A meta-analysis to correlate lead bioavailability and bioaccessibility and predict lead bioavailability

Zhaomin Dong[a,b], Kaihong Yan[a,b], Yanju Liu[a,b], Ravi Naidu*[,a,b], Luchun Duan[a,b], Ayanka Wijayawardena[a,b], Kirk T. Semple[c], Mohammad Mahmudur Rahman[a,b]

[a] *Global Center for Environmental Research (GCER), The Faculty of Science and Information Technology, University of Newcastle, University Drive, Callaghan, NSW 2308, Australia*

[b] *Cooperative Research Centre for Contamination Assessment and Remediation of the Environment (CRC CARE), Mawson Lakes, SA 5095, Australia*

[c] *Lancaster Environment Centre, Lancaster University, LA1 4YQ Lancaster, United Kingdom*

*corresponding author: Ravi Naidu;
ATC Building, Global Center for Environmental Remediation, Faculty of Science and Information Technology, University of Newcastle, Callaghan, NSW 2308
Email: ravi.naidu@newcastle.edu.au
Phone : +61 02 4913 8705

Contents

Abstract

Defining the precise clean-up goals for lead (Pb) contaminated sites requires site-specific information on relative bioavailability data (RBA). While *in vivo* measurement is reliable but resource insensitive, *in vitro* approaches promise to provide high-throughput RBA predictions. One challenge on using *in vitro* bioaccessibility (BAc) to predict *in vivo* RBA is how to minimize the heterogeneities associated with *in vivo*-*in vitro* correlations (IVIVCs) stemming from various biomarkers (kidney, blood, liver, urinary and femur), *in vitro* approaches and studies. In this study, 252 paired RBA-BAc data were retrieved from 9 publications, and then a Bayesian hierarchical model was implemented to address these random effects. A generic linear model (RBA (%) = (0.87 ± 0.16) × BAc + (4.70 ± 2.47)) of the IVIVCs was identified. While the differences of the IVIVCs amongst the *in vitro* approaches were significant, the differences amongst biomarkers were relatively small. The established IVIVCs were then applied to predict Pb RBA of which an overall Pb RBA estimation was 0.49 ± 0.25. In particular the RBA in the residential land was the highest (0.58 ± 0.19), followed by house dust (0.46 ± 0.20) and mining/smelting soils (0.45 ± 0.31). This is a new attempt to: firstly, use a meta-analysis to correlate Pb RBA and BAc; and secondly, estimate Pb RBA in relation to soil types.

KEY WORDS： lead, bioavailability, bioaccessibility, meta-analysis, soil

1. Introduction

Lead (Pb) exposure in children is of worldwide concern, and soil and house dust have been considered a significant exposure pathway because Pb may be directly ingested and indirectly absorbed (Levin et al. 2008; Mielke and Reagan 1998). Incorporating Pb bioavailability, i.e. the fraction of an ingested dose that crosses the gastrointestinal epithelium and becomes available for distribution to internal tissues and organs, into human health and ecological risk assessment is increasingly acknowledged (Naidu et al. 2015; Ortega Calvo et al. 2015). The U.S. Environmental Protection Agency (EPA) suggests an overall relative bioavailability (RBA) in soil with reference to water and food is about 60% (U.S. EPA 2007). However, many studies have reported that Pb bioavailability varies extensively with the type of soils (Casteel et al. 2006; Li H et al. 2014; Li et al. 2015; Wijayawardena et al. 2015). For example, Li et al. (2015) and Li H et al. (2014) reported that Pb RBA ranged from 51% to 60% for farming soils, 31% to 84% for smelter soils, 7% to 26% for mining soils, and 29% to 60% for house dusts, respectively. Since Pb bioavailability may vary among soil types (Oliver et al. 1999; U.S. EPA 2007; Wijayawardena et al. 2015), it is necessary to use type-specific RBA to define the accurate clean-up goals for specific contaminated sites.

In vivo and *in vitro* approaches, are commonly employed to estimate Pb RBA. Although *in vivo* measurements can directly provide reliable information on Pb RBA (Casteel et al. 2006; Hettiarachchi et al. 2003), only limited information is available because it is time-consuming and expensive. Considering *in vitro* measurements are rapid, economical and reproducible, *in vitro* bioaccessibility (BAc) (Ruby et al. 1993) approaches promise to provide high-throughput RBA predictions if the correlation between *in vivo* RBA and *in vitro* BAc (IVIVC) can be validated. A challenge when using *in vitro* BAc to predict *in vivo* RBA is how to minimize the heterogeneities of IVIVCs. For example, five *in vitro* methods, namely the Relative Bioavailability Leaching Procedure (RBALP), unified BioAccessibility Research Group Europe (BARGE) method (UBM), Solubility Bioaccessibility Research Consortium assay (SBRC), Physiologically Based Extraction Test (PBET), and the In Vitro digestion model (RIVM), have been widely utilized for determining *in vitro* bioaccessibility (BAc) (Casteel et al. 2006; Dodd et al. 2013; Juhasz et al. 2009; Juhasz et al. 2013; Kesteren et al. 2014; Ruby et al. 1996). The IVIVCs based on each *in vitro* method have been previously reported (Casteel et al. 2006; Denys et al. 2007;

Denys et al. 2012; Deshommes et al. 2012; Kesteren et al. 2014; Li H et al. 2014; Schroder et al. 2004; Smith et al. 2011): results from U.S. EPA have documented that Pb RBA can be reliably estimated using RBALP assay and reported a regression equation (*RBA=0.878×BAc-0.028*) relating *in vitro* BAc to *in vivo* RBA (U.S. EPA 2007). Another study indicated experimental BAc based on SBRC is higher than the observed RBA when using rat model (Li H et al. 2014). A closer examination of these statistical relationships shows uncertainties do exist as exemplified by a fitted coefficient which has been reported as ranging widely from 0.39~1.87 (Deshommes et al. 2012).

These uncertainties mostly stem from various *in vitro* measurements, different biomarkers, inter-laboratory variances, and model selections. For example, Yan et al. (2015) measured BAc on the same soils using different *in vitro* approaches. Further analysis showed that Pb BAc based on the RBALP and SRBC, RIVM models were comparable, while the slopes between RBALP and UBM can be up to 1.21 (Yan et al. 2015). Meanwhile, *in vivo* RBAs based on different biomarkers, including blood area under curve (AUC), liver, kidney and femur do not agree precisely with each other (Li H et al. 2014; U.S. EPA 2007). By integrating all the raw data, a meta-analysis promises to: firstly, determine the heterogeneities of IVIVCs, and secondly, to produce a comprehensive extrapolation (Axelrad et al. 2007; Whitehead 2002).

In this study, paired BAc-RBA data, type-specific BAc and RBA data were retrieved from published reports. The objective of this study was to estimate Pb RBA with reference to soil types. This was achieved via two steps using: 1) meta-analysis to establish the IVIVCs and 2) established IVIVCs to predict RBA. The study presented here provides Pb site-specific RBA estimation to assist in Pb risk assessment and management.

2. Materials and Methods

2.1. Process for estimating site-specific Pb RBA.

As shown in Figure 1, the procedure for estimating type-specific Pb RBA consisted of three steps. In the first step, three types of data (paired BAc-RBA, type-specific RBA, type-specific BAc) were collected. Using ‘lead’ & ‘bioavailability’ & ‘bioaccessibility’ as the keywords, an extensive literature search (for analyses published between 1950 and 2015) was done and checked by the two co-authors (databases included Pubmed, Web of Science, Medline). The BAc-RBA paired data based on IVG and PBET were not considered in this study because no significant correlations were reported between such

two methods and other *in vitro* approaches (Yan et al. 2015). Meanwhile, the BAc data above 100% were omitted. Finally, the BAc measurements included four *in vitro* methods, RBALP, SBRC, UBM and RIVM. Two different solid: liquid ratio (1:37.5 and 1: 375) were used in the RIVM approaches. Five biomarkers, namely blood area under curve (AUC), liver, kidney, femur and urinary were also selected for indicating RBA. Since *in vitro* experimental parameters (pH, solid: liquid ratio and other factors) will influence the BAc measurements (Ryan et al. 2004), the procedures for each *in vitro* methods were identical in the pilot study (Yan et al. 2015). In the second step, the IVIVCs were developed by using a Bayesian hierarchical random-effects model and paired BAc-RBA data, and later the developed IVIVCs were used to convert type-specific BAc data into predicted RBA. In step 3, the RBA data, including the predicted RBA and reported RBA, was classified according to environmental media types. These media types were clustered into four categories: house dust, mining and smelting sites, residential land and others. The 'others' here included the soil samples from shooting range, incinerator, landfill, gasworks, etc. The BAc data were omitted when both the BAc and RBA data became available for the same soil samples. Table 1 summarizes the data collection, and all the raw data are available in Supplemental Material (SM) Tables S1, S2 and S3. It should be noted that some data in Table S1 and S3 were shared.

2.2. Meta-analysis.

When raw data are available a meta-analysis using a hierarchical approach is possible and this strategy can be used to address the effects from various factors (Whitehead 2002). In this study, a hierarchical random-effects model was employed, which is commonly utilized to combine relevant information from different studies (Axelrad et al. 2007). Here the heterogeneities were identified, consisting of three types, (i) individual effects, to represent the treatment differences from inter-lab, operations and other factors (ii) *in vitro* method effects and (iii) biomarker effects. This treatment can distinguish between heterogeneities to establish a 'real' link between the independent and dependent variables (Axelrad et al. 2007).

According to part 2 (meta-analysis) in Figure 1, the measured RBA (y_{ij}) was assumed to be of normal distribution with expected RBA (θ_{ij}) and individual variance (s_{ij}^2) (Whitehead 2002):

$$y_{ij} \sim N(\theta_{ij}, s_{ij}^2) \quad (1)$$

Here subscripts *i* and *j* represent different methods and endpoints, respectively. The variance of RBA has been observed to be a function of increasing RBA, which is referred to as heteroscedasticity (U.S. EPA 2007). To handle the heteroscedasticity, an option termed as an "external'' variance, which is aimed to establish the relationship between variance and RBA has been recommended by U.S. EPA (2007) and adopted in this study. The detailed descriptions for variance estimation (s_{ij}^2) are provided in the SM.

Using expected RBA (θ_{ij}) in Equation (1), the remaining heterogeneities can be described using Equations (2) and (3)

$$\theta_{ij} \sim N(\mu_{ij}, \delta^2) \quad (2)$$

$$\mu_{ij} = \beta + \alpha_0 \times x_{ij} + (\gamma_i + \lambda_j) \times x_{ij} \quad (3)$$

where the expected RBA (θ_{ij}) was assumed to be the normal distribution with 'real' RBA (μ_{ij}) and population variance δ^2, which accounts for the model residuals. A linear algorithm, as illustrated in Equation (3), has been applied to link the 'real' RBA (μ_{ij}) and BAc (x_{ij}). β and α_0 are the intercept and overall coefficient, respectively. (Yan et al. 2015). The remaining coefficients γ and λ account for the random effects from *in vitro* approaches and endpoints. Since data were collected based on different *in vitro* methods, BAc (x_{ij}) was firstly adjusted by using the established correlations among *in vitro* methods (SM Table S4) (Yan et al. 2015). Both the raw data and adjusted BAc data are provided in SM Table S1. Similarly, λ has been utilized to represent the endpoint random effects.

The probabilistic and deterministic methods are both employed for mathematical modelling and parameter optimization. In this study all the objective parameters were fitted via Bayesian inference, a commonly used probability method. Compared to deterministic methods, the advantages of the Bayesian inference have been well summarized: interval estimation, the use of prior information and constraint test for parameters (Xu et al. 2006). All the procedures were simulated by *matbugs*, a Matlab (version 2012b) interface to WinBUGS that can execute Bayesian inference. Three Monte Carlo Markov Chains (MCMC) were simultaneously run until convergence was achieved. A Gibbs sampler was employed to obtain the parameters in each model. The pseudo-code for the simulation is provided in SM.

2.3. Model comparisons.

The main objective of the curve-fitting is to find a mathematical model that fits the collected data reasonably well. However, the model itself neither has a mechanistic basis nor biological meaning. It is generally not appropriate to choose the form of the dose-response models based on only one function. In fact it is prudent to make the choice based on the weight of observations across many different regressions. Two alternative non-linear models (two parameters exponential, Equation (4); three parameters exponential, and Equation (5)) were also evaluated in this study (U.S. EPA 2007):

$$\mu_{ij} = \beta + \alpha_0 \times \exp(x_{ij}) + (\gamma_i + \lambda_j) \times \exp(x_{ij}) \quad (4)$$

$$\mu_{ij} = \beta + \alpha_0 \times \exp(c \times x_{ij}) + (\gamma_i + \lambda_j) \times \exp(c \times x_{ij}) \quad (5)$$

2.4. Robustness analysis.

A Jackknife resampling approach was employed to assess data bias in this study. The Jackknife estimator of a parameter is achieved by systematically leaving out each observation from a dataset and calculating the estimate, which is commonly used to estimate the bias and the standard error of statistics (Wu 1986). According to the various *in vitro* methods, studies and biomarkers, the raw data in SM Table S1 was classified into 29 groups. With 1-deleted group in turn, the meta-analysis was re-run to obtain the objective parameters.

2.5. Type-specific bioavailability estimations.

The established IVIVCs were used to convert BAc data into RBA. Both the predicted RBA and collected RBA data (Figure 1, step 3) were applied to statistically summarize RBA according to its type.

3. Results

3.1. Data preparations and descriptions for meta-analysis.

As summarized in Table 1, 252 paired RBA-BAc data points were collected from 9 published reports. It is worth noting that the collected BAc data only included the data from gastric phase: the data from intestinal phase were excluded given the BAc under intestinal phase was always reported with BAc data for gastric phase for the same material. Including the BAc data from intestinal phase for the same material would outweigh this material's impact. All the collected data are shown in Figure 2 and SM Figure S1.

The number of collected data based on RBALP (Bannon et al. 2009; U.S. EPA 2007) was 104 (Table 1, Figure 2 legend cycle) and was the largest of all the methods considered for this study. Denys et al. (2012) have reported *in vivo* data and *in vitro* UBM data for 16 soils, however, the concentrations of 6 soils were beyond the linear range of the *in vitro* correlations (Yan et al. 2015). Consequently only 10 of the 16 soils have been included in this study, and the biomarker urine was only utilized in Denys et al. (2012). Some RBA data for the UBM (Wragg et al. 2011) and RIVM *in vitro* methods were derived from an *in vivo* study conducted by U.S. EPA (2007).

The mean BAc in the *in vitro* methods varied as follows: SBRC (69%)> RIVM (62%-69%) >RBALP (64%) > UBM (37%). While the reported mean BAc based on UBM was significantly lower than the other three groups (Mann-Whitney test, $p<0.001$), no significant differences emerged among the other groups. While the ratio of RBA/BAc-$_{RBALP}$ and RBA/BAc-$_{UBM}$ was slightly higher than 1, the ratio of RBA/BAc-$_{SBRC}$ was approximately 0.57 (Table 1).

The raw RBA-BAc data used for meta-analysis have been presented in Figure 2. The size, color and style represent the variance, biomarker and *in vitro* method, respectively. It indicated that the size of the points with higher RBA was significantly larger than that with lower RBA, suggesting there may be a significant positive link between the variance and reported RBA. This linkage has been examined (Figures S2 and S3), and a function between variance (s^2) and RBA (x) was done as follows:

$$Ln(s^2) = (1.65 \pm 0.33) \times Ln(x) - (4.10 \pm 0.09) \quad (6)$$

This current study showed that the random effect of variance estimations using femur is slightly higher than for the other biomarkers. In particular, the random effects of variances from femur samples were positive (0.65). Conversely, blood AUC (-0.28), liver (-0.26), kidney (-0.17), and urine (0.0036) yielded negative random effects. The estimated variances were applied to Equation (1) for further meta-analysis.

3.2. Meta-analysis, established IVIVCs and model comparisons.

As stated above, the estimated variance (Equation (6)) alongside raw data were employed to help execute the hierarchical random effects model (Equations 1-3). The Gelman-Rubin (G-R) diagnostic method tested the convergence of the Monte Carlo sampling. By running three

parallel chains at any random start points, the results of the three MCMC chains should be similar. The idea of G-R test is that if the simulated MCMC has reached convergence, the within-run variation should be roughly equal to between-run variation (Xu et al. 2006).The simulation was considered to be converged when the Corrected Scale Reduction Factors (*R*) was < 1.20 (Xu et al. 2006). In this study, the reverse sampling simulations converged to $R <$ 1.10 for all population parameters.

A random chain was chosen for shaping the population posterior distribution to obtain objective parameters in Equation 3, and therefore the IVIVC was developed as Equation (7).

$$RBA(\%) = (0.87 \pm 0.16) \times BAc + (4.70 \pm 2.47) + g(method, endpo\,\mathrm{int}) \times BAc \quad (7)$$

where function *g* represents the random effect from various methods and endpoints. As shown in Table 2 and Figure S4, the random effect for RIVM (1:37.5 Solid/Liquid ratio, termed as S/L ratio) was 0.32, which was the highest ratio. This was followed by RBALP (0.075), UBM (-0.018), and RIVM (1:375 S/L ratio) (-0.038) and SBRC (-0.37). Compared to the square of mean random effect for *in vitro* methods (0.25), this square of mean for biomarkers was much lower (0.0069). In particular, the random effects for liver was the highest (0.039), followed by blood AUC (0.018), urine (0.017), kidney (-0.018) and femur (-0.067).

The Jackknife re-sampling approach was employed to address the data bias from each group. As a result, the means of re-simulated intercept and slope were estimated to be 4.77 ± 0.024 and 0.87 ± 0.0014, respectively (SM Figure S5). This low variation of coefficient (CV) for the intercept and slope (0.49% for intercept and 0.16% for slope) indicated all the groups may exert a limited influence on the model simulations.

The alternative exponential models may potentially fit the dose response curve because Figure 2 suggested a higher slope for the higher RBA (the right side). Thus, two alternative non-linear models (Equations (4) and (5)) were also employed for model comparisons. As seen in Table S5, the two exponential models fit slightly better than the linear model (lower deviance information criterion). However, the improvement was below 1% and this was not enough to conclude that a non-linear fit is preferable to a linear model. Furthermore when using the power model to link RBA and BAc, the predicted RBA was not convergent when BAc was high. Considering the linear model has the most sophisticated theory and judgement system, the linear model was employed in the present study. As more data

become available in the future, the relationship between BAc and RBA will be reassessed and the model selection will be reviewed as necessary.

3.3. Type-specific Pb bioavailability estimations.

A total of 98 datasets for type-specific RBA and 105 datasets for type-specific BAc were retrieved (Table 1 and SM Table S2). In particular, 43, 3, 31 and 28 data were collected based on RBALP, RIVM, SBRC and UBM *in vitro* methods, respectively. RBA rank across four soil types differed from the rank concerning BAc. For example, the RBA for the residential land was the highest (62%), while the BAc for this type of soils (52%) was lower than house dust (57%).

Using the established IVIVC and parameters in Table 2, the RBA for Pb was estimated according to different types of soils. No significant relationship (p=0.13) was observed for the Pb concentration (log-transformation) and RBA. The predicted RBA from BAc was 46 ± 18 %, while the published RBA was 52 ± 31%. Using the Mann-Whitney test, no significant difference (p=0.32) was found between the two types of RBA, which may confirm that the prediction based on BAc was comparable to the RBA based on *in vivo* studies. The boxplots for different soils across the data source are shown in Figure 3. An overall RBA was estimated to be 49 ± 25% (median: 47%), and the RBA for different types of soils are in the 45%-60% range. In particular, the RBA for the residential land was the highest (58 ± 19%, median: 58%), followed by house dust (46 ± 20%, median: 44%) and mining/smelting soils (45± 31%, median: 36%). The RBA for other soil types are 45 ± 24% (median: 45%). Meanwhile, the median RBA for the residential land, house dust, mining/smelting soils and other types were 58%, 44%, 36% and 45%, respectively. Various mining and smelting types may result in the high CV and differences between mean estimation (45%) and median estimation (36%) of the mining/smelting's RBA. Significant differences were found between residential land and house dust (M-W U test, p<0.05), residential land and other soils (M-W U test, p<0.05).

4. Discussion

4.1. Implications of RBA-dependent variance.

Usually, the ordinary linear squares regression (OLS) is employed to correlate RBA and BAc (Deshommes et al. 2012; Li H et al. 2014). With the OLS regression, the variances of the responses should be independent of the RBA (termed as homoscedasticity). However, Equation (6) indicated that this assumption is generally not satisfied, at least in this case.

Casteel et al. (2006) have similarly estimated the link between RBA and variance. Furthermore the coefficient and intercept for different biomarkers have been reported with ranges of 1.55 to 2.10 and -2.60 to -1.32, respectively (Casteel et al. 2006), while the 95% confidential interval (CI) for the coefficient and intercept in our study were estimated to be 0.96 to 2.30 and -4.29 to -3.92, respectively. The slope (1.65) we simulated here is within previous range, while the intercept (-4.10) was lower. However, the 95% CI for slope in this study (0.96 to 2.30) was wider than previous study and this may be due to heterogeneities from studies and biomarkers.

Thus, considering the 'RBA-dependent variances', the weighted linear squares (WLS) regression has been recommended (Casteel et al. 2006; U.S. EPA 2007). In this study, according to the various studies and *in vitro* methods, we have clustered all data into 8 groups (SM Table S6), and both the WLS and OLS have been applied to re-examine the data from each group to compare the two regressions. The mean RBA data for each group, i.e. the average bioavailability of the multiple biomarkers (if available) were applied, since differences among biomarkers were insignificant as demonstrated in this study. As shown in Table S6, 5 of the 8 functions were significant for the OLS and WLS approaches. The slopes were in the 0.61~1.08 range for OLS, while the values for WLS were all below 1. Particularly, in the estimate when using WLS, the highest coefficient was found for RBALP (0.84), followed by UBM (0.80), SBRC (0.44 - 0.78) and RIVM with 1:375 S/L ratio (0.70).

It is noted that the coefficients based on the WLS were all below the values under OLS (paired t test, p=0.008). For example, the simulated coefficients decreased to 18%, 15%, 16% and 26% for the raw data collected from U.S. EPA (2007), Li H et al. (2014), Denys et al. (2012) and Oomen et al. (2006), respectively. This may be explained by the difference in RBA/BAc slopes at the lower and higher BAc. As shown, the RBA/BAc slope when BAc is higher (>50%) (Figure 2) is steeper than the RBA/BAc slope when BAc is lower (<50%). Additionally, when the WLS was employed, this points to the BAc below 50% being weighted more than the BAc above 50% (since their variance was relatively low as stated in Equation (6)), which resulted in a lower simulated slope (relative to OLS)). However, to the best of our knowledge, there is no explicit explanation for the different slopes between the higher and lower RBA. This may be due to the fact that when BAc is higher, the *in vitro* methods are not able to extract the proportionate bioaccessible fraction. This has been

partly proven by the fact that the extraction abilities of the UBM and RIVM (1:37.5 S/L ratio) methods are limited when the bioaccessible fraction is high (Yan et al. 2015). Consequently, if extractability is limited when the bioaccessible fraction increases, the ratio between RBA and BAc may increase (as shown on right side of Figure 2). Therefore, our study suggests the IVIVCs using traditional OLS may need to be adjusted if the measured BAc crosses from a low value to a high BAc.

While the OLS assumed the variance is independent, the WLS approach considers the magnitude of variance would increase with an increase in dose/response to overcome this 'heteroscedasticity'. A non-parametric method has been conducted in a previous study (Denys et al. 2012): this method used a repeated medians approach which specifically does not make any assumptions that the error is associated with the Y axis or that the residuals should be normally distributed. In this study, the strategy to treat the 'heteroscedasticity' is to use a normal distribution to account for variance. Meanwhile, the hierarchical model used in this study may be more informative, since it is also capable of separating the random effect from *in vitro* approaches and biomarkers (Equation 7 and Table 2).

4.2. Comparisons of IVIVCs.

Although previous *in vivo* RBA from different biomarkers results do not agree precisely (Casteel et al. 2006; U.S. EPA 2007) with each other, and we believe such differences are emerging from measurement and intra-species differences. Theoretically, using tissue concentration and blood concentration to estimate RBA and absolute bioavailability (ABA) should result in the same estimates. This study also demonstrated the differences among the biomarkers may be ignorable (Table 2), a finding that agrees with a recent study conducted on Arsenic (Li J et al. 2016).

In this study, the generic coefficient for IVIVC was estimated to be 0.87 (95% CI: 0.55~1.19, Equation (7)). Although we have used the prior information to minimize the impact from *in vitro* methods (Table S4), the coefficient for RIVM (1:37.5 S/L ratio, 1.19), RBALP (0.95), UBM (0.85), RIVM (1:375 S/L ratio, 0.84) and SBRC (0.52) were considerably different. Denys et al. (2012) suggested the slope should be between 0.8 and 1.2. In this case, while the slopes based on RBALP (0.95), RIVM (1:37.5 S/L ratio, 1.19), UBM (0.85) and RIVM (1:375 S/L ratio, 0.84) were within this range, the only slopes based on SBRC (0.52) were slightly lower than the baselines. Surprisingly, previous studies indicated that the procedures

for RBALP and SBRC were identical (Yan et al. 2015), however, huge differences in the coefficient for RBALP (0.95) and SBRC (0.52) were observed in this study. The RBALP data largely derived from U.S. EPA (2007) yields a slope of 0.88 (Table 3), which is close to the estimation in the present study (0.95). Regarding another issue, most SBRC data originated from (Li H et al. 2014), and these two studies suggested low coefficient values (0.40 - 0.61, Table 3). This is a visual case explaining why it may be necessary to employ meta-analysis to consider research from inter-labs in order to achieve a reasonable result. The regression for SBRC and RBALP should not differ substantially from each other, while inter-study variances result in huge heterogeneities. More *in vitro-in vivo* experiments in the future may confirm that IVIVCs based on the two *in vitro* approaches do not differ, however, this judgement is not validated in current experiments.

Previous IVIVCs based on the RBALP, UBM, SBRC and RIVM are summarized in Table 3. With the exception of Oomen et al. (2006), all the RBA/BAc slopes from other studies were below 1, which may indicate that per RBA change is more conservative than per BAc change. On another issues, the intercepts among the IVIVCs were reported as having a large range, from -0.028 to 30.21 (Table 3), while Denys et al. (2012) asserted the intercept should not be significantly different from 0. The difference between previous IVIVCs and developed IVIVCs in this study is we have integrated these reported IVIVCs to address the random effects. In this way, a less biased IVIVC is expected (Equation 7). It is in the meantime convenient to convert the BAc data into RBA data by choosing the appropriate parameters for the selected endpoints and *in vitro* approaches (Table 2).

Some limitations have been acknowledged in establishing IVIVCs. For example, the sample size amongst the *in vitro* approaches differed. The sample size of the RBALP approach was 104, while this value was only 12 for RIVM (1:37.5 S/L ratio). Such discrepancy would impact on the reliability and stability of the estimate for RIVM. Also, in this study, we did not consider the inter-species uncertainties of RBA. An underlying assumption here is the relative absorption ability among species should be the same. This assumption should be validated using various animals for the same soil, however, consistent data are presently not available. Another limitation here is that mining and smelting represent two different types of anthropogenic activities, causing different Pb speciation thereby variable Pb bioavailability in soil. However, since some previous studies mixed the two soil types, to compare between Pb bioavailability between mining and smelting impacted soils may require further investigation

in the future. Also, the variations of BAc were not considered in the analysis. For example, in some cases the CV of 46.8% was observed for Pb during inter-laboratory assessment of the UBM (Wragg et al. 2011). However, the variations of BAc are much lower than that of RBA as presented in SM Figure S1 in most cases. Thus, such a consideration may wield limited influence on the results.

While these limitations may result in some error or bias in our study. a major aim of this study was to minimize reducible uncertainties when establishing IVIVCs. Based on all the available data and computational techniques, this study provides an informed attempt to better understand the relationships between RBA and BAc.

4.3. Implications of RBA predictions.

The RBA for residential land was observed to be higher than the other types. Of the collected data for residential land, the median Pb concentration was summarized as 1200 mg/kg. Therefore the daily soil intake for children can be up to 33.6 µg per day, based on IEUBK model simulation (model assumption: daily consumption for soil is 100mg) (U.S. EPA 2007). This value can increase to 3 to 6 µg/dL blood level for children aged 0.5 to 6, which contributes considerably when children are exposed to such levels of Pb.

On the basis of type-specific RBA analysis conducted in this study, the soil types may not provide useful RBA predictions: only the differences between residential land and house dust, other soils were significant. However, an overall RBA estimation of 49 % in this study differed that from the RBA value of 60% that was selected by U.S. EPA in the IEUBK model (U.S. EPA 2002). This estimation indicated that the previous standard may be a conservative strategy. The lower estimate of RBA in this study may benefit the relevant stakeholder when establishing the clean-up goal and environmental regulations. For example, the IEUBK model helps the standard setting for soil (U.S. EPA 1998). Therefore a hazard standard of 400 mg/kg by weight in play areas and an average of 1200 mg/kg in bare soil in the remainder of the yard were released (U.S. EPA 1998). If the lower RBA presented in this study can be updated in the IEUBK model, the outcome may be a more tolerable Pb exposure criterion. In reality, the site-types information may be unclear or contaminations may be from multiple sources. It is recommended that a prior assessment of site-specific BAc be undertaken, and then RBA can be predicted by applying the IVIVCs.

To the authors' knowledge, this is the first time available data have been used to underpin IVIVCs analysis, and the robustness, reliability and comparisons of the established IVIVCs have been documented. Currently, developed IVIVCs still require future validation using *in vivo* experiments, and a validated IVIVC can be anticipated to help predict RBA, together with *in vitro* measurements. Meanwhile, RBA estimations presented here for different soil categories are simply empirical judgments. It should be noted that soils constitute variable material from site to site, and thus the RBA estimations should be treated with much caution in practice. In summary, this study is a new approach to estimating soil RBA according to soil types. Estimation of type-specific RBA can help: firstly, evaluate the potential risk arising from Pb exposure; and secondly, determine more precisely the clean-up goal.

5. Acknowledgements

We would like to thank the Cooperative Research Centre for Contamination Assessment and Remediation of the Environment (CRC CARE) for funding support.

6. References

Axelrad DA, Bellinger DC, Ryan LM, Woodruff TJ. 2007. Dose-response relationship of prenatal mercury exposure and IQ: An integrative analysis of epidemiologic data. Environ Health Perspect 115:609-615.

Bannon DI, Drexler JW, Fent GM, Casteel SW, Hunter PJ, Brattin WJ, et al. 2009. Evaluation of small arms range soils for metal contamination and lead bioavailability. Environ Sci Technol 43:9071-9076.

Casteel SW, Weis CP, Henningsen GM, Brattin WJ. 2006. Estimation of relative bioavailability of lead in soil and soil-like materials using young swine. Environ Health Perspect 114:1162-1171.

Denys S, Caboche J, Tack K, Delalain P. 2007. Bioaccessibility of lead in high carbonate soils. J Environ Sci Heal A 42:1331-1339.

Denys S, Caboche J, Tack K, Rychen G, Wragg J, Cave M, et al. 2012. In vivo validation of the unified BARGE method to assess the bioaccessibility of arsenic, antimony, cadmium, and lead in soils. Environ Sci Technol 46:6252-6260.

Deshommes E, Tardif R, Edwards M, Sauvé S, Prévost M. 2012. Experimental determination of the oral bioavailability and bioaccessibility of lead particles. Chem Cent J 6:138-168.

Dodd M, Rasmussen PE, Chénier M. 2013. Comparison of two in vitro extraction protocols for assessing metals' bioaccessibility using dust and soil reference materials. Hum Ecol Risk Assess 19:1014-1027.

Hettiarachchi GM, Pierzynski GM, Oehme FW, Sonmez O, Ryan JA. 2003. Treatment of contaminated soil with phosphorus and manganese oxide reduces lead absorption by sprague-dawley rats. J Environ Qual 32:1335-1345.

Juhasz AL, Weber J, Smith E, Naidu R, Marschner B, Rees M, et al. 2009. Evaluation of SBRC-gastric and SBRC-intestinal methods for the prediction of in vivo relative lead bioavailability in contaminated soils. Environ Sci Technol 43:4503-4509.

Juhasz AL, Smith E, Weber J, Rees M, Kuchel T, Rofe A, et al. 2013. Predicting lead relative bioavailability in peri-urban contaminated soils using in vitro bioaccessibility assays. Journal of Environmental Science and Health, Part A 48:604-611.

Kesteren PCEv, Walraven N, T. Schuurman, R. A. Dekker, R. Havenaar, A.J.H. Maathuis , et al. 2014. Bioavailability of lead from Dutch made grounds, a validation study. National Institute for Public Health and the Environment, Ministry of Health, Welfare and Sport, Dutch.

Levin R, Brown MJ, Kashtock ME, Jacobs DE, Whelan EA, Rodman J, et al. 2008. Lead exposures in US children, 2008: Implications for prevention. Environ Health Perspect 116:1285-1293.

Li H, Cui XY, Li K, Li J, Juhasz Albert L, Q ML. 2014. Assessment of in vitro lead bioaccessibility in house dust and its relationship to in vivo lead relative bioavailability. Environ Sci Technol 48:8548-8555.

Li J, Li K, Cave M, Li H-B, Ma LQ. 2015. Lead bioaccessibility in 12 contaminated soils from China: Correlation to lead relative bioavailability and lead in different fractions. Journal of hazardous materials 295:55-62.

Li J, Li C, Sun HJ, Juhasz AL, Luo J, Li HB, et al. 2016. Arsenic relative bioavailability in contaminated soils: Comparison of animal models, dosing schemes, and biological end points. Environ Sci Technol 50:453-461.

Mielke HW, Reagan PL. 1998. Soil is an important pathway of human lead exposure. Environ Health Perspect 106:217-229.

Naidu R, Channey R, McConnell S, Johnston N, Semple KT, McGrath S, et al. 2015. Towards bioavailability-based soil criteria: Past, present and future perspectives. Environ Sci Pollut Res 22:8779-8785.

Oliver DP, McLaughlin MJ, Naidu R, Smith LH, Maynard E, Calder I. 1999. Measuring Pb bioavailability from household dusts using an in vitro model. Environ Sci Technol 33:4434-4439.

Oomen AG, Brandon EFA, Swartjes FA, Sips A. 2006. How can information on oral bioavailability improve human health risk assessment for lead-contaminated soils? Implementation and scientific basis. Ministry of Housing, Spatial Planning and the Environment (VROM), Directorate-General for Environmental Protection, Directorate of Soil, Water and Rural Area (BWL).

Ortega Calvo JJ, Harmsen J, Parsons JR, Semple KT, Aitken MD, Ajao C, et al. 2015. From bioavailability science to regulation of organic chemicals. Environ Sci Technol 49:10255-10264.

Ruby MV, Davis A, Link TE, Schoof R, Chaney RL, Freeman GB, et al. 1993. Development of an in vitro screening test to evaluate the in vivo bioaccessibility of ingested mine-waste lead. Environ Sci Technol 27:2870-2877.

Ruby MV, Davis A, Schoof R, Eberle S, Sellstone CM. 1996. Estimation of lead and arsenic bioavailability using a physiologically based extraction test. Environ Sci Technol 30:422-430.

Ryan JA, Scheckel KG, Berti WR, Brown SL, Casteel SW, Chaney RL, et al. 2004. Peer reviewed: Reducing children's risk from lead in soil. Environ Sci Technol 38:18A-24A.

Schroder J, Basta N, Casteel S, Evans T, Payton M, Si J. 2004. Validation of the in vitro gastrointestinal (IVG) method to estimate relative bioavailable lead in contaminated soils. J Environ Qual 33:513-521.

Smith E, Kempson IM, Juhasz AL, Weber J, Rofe A, Gancarz D, et al. 2011. In vivo–in vitro and XANES spectroscopy assessments of lead bioavailability in contaminated periurban soils. Environ Sci Technol 45:6145-6152.

U.S. EPA. 1998. Hazard standard risk analysis -- TSCA section 403. Washington:U.S. Environmental Protection Agency,.

U.S. EPA. 2002. User's guide for the integrated exposure uptake biokinetic model for lead in children (IEUBK) windows® version – 32 bit version. Washington, DC: U.S. Environmental Protection Agency.

U.S. EPA. 2007. Estimation of relative bioavailability of lead in soil and soil-like materials using in vivo and in vitro methods, OSWER 9285.7-77. Washington, DC: U.S. Environmental Protection AgencyWashington,.

Whitehead A. 2002. Meta-analysis of controlled clinical trials:John Wiley & Sons.

Wijayawardena MA, Naidu R, Megharaj M, Lamb D, Thavamani P, Kuchel T. 2015. Using soil properties to predict in vivo bioavailability of lead in soils. Chemosphere 138:422-428.

Wragg J, Cave M, Basta N, Brandon E, Casteel S, Denys S, et al. 2011. An inter-laboratory trial of the unified BARGE bioaccessibility method for arsenic, cadmium and lead in soil. Sci Total Environ 409:4016-4030.

Wu CFJ. 1986. Jackknife, bootstrap and other resampling methods in regression-analysis. Annals of Statistics 14:1261-1295.

Xu T, White L, Hui D, Luo Y. 2006. Probabilistic inversion of a terrestrial ecosystem model: Analysis of uncertainty in parameter estimation and model prediction. Glob. Biogeochem. Cycles 20, GB2007.

Yan K, Dong ZM, Liu YJ, R N. 2015. Quantifying statistical relationships between commonly used in vitro models for estimating lead bioaccessibility. Doi: 10.1007/s11356-015-5947-8. Environ Sci Pollut Res:1-10.

List of Tables

574 **Table 1.** Summary of data collected from the literature

Paired BAc-RBA data for IVIVC			RBA (%)	BAc(%)
Methods	Biomarker	N	Mean (Median)	
RBALP	Blood/liver/kidney/femur	104	66 (70)	64 (71)
SBRC	Blood	29	40 (43)	69 (74)
UBM	Liver/kidney/femur/urine	67	49 (39)	37 (31)
RIVM[a]	Blood/liver/kidney/femur	40	52 (56)	62 (83)
RIVM[b]	Blood/liver/kidney/femur	12	72 (82)	69 (68)
Type-specific BAc and RBA data			RBA (%)	BAc(%)
	Data type[c]	N	Mean (Median)	
House Dust	Both 1 and 2	45	50 (52)	57 (66)
Residential	Both 1 and 2	59	62 (58)	52 (53)
Mining/Smelter	Both 1 and 2	77	50 (48)	40 (37)
Others	Both 1 and 2	22	38 (27)	68 (64)

575 The raw data are available in Supplementary Materials Table S1, S2.
576
577 Abbreviations. BAc: bioaccessibility; RBA: relative bioavailability; IVIVC: *in vitro* and *in*
578 *vivo* correlation; RBALP: relative bioaccessibility leaching procedure; SBRC:
579 Solubility/Bioavailability Research Consortium; UBM: BARGE Unified Bioaccessibility;
580 RIVM: National Institute for Public Health and Environment method; N: sample number
581 Note: a, S/L ratio is 1:375; b, S/L ratio is 1:37.5; c: 1 is type-specific BAc and 2 is
582 type-specific RBA data.
583

584 **Table 2.** Posterior estimations for model parameters, using Bayesian inference
585

Parameter		Mean (Median)	SD	95% CI
intercept (β)		4.70 (4.69)	2.47	(-0.13, 9.56)
coefficient (α_0)		0.87 (0.87)	0.16	(0.55, 1.19)
study effect (γ)	RBALP (γ_1)	0.075 (0.075)	0.15	(-0.23, 0.39)
	SBRC (γ_2)	-0.37 (-0.36)	0.16	(-0.70, 0.056)
	UBM (γ_3)	-0.018 (-0.018)	0.15	(-0.33, 0.30)
	RIVM[a] (γ_4)	-0.038 (-0.037)	0.15	(-0.35, 0.28)
	RIVM[b] (γ_5)	0.32 (0.32)	0.16	(-0.0088, 0.67)
biomarker effect (λ)	Blood (λ_1)	0.018 (0.018)	0.055	(-0.086, 0.13)
	Liver (λ_2)	0.039(0.037)	0.053	(-0.061, 0.13)
	Kidney (λ_3)	-0.018 (-0.015)	0.054	(-0.13, 0.079)
	Femur (λ_4)	-0.067 (-0.061)	0.057	(-0.19, 0.023)
	Urine (λ_5)	0.017 (0.014)	0.066	(-0.11, 0.16)

586 The parameter definitions are provided in Equation 3.
587
588 Abbreviations. SD: standard deviation; CI: confidential interval; IVIVC: *in vitro* and *in*
589 *vivo* correlation; RBALP: relative bioaccessibility leaching procedure; SBRC:
590 Solubility/Bioavailability Research Consortium; UBM: BARGE Unified Bioaccessibility;
591 RIVM: National Institute for Public Health and Environment method.
592 Note: a, Solid/Liquid ratio is 1:375; b, Solid/Liquid ratio is 1:37.5. The parameters (β, α_0, γ
593 and λ) were defined in Equation 3.

Table 3. Summary of reported IVIVCs

Sample descriptions (sample size)	*In vivo* animal/biomarker	*In vitro* model	IVIVCs[d]	Reference
EPA region VIII (n=19)	Swine/blood	RBALP[c]	$y = 0.88x - 0.028$. $r^2 = 0.93$	(U.S. EPA 2007)
Soils (n=12)	Mice/blood	RBALP	$y = 0.69x + 30.21$. $r^2 = 0.78$	(Smith et al. 2011)
Farming, mining and smelter soils in China (n=12)	Mice/blood	SBRC	$y = 0.40x + 14.0$. $r^2 = 0.43$	(Li et al. 2015)
House dust (n=24)	Mice/blood	SBRC	$y = 0.61x + 3.15$. $r^2 = 0.68$	(Li et al. 2014)
Farming, mining and smelter soils in China (n=12)	Mice/blood	UBM	$y = 0.80x + 9.99$. $r^2 = 0.67$	(Li et al. 2015)
Mining and smelter soils in Europe (n=16)	Swine/urine, bone, kidney and liver	UBM	$y = (0.6 \text{ to } 1.2)x+(0 \text{ to } 5)$. $r^2 > 0.6$	(Denys et al. 2012)
Jasper Yard soils, residential soils, slag soils (n=12)	Swine/blood	UBM	$y = 0.78x$, $r^2 = 0.61$	(Wragg et al. 2011)
EPA Region VIII, Bunker hill (n=10)	Swine/blood	RIVM[a] (0.06)	$x = 1.08y$, $r^2 = 0.68$	(Oomen et al. 2006)
EPA Region VIII, Bunker hill (n=7)	Swine/blood	RIVM[b] (0.6)	$x = 0.79y$, $r^2 = 0.95$	(Oomen et al. 2006)

Abbreviations. IVIVCs: *in vitro* and *in vivo* correlations; RBALP: relative bioaccessibility leaching procedure; SBRC: Solubility/Bioavailability Research Consortium; UBM: BARGE Unified Bioaccessibility; RIVM: National Institute for Public Health and Environment method.

Note: a, Solid/Liquid ratio is 1:375; b, Solid/Liquid ratio is 1:37.5; c, based on weighted linear regression; d, *x*: bioaccessibility and *y*: bioavailability.

List of Figures

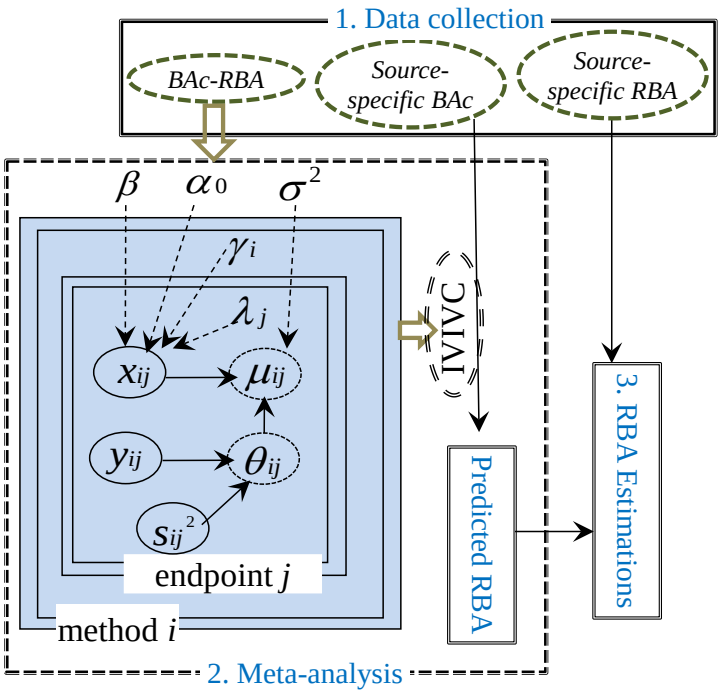

Figure 1. Framework for estimating lead bioavailability. Model definition is provided in Equation 3. Abbreviations. BAc: Bioaccessibility; RBA: relative bioavailability; *x*: adjusted bioaccessibility; *y*: measured RBA; θ: expected RBA; μ: real RBA; s^2: individual variance; σ^2: population variance; β: intercept; α_0: overall coefficient; γ: absolute coefficient differences among methods; λ: absolute coefficient differences among endpoints; IVIVC: *in vitro* and *in vivo* correlation.

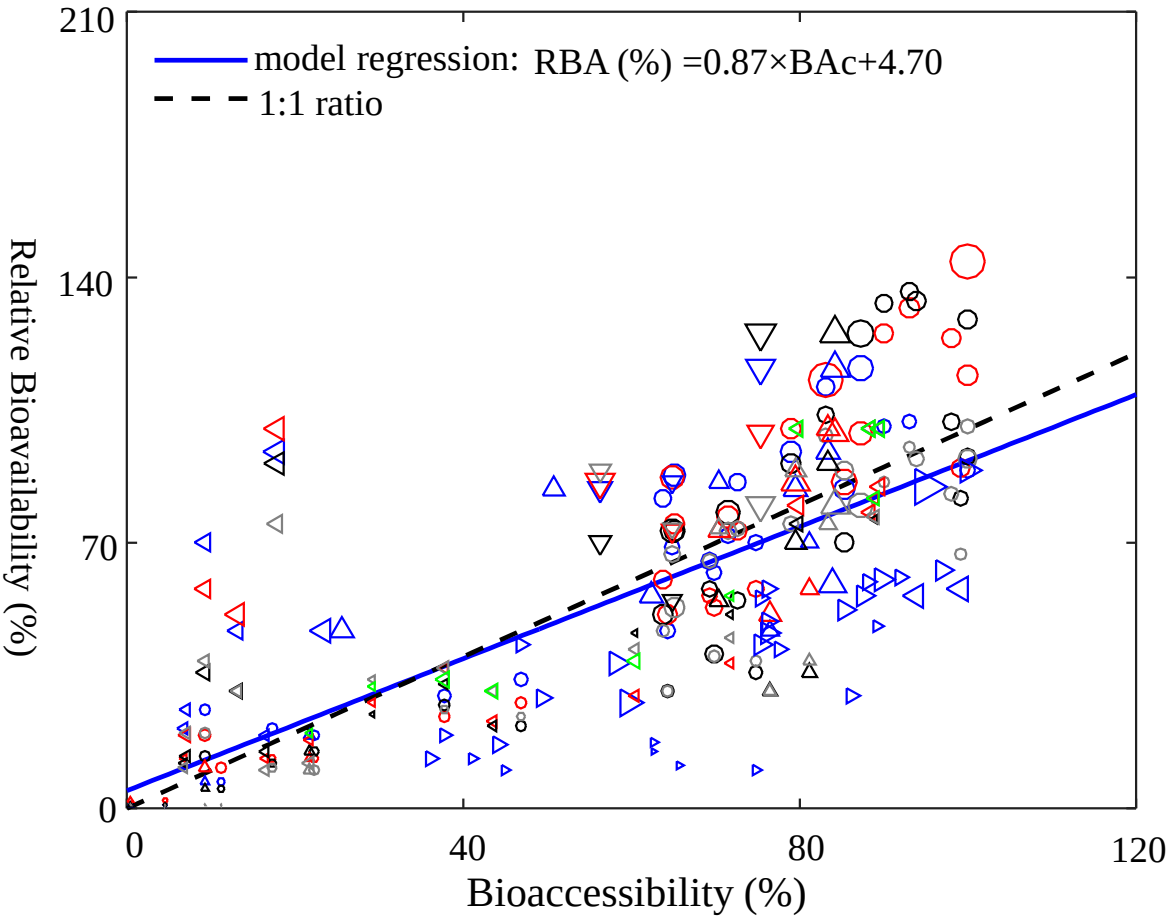

Figure 2. The scatter plot for bioaccessibility and relative bioavailability.
Color: Blue (blood); Red (liver); Black (kidney); grey (femur); Green (urine).
Method: circle (RBALP); Right-pointing triangle (SBRC); Left-pointing triangle (UBM); Upward-pointing triangle (RIVM,S/L ratio = 1:375); Downward-pointing triangle (RIVM,S/L ratio =1:37.5)
The marker size was plotted based on the standard error of separate bioavailability (5/12 inch per standard error).

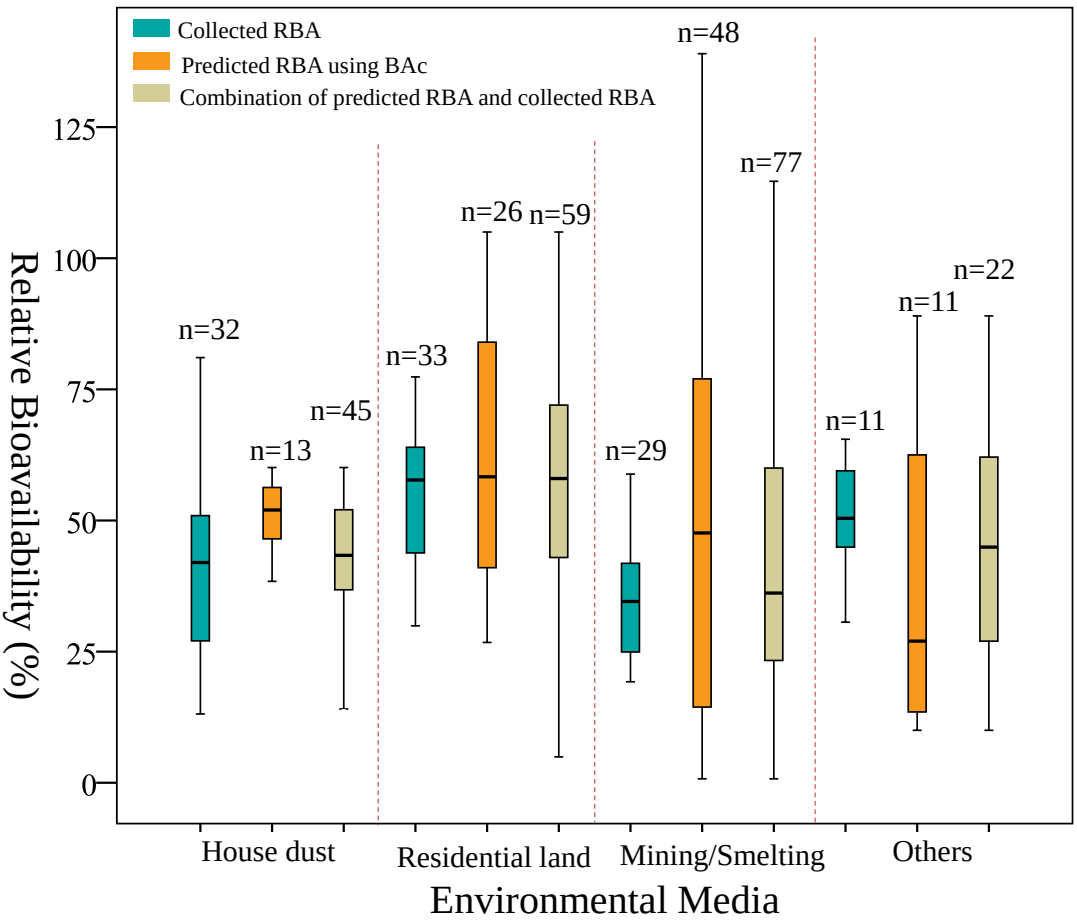

Figure 3. Boxplots for the type-specific Pb RBA

For each box the central mark is the median, the edges of the box are the 25th and 75th percentiles, and the whiskers represent the most extreme data points without consideration of outliers.

Abbreviations. BAc: Bioaccessibility; RBA: relative bioavailability; n, sample size.