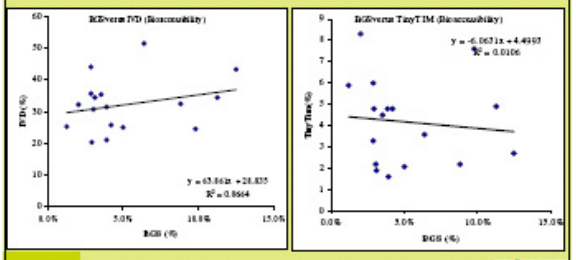
Effect (ultra)filtration on bioaccessibility



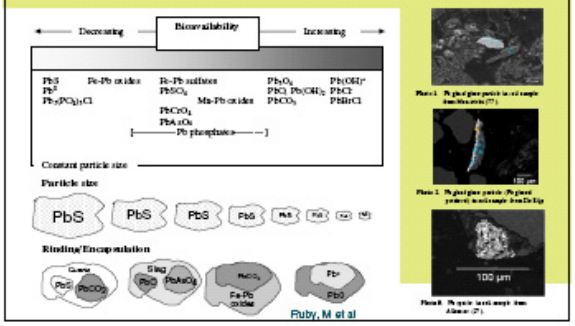
Filtration (Tiny-TIM) / centrifugation (IVD)
Comments by reviewers

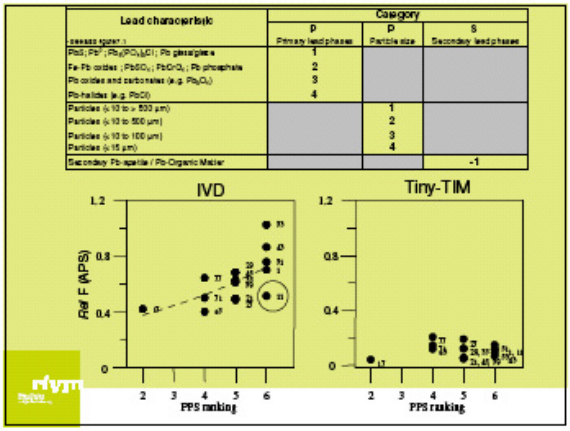
- Early exclusion of residual particles, larger molecules and complexes does not allow the slow release of Pb from soil particles, including organo-Pb complexes, which occur in the stomach. These complexes are soluble in a child's stomach
- Lead complexes > 10kd are excluded
  - Effects reversible lead-to-protein binding are excluded
  - Carrier mediated absorptive transport in the upper small intestine is excluded
  - Children absorb lead to a much greater extent then adults (active transport?)
- Additional studies using juvenile swine would provide valuable information on over/underestimation of both models

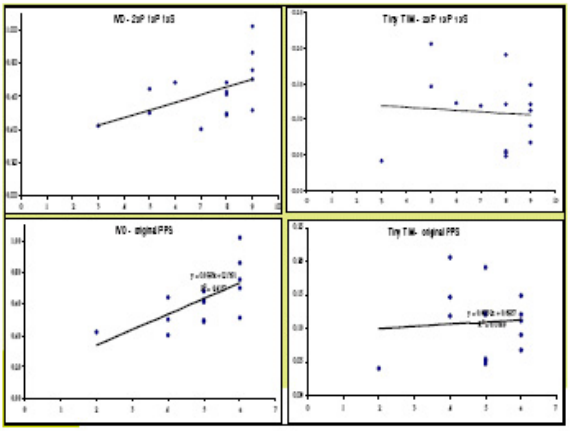
Relation with BGS extraction method



PPS analysis
Primary lead – Particle size – Secondary lead





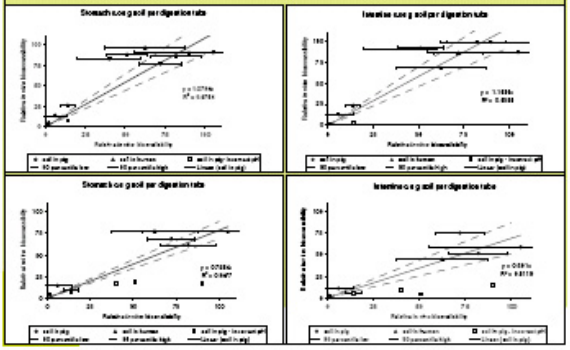


Tiny-Tim and IVD Comparison with in vivo experiments

Soil	physiological state	Lead concentration soil (mg/kg)	bioaccessibility (%) Tiny-TIM	bioaccessibility (%) IVD default: 0.6 gram soil	Oral bioavailability (%) in vivo human Child	Oral bioavailability (%) in vivo human Adult	Oral bioavailability (%) in vivo animals
Flandaca[1]	fasted	614	10 ± 3	60 ± 9			
Oker 11[1]	fasted	5153	4 ± 1	29 ± 2			1.7 ± 0.3 (55% RBA)[1]
Montana 27 10[2]	fasted	5532	2.2 ± 0.5	8.4 ± 2.8 (1.06 gram: 43 ±6)			0.7 ± 0.3[2]
Montana 27 11[2]	fasted	1040	17 ± 3	11 ± 2			
Bunker Hill[3]	fasted	2924	32.5 ± 5	31.8 ± 2 (1:100)		36.2 ± 5[3]	
Bunker Hill[3]	fasted	2924		47.4 ± 2 (1:100); 1.06 gram)			
Bunker Hill[3]	fed	2924	7.0 ± 1.5	33.9 ± 2 (1:100)		2.5 ± 1.1[3]	
Bunker Hill[3]	fed	2924		39.8 ± 2 (1:100); 1.06 gram)			
Dhvani Pair[1]	fed	low	16.4	53.5 ± 12.2	42[1]		
Dhvani Pair[1]	fed	high	69	49.6 ± 14.9	53[1]		
Dhvani Hoodoo[1]	fed	low	12.2	48.2 ± 1.2			
Dhvani Hoodoo[1]	fed	high	64	57.6 ± 2.1			
Spike solution[1]	fasted	low	65.8	53.5 ± 0.1	range 29-71[1]		
Spike solution[1]	fasted	high	56.2	51.4 ± 0.3			
USEPA soils[4]				IVD: 0.06 gram[11]	IVD: 0.6 gram[11]		In vivo pigs (%RBA)[6]
Jasper LL Yard	fasted	4050	nd	66.3 ± 14.3	14.4		90 ± 28
Murray smelter slag	fasted	11500	nd	91.7 ± 3.4	9.8		40 ± 20
Jasper HL mill	fasted	6940	nd	96.4 ± 2.8	51.7		82 ± 16
Midvale slag	fasted	8170	nd	3.1 ± 0.1	5.9		14 ± 4
Butte soil	fasted	8530	nd	23.8 ± 0.7	5.7		14 ± 4
California Gulch Fe/Mn PbO	fasted	4320	nd	87.3 ± 3.4	59.1		105 ± 50
Murray smelter soil	fasted	3200	nd	92.5 ± 2.9	4.1		51 ± 13
HIST palm + soil	fasted	8350	nd	86.3 ± 2.3	75.3		72 ± 14
Galenae rriched soil	fasted	11200	nd	1.7 ± 0.2	2.9		1 ± 1
California Gulch Tailings	fasted	1270	nd	12.9 ± 0.4	11.1		6 ± 6
Bunker Hill	fasted	2924	32.5 ± 5	68.6 ± 6.2	44.7 ± 7.6		62 ± 25 (%RBA), human

Literature values of lead from food and solution

IVD: Stomach and intestinal bioaccessibility



IVD: 0.6 gram	lead in soil (mg/kg)	Relative bioaccessibility (%) Stomach	Relative bioaccessibility (%) Intestine	Relative bioavailability (%) In vivo pigs (%RBA)[6]
Jasper LL Yard	4050	17.1	14.4	90 ± 28
Murray smelter slag	11500	16.8	9.8	40 ± 20
Jasper HL mill	6940	61.3	51.7	82 ± 16
Midvale slag	8170	8	5.9	14 ± 4
Butte soil	8530	9.4	5.7	14 ± 4
California Gulch Fe/Mn PbO	4320	77	59.1	105 ± 50
Murray smelter soil	3200	19.6	4.1	51 ± 13
HIST palm + soil	8350	68.4	75.3	72 ± 14
Galenae rriched soil	11200	4.3	2.9	1 ± 1
California Gulch Tailings	1270	14.4	11.1	6 ± 6
Bunker Hill	2924	77.8 ± 2.4	44.7 ± 7.6	62 ± 25 (%RBA), human
IVD: 0.06 gram	**lead in soil (mg/kg)**	**Stomach**	**Inseine**	**In vivo pigs (%RBA)[6]**
Jasper LL Yard	4050	89.7 ± 8.9	92.3 ± 14.3	90 ± 28
Murray smelter slag	11500	82.5 ± 2.7	91.7 ± 3.4	40 ± 20
Jasper HL mill	6940	85.7 ± 2.5	96.4 ± 2.8	82 ± 16
Midvale slag	8170	7.0 ± 0.5	3.1 ± 0.1	14 ± 4
Butte soil	8530	25.9 ± 0.9	23.8 ± 0.7	14 ± 4
California Gulch Fe/Mn PbO	4320	90.4 ± 2.6	87.3 ± 3.4	105 ± 50
Murray smelter soil	3200	87.3 ± 2.6	92.5 ± 2.9	51 ± 13
HIST palm + soil	8350	76.3 ± 2.4	86.2 ± 2.3	72 ± 14
Galenae rriched soil	11200	4.0 ± 0.7	1.7 ± 0.2	1 ± 1
California Gulch Tailings	1270	13.3 ± 0.5	12.9 ± 0.4	6 ± 6
Bunker Hill	2924	95.8 ± 0.6	68.6 ± 6.2	62 ± 25 (%RBA), human

RIVM report 711701042/2008 : How can information on oral bioavailability improve human health risk assessment for lead-contaminated soils? Implementation and scientific basis
A.G. Oomen, E.F.A. Brandon, F.A. Swartjes, A.J.A.M. Sips

The particle size

- < 2 mm: Standard practice in soil/agricultural science
- < 250 µm: Normally used in human health studies

In general, the smaller particles stick top children's hands

The made ground samples were " smashed" to 2 mm

7. Final discussion

Workshop *In vitro* modelling of bioaccessibility of lead in soil

Final discussion using propositions

Martine Bakker and Ad Ragas

Proposition 1

- The costs of *in vitro* methods for measuring bioaccessibility are low compared to those of remediation
- To develop a regulatory framework for the bioavailability of substances in soil it is not necessary to have a simple, cheap model with high sample throughput.

Proposition 2

- Local authorities may decide for themselves which model they use to measure bioavailability: IVD, TIM or USEPA model

Proposition 3

- A generic, conservative, default bioavailability of lead in soil (e.g. 0.6) should be used in the regulations instead of values generated with a model.

0.6

default

Proposition 4

- The selection of the best *in vitro* model should be done on how it correlates with *in vivo* data

Proposition 5

- For both IVD and TIM more validation is needed
- This should be done by measuring the bioavailability of lead in made grounds in young swine

Proposition 6

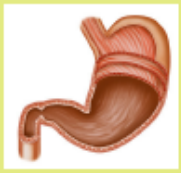
- IVD likely overestimates bioaccessibility
- TIM likely underestimates bioaccessibility
- But that is not a problem as long as it correlates with *in vivo* data

Proposition 7

- IVD could be used but
  - a higher solid/liquid ratio should be applied
  - The pH should be better maintained

Proposition 8

- TIM could be used but
  - A lower solid/liquid ratio in the gastric phase is needed



Proposition 9

- The appointed model should be evaluated for other substances than lead

Proposition 10

- For risk assessment of metals in soil, other exposure routes (ingestion and inhalation of (house) dust) should be included

Proposition 11

- All the different 'bioaccessibility terms' (absolute, relative to PbAc, relative to food) make the discussion of this subject overly complex
- In the future we'll use the term **absolute bioaccessibility** as much as possible

Appendix 4 List of workshop participants

Organisation
Dr. Martine Bakker, RIVM
Dr.Werner Hagens, RIVM
Dr. Rob Havenaar, TNO
Dr. Mans Minekus, TNO
Dr. Ad Ragas, Radboud University Nijmegen (chair)

External reviewers
Dr. Mike Beringer: US-Environmental Protection Agency
Dr. Karen Bradham: US-Environmental Protection Agency
Dr. Stan Casteel, DVM, DABVT: University of Missouri
Dr. Pat Rasmussen: Health Canada
Dr. Sohel Saikat: British Health Protection Agency

Others
Dr. Bert-Jan Baars, RIVM
Dr. Marc Cave, British Geological Survey
Drs. Marlies ten Hove, Technical Committee on Soil Protection
Ir. Johannes Lijzen, RIVM
Dr. Agnes Oomen, RIVM
Ir. Joke Wezenbeek, Grontmij/Ministry of Housing, Spatial Planning and the Environment
Dr. Tom van de Wielen, Ghent University
Dr. Joanna Wragg, British Geological Survey

Absent
Ing. N.J. Molenaar, Ministry of Housing, Spatial Planning and the Environment
Ir. A.B. Roeloffzen, DCMR Milieudienst Rijnmond
Ir. N. Walraven, Geoconnect

RIVM

National Institute
for Public Health
and the Environment

P.O. Box 1
3720 BA Bilthoven
The Netherlands
www.rivm.com