

1 **Independent data validation of an in vitro method for prediction of relative**
2 **bioavailability of arsenic in contaminated soils**

3

4 Karen D. Bradham^{a*}, Clay Nelson^a, Albert L. Juhasz^b, Euan Smith^b, Kirk Scheckel^c, Daniel
5 R. Obenour^d, Bradley W. Miller^c, David J. Thomas^a

6

7 ^aUnited States Environmental Protection Agency, Office of Research and Development,
8 Research Triangle Park, NC 27711, USA.

9

10 ^bCentre for Environmental Risk Assessment and Remediation, University of South Australia,
11 Mawson Lakes, SA 5095, Australia.

12

13 ^cUnited States Environmental Protection Agency, Office of Research and Development,
14 Cincinnati, OH 45224, USA.

15

16 ^dNorth Carolina State University, Department of Civil, Construction, and Environmental
17 Engineering, Raleigh, NC 27695, USA.

18

19 *Corresponding Author: bradham.karen@epa.gov, 109 TW Alexander Drive, MD-D205-05,
20 Research Triangle Park, North Carolina; phone: (919) 541-9414; fax: (919) 541-3527.

21

22 **Keywords**

23 Arsenic, Bioaccessibility, Bioavailability, Correlation, Relative Bioavailability, SBRC

24

25

26 **Abstract**

27 In vitro bioaccessibility assays (IVBA) estimate arsenic (As) relative bioavailability (RBA) in
28 contaminated soils to improve accuracy in human exposure assessments. Previous studies
29 correlating soil As IVBA with RBA have been limited by use of few soil types and sources of
30 As, and the predictive value of As IVBA assays have not been validated using an independent
31 set of As-contaminated soils. In this study, a robust linear model was developed to predict As
32 RBA in mice using an IVBA assay and the predictive capability of the model was
33 independently validated using a unique set of As-contaminated soils. Forty As-contaminated
34 soils varying in soil type and contaminant source were included in this study, with 31 soils
35 used for initial model development and nine soils used for independent model validation.
36 The initial model reliably predicted As RBA values in the independent data set, with a mean
37 As RBA prediction error of 5.4%. Following validation, 40 soils were used for final model
38 development, resulting in a linear model with the equation: $RBA = 0.65 * IVBA + 7.8$ and R^2
39 of 0.81. The in vivo-in vitro correlation and independent data validation presented provide
40 critical verification necessary for regulatory acceptance in human health risk assessment.

41

42

43

44

45

46

47

48

49

50

51 **Introduction**

52 Arsenic (As) is the most frequently occurring contaminant on the Priority List of Hazardous
53 Substances, which lists substances of greatest public health concern to people living at or
54 near U.S. National Priority Listing sites.¹ Human exposure to As via ingestion of As-
55 contaminated soils can have serious health impacts including increased cancer risk.^{2,3,4}
56 Accurate assessment of human health risks from exposure to As-contaminated soils depends
57 on estimating its bioavailability, defined as the fraction of ingested As absorbed across the
58 gastrointestinal barrier and available for systemic distribution and metabolism. Arsenic
59 bioavailability varies among soils and is influenced by site-specific soil physical and
60 chemical characteristics and internal biological factors. U.S. Environmental Protection
61 Agency (USEPA) guidance describes the need for development of soil As bioavailability
62 methods and data to improve the accuracy of human exposure and risk calculations at As-
63 contaminated sites.⁵

64

65 Difficulties inherent in measuring site-specific soil arsenic bioavailability in humans⁶ have
66 prompted development of in vivo animal bioassays to determine As relative bioavailability
67 (RBA) in soil.⁷⁻¹⁴ Although, mice and humans differ in metabolism and disposition of
68 arsenicals, similarities are sufficient to permit use of mouse data to create physiologically
69 based pharmacokinetic models that can be scaled for humans⁷. For these assays, the
70 bioavailability of soil As is expressed relative to the bioavailability of a completely water
71 soluble form of As (i.e., sodium arsenate). Currently, the USEPA requires the use of in vivo
72 models for assessing the RBA of As- contaminated soils.¹⁵ However, time and cost
73 considerations often limit their use in risk assessment and result in the use of default values
74 for As RBA.⁷

75

76 As an alternative to in vivo bioassays, in vitro bioaccessibility (IVBA) assays have been
77 developed to measure the extent of As solubilization in simulated gastrointestinal fluids.^{7, 12,}
78 ¹⁶⁻²¹ IVBA assays are attractive alternatives to in vivo assays because they are cost-effective
79 and reduce reliance on animal studies. A prime assumption underlying these IVBA assays is
80 that the fraction of As solubilized in vitro is similar to the fraction of As that can cross the
81 gastrointestinal barrier.²² If an IVBA method is an appropriate surrogate, then it must be
82 shown to reliably predict in vivo RBA.⁵ While some studies have examined the relationship
83 between As RBA and IVBA,^{7,12,16,19,20,23} validation of this relationship using an independent
84 set of soils is the next critical step for regulatory acceptance.

85

86 Multiple in vivo animal models and in vitro methods have been proposed to assess As RBA
87 and IVBA, respectively, in contaminated soils.^{7,12,16,19,20,23} A recent study described a mouse
88 assay as a cost effective and reproducible alternative to other animal assays.^{7,8} Until recently,
89 precision of As RBA estimates determined from repeated assays of the same soils had not
90 been reported for any animal model. Low between-assay variation in urinary excretion
91 fraction (UEF) and RBA estimates in the mouse assay results in a highly reproducible,
92 inexpensive in vivo model.⁸ A strong relationship was noted between As RBA estimated
93 from the mouse assay and As IVBA determined using a simplified gastric phase method⁷
94 hereafter referred to as the Solubility/Bioavailability Research Consortium (SBRC) method.²⁴
95 A study evaluating the correlation between the mouse model and 5 commonly employed in
96 vitro methods, which varied in operational parameters from simplified gastric methods²⁴ to
97 complex physiological methods aimed at replicating human digestive systems,²⁵ reported that
98 the strongest correlation was found between the results obtained with the mouse model and
99 the SBRC method.²⁶ A similar study comparing results from a juvenile swine model and the
100 SBRC method also found a strong correlation.²⁰

101
102
103
104
105
106
107
108
109
110
111
112
113
114

The objective of this study was to build upon a previously published linear regression model⁷ to predict As RBA in mice using an IVBA assay and to develop a more robust model across multiple soil types, As contaminant sources, and As concentrations. A second objective was to validate the predictive capability of this model using an independent set of As-contaminated soils. Although earlier studies have evaluated correlations between As RBA and IVBA, these studies have lacked model validation using an independent set of soils and been limited with respect to variety of soil types and contaminant sources used to construct the model. Validation of model performance using data independent to those used to construct the model is imperative for IVBA data to be used routinely for incorporation into human health risk assessments.²⁷ This is particularly important because the predictive capability of the model may be overestimated when evaluated solely with samples used to construct the model.²⁸

115 **Materials and Methods**

116 *Test Soils and Standard Reference Materials*

117 This study used 37 As-contaminated soils in which As was introduced by mining and
118 smelting, pesticide or defoliant use in agricultural or orchard sites, railway corridors, cattle
119 tick dip sites, or occurred as a natural soil constituent. Standard reference materials (SRMs),
120 SRM 2710 and SRM 2710a (National Institute of Standards and Technology), and a USEPA
121 reference material were also evaluated. No soils spiked or amended with As were included in
122 this study. All test soils were collected from the top 1-2” of soil, dried (<40°C), sieved to
123 <250 µm, homogenized, and riffled²⁹ for mixing and splitting samples.

124

125 Total As concentrations in test soils and SRMs were determined by Instrumental Neutron
126 Activation Analysis (INAA) at the Department of Nuclear Engineering, North Carolina State
127 University, Raleigh. The mean As mass detection limit was 0.035 µg (approximately 0.2 µg/g
128 soil). Additional soil element concentrations (Al, Fe, Mn, and P) were determined by
129 microwave digestion in accordance with USEPA SW-Method 3051 with analysis by
130 Inductively Coupled Plasma-Atomic Emissions Spectroscopy in accordance with USEPA
131 SW-Method 6010C.

132

133 A subset of test soils (soils 1-5; 8-27) were also characterized for As speciation using the
134 Materials Research Collaborative Access Team’s (MRCAT) beamline 10-ID, Sector 10 at the
135 Advanced Photon Source (APS), Argonne National Laboratory (ANL), Argonne, IL.
136 Additional information on As speciation determination is provided in Supporting
137 Information.

138

139 ***Assessment of As relative bioavailability***

140 Arsenic RBA was determined using an in vivo mouse model.^{7,8} All assays were performed in
141 four- to six-week-old female C57BL/6 mice (Charles River Laboratories, Raleigh, NC).

142 Additional details on mouse assay methodology is provided in Supporting Information.

143 Data from each mouse assay were used to calculate the urinary excretion fraction (UEF) of

144 As from ingestion of an amended diet as the ratio of cumulative excretion of As in urine (μg)

145 to cumulative dietary intake of arsenic (μg) as shown in Equation 1:

146
$$UEF\% = 100 \cdot \frac{\text{Cum Urinary As}}{\text{Cum As Dose}} \quad \text{Eq. (1)}$$

147 Arsenic RBA was calculated as the ratio of the UEF for As in a specific soil-amended diet to

148 the UEF for As in a diet containing sodium arsenate heptahydrate (see Equation 2):

149
$$RBA\% = \frac{UEF \text{ Soil}}{UEF \text{ Arsenate}} \quad \text{Eq. (2)}$$

150 Each UEF in Equation 2 is derived from multiple estimates of UEF for groups of three mice

151 housed together in a single metabolic cage (the unit of measure in the assay is data from a

152 single cage).

153

154 ***Assessment of As bioaccessibility***

155 Arsenic IVBA was determined using the SBRC in vitro method (USEPA Method 9200.86-

156 2).²⁴ See Supporting Information for additional details on IVBA methodology.

157

158 Arsenic IVBA was calculated and expressed on a percentage basis according to equation 3.

159

160
$$\text{As bioaccessibility (\%)} = \left(\frac{\text{in vitro As}}{\text{total As}} \right) \times 100 \quad \text{Eq. (3)}$$

161

162 Where:

163

164 In vitro As = As extracted during the in vitro assay

165 Total As = Amount of As in the contaminated soil used for bioaccessibility determination

166

167 ***As relative bioavailability prediction: model formulation and validation***

168 In this study, the correlation between As RBA and IVBA was determined for 40 soils. A
169 training set of 31 soils (#1-31) were used for initial model development, and nine additional
170 soils (#32-40), previously described by Juhasz et al,¹² were used to independently validate the
171 in vivo-in vitro correlation. The soils used for the independent data validation are from
172 Australia and contain As from different contamination sources (e.g., cattle dip), mineralogy,
173 and As concentrations versus the soils from the U.S. Following validation, the regression
174 model was then fitted using all 40 soils.

175

176 A Bayesian hierarchical approach to linear regression was used to evaluate the ability of the
177 SBRC in vitro assay to predict As RBA in the form:

178

$$179 \quad \text{RBA (\%)} = a + (b) \text{ IVBA (\%)} + \epsilon$$

180 where,

181 a = y-intercept

182 b = slope

183 ϵ = normally distributed prediction error

184

185 This approach has the advantage over simple linear regression of accounting for variation
186 among replicate measurements of individual soils, as well as variation among different soils

187 (see Supporting Information for a more detailed summary of the Bayesian model
188 formulation).

189

190 The predictive capability of the model was assessed using the coefficient of determination
191 (R^2) and absolute error (AE) in As RBA prediction. Here, R^2 is defined as the fraction of the
192 variance in the observations that is resolved by the model predictions (i.e., the means of the
193 predictive distributions) relative to a null (constant-only) model.³⁰ AE is defined as the
194 absolute percent difference between the model-predicted As RBA value and the As RBA
195 value observed in the mouse assay.

196

197 To evaluate model robustness, model parameters (slope and y-intercept) and As RBA
198 prediction performance under each model development scenario (training data only,
199 independent validation data only, and all soils) were compared. A “leave-one-out” cross
200 validation (CV) was also performed for the overall (i.e., 40 soil) fitted model to further assess
201 model robustness by evaluating model prediction “out-of-sample” over a wider range of
202 observations.³¹ In this case, As RBA for each of the 40 soil types was predicted, in turn, after
203 removing the target soil type from the observation data set and re-calibrating the model based
204 on the remaining 39 observations.

205

206 **RESULTS**

207 *Test Soils and Standard Reference Materials*

208 Test soils and SRMs displayed a range of As and other elemental concentrations, pH values,
209 and speciation (Tables SI-1 and SI-2). Total As concentration in test soils ranged from 108 to
210 6,899 mg kg⁻¹ (Table 1). The concentration of major elements, including aluminum (Al) and
211 iron (Fe) ranged from 0.7 - 72.1 g/kg and 14.4 - 276.2 g/kg, respectively. The concentration

212 of manganese (Mn) and phosphorous (P) ranged from 0.5 - 9,321 and 4 - 6,745 mg kg⁻¹,
213 respectively. Soil pH ranged from 2.2 to 8.8. Arsenic speciation was categorized into three
214 coordination environments, As(V)-oxygen bonding (arsenate sorbed to oxides and scorodite),
215 As(III)-oxygen bonding (arsenite sorbed to oxides and schneiderhöhnite), and As-sulfide
216 bonding (realgar, loellingite, and arsenopyrite). With respect to speciation, mining soils had
217 varying ratios of all three arsenic coordination environments (Table SI-2). Mining Soil 12
218 was mostly As(III)-oxygen bonding (26%) and As-sulfide bonding (60%), but the remaining
219 mining soils contained mostly As(V)-oxygen bonding species. Orchard soils were
220 predominately sorbed As(V) phases with the exception of Soil 21, which had about 10%
221 sorbed As(III). The reference material soils were predominantly sorbed As(V) and scorodite
222 with minor addition of sorbed As(III).

223

224 *As relative bioavailability and bioaccessibility in test soils*

225 Arsenic RBA values observed in the mouse assay ranged from 1.9-52.8 % (Table 1). Arsenic
226 IVBA in test soils and SRMs ranged from 0.0-74.3 % (Table 1), while sodium arsenate IVBA
227 was 100%. In addition to a strong correlation with As RBA values, acceptable within-
228 laboratory repeatability and between-laboratory reproducibility must be established in order
229 for an in vitro method to be accepted.⁵ Although, this study was not designed as an inter-
230 laboratory trial, information is provided in the discussion regarding the repeatability and
231 reproducibility of the SBRC method.

232

233 The current study provided SBRC values for 23 soils determined at two independent
234 laboratories. Observed standard deviations (SDs) ranged from 0.1 to 6.7%. Comparison of

235 between-lab variability resulted in a strong correlation (slope = 1.0; y-intercept = 3.7; $R^2 =$
236 0.92) (Figure 1) indicating that the assay was reproducible.

237

238 ***Regression Model Performance – Utility of in vitro bioaccessibility data for predicting As***
239 ***relative bioavailability***

240 An initial linear model was developed using the training data set (n=31 soils) to evaluate the
241 ability of IVBA values to predict As RBA in mice. The initial linear model had a slope of
242 0.67 (standard error (SE) of 0.06) and y-intercept of 7.1 (SE of 1.8) (Table 2). Goodness of
243 fit, as measured by R^2 , was 0.83. This finding is similar to studies by Juhasz et al.^{20,21} and
244 Brattin et al.,¹⁹ which have reported strong correlations with the SBRC gastric method
245 correlation and in vivo RBA swine data ($R^2=0.75$ and $R^2 = 0.72$, respectively). Bradham et
246 al.⁷ reported a strong correlation between the SBRC method and in vivo RBA mouse data
247 ($R^2=0.92$) for mining soils.

248

249 For independent validation, this initial linear model was used to predict As RBA for nine
250 additional soils (#32-40) with comparison to measured values. The model accurately
251 predicted As RBA for all nine soils in the validation set with a mean and median absolute
252 error (AE) of 5.4 and 6.0% respectively (range of 1.7 to 8.4%) (Table 3). The R^2 for the
253 validation predictions was 0.73.

254

255 Following independent validation, all 40 soils were fitted to an updated linear regression
256 model (Figure 2 and Table 2). Parameters for this model were similar to the initial model
257 with a slope of 0.65 (SE of 0.05), y-intercept of 7.8 (SE of 1.6), and R^2 of 0.81 ($R = 0.91$).
258 Mean and median AE in As RBA prediction across all 40 soils were 4.9% and 4.8%,
259 respectively. In addition, 39 of the 40 predicted As RBA values were within 10% of the

260 RBA value observed in the mouse assay; only soil 17 (AE of 16.7%) exceeded the target
261 range. A potential explanation for poor agreement between IVBA and RBA in soil #17 is
262 that high Al levels observed in this soil (66.9 g kg^{-1}) differentially influenced As dissolution
263 in vitro versus in vivo due to either pH specific sorption kinetics or the influence of organic
264 matter in mouse diet on As sorption onto Al surfaces, resulting in the observation of low As
265 IVBA values relative to RBA. Interestingly, the only soil in the data set with higher
266 aluminum levels, soil #36 (72.1 g kg^{-1}) also had a much lower As IVBA than RBA (9.0%
267 versus 21.5% respectively).

268

269 To evaluate model robustness, slope and y-intercept model parameters and As RBA
270 prediction accuracy were compared under the multiple model development scenarios used in
271 this study (i.e., 1. training data set, 2. independent data set, and 3. all 40 soils) (Table 2).
272 Model performance was consistent across all scenarios, with slope and intercept values all
273 within one SE of each other. Mean and median AE in As RBA predictions were within 0.0%
274 and 0.8%, respectively.

275

276 Results of the leave-one-out cross validation (CV) were used to assess model robustness by
277 estimating the model's ability to predict "out-of-sample" across all 40 soils used for model
278 development. This approach also showed consistent estimates in slope and y-intercept across
279 the CV model runs. Slope varied from 0.63 to 0.67 and y-intercept varied from 7.2 to 8.3.
280 Overall model goodness of fit (R^2) for the CV predictions was 0.79 (compared to 0.81 for the
281 "full" model fit to all 40 soils).

282

283 **Discussion**

284 The RBA values reported in this study fall within the As RBA range previously reported for
285 juvenile swine and monkey bioassays.^{11,13,16,19,20,23,32,33} A recent study by the USEPA
286 Technical Review Workgroup Bioavailability Committee compiled all available estimates of
287 soil As RBA across juvenile swine, primate, and mouse assays (103 As RBA values) and
288 reported that only 5% of As RBA values were greater than 60% (USEPA 2012). Based on
289 these studies, As RBA values reported in this study were consistent with these findings and
290 represent a wide range in As RBA values. Differences in bioavailability values for different
291 soils may be largely determined by As mineralogy and physical and chemical properties of
292 soils (Table SI-1) that influence solubility of As in the gastrointestinal system.²² Studies have
293 shown that As bioaccessibility extractability accounts for much of the variability in RBA
294 estimates obtained from the animal bioassays, including the mouse, swine, and monkey
295 assays.^{7,9,11,12,14,23} Some clay minerals contain ferrous and ferric iron that, upon release via
296 weathering, will form iron oxides and hydroxides in soil environments,³⁴ which sorb As
297 reducing As bioavailability. Similar processes have also been identified for aluminum and
298 manganese oxides in soils.^{35,36} Lower As RBA estimates were observed for soils containing
299 sulfide forms of As (realgar or arsenopyrite), which may reflect slow dissolution kinetics of
300 these mineral species. Additional studies would be useful to identify other metals and
301 metalloids in soils that are potential modifiers of As bioavailability and bioaccessibility and
302 to determine concentration dependencies of these interactions.

303

304 Comparison of between-lab variability resulted in a strong correlation (slope = 1.0; y-
305 intercept = 3.7; $R^2 = 0.92$) (Figure 1) indicating that the assay was reproducible.

306

307 Results of the between-lab variability in As bioaccessibility values using the SBRC method
308 support previously published observations that the SBRC method is reproducible between

309 labs. Juhasz et al.²⁶ previously demonstrated a strong relationship (slope = 1.12; y intercept =
310 0.61; $R^2 = 0.98$) between SBRC As bioaccessibility measurements made in their study with
311 data previously published by Bradham et al.⁷ Koch et al.³⁷ conducted an extensive round
312 robin study evaluating 17 bioaccessibility methods, including the SBRC method, by 14
313 laboratories for NIST SRM 2710. For the SBRC method, the between-lab reproducibility SD
314 was 9.5-13% and the individual lab reproducibility SD was 5-8%. A recent study by Brattin
315 et al.¹⁹ reported results of a 4 laboratory comparison of the SBRC method resulting in a
316 within-lab precision of less than 3% (SD) and average SD of 0.8% for the 4 labs. The
317 between-lab variation resulted in an overall average of 3% SD,¹⁹ illustrating the performance
318 of the SBRC in vitro assay.

319

320 Taken together, comparisons of the multiple model development scenarios along with results
321 of the cross validation indicate that model performance is robust with regards to both model
322 parameterization (slope and y-intercept) and As RBA prediction accuracy, as measured by
323 mean and median AE. It is important to note that some range of uncertainty or variability in
324 actual As RBA relative to model predicted As RBA can be expected, due to authentic inter-
325 sample variability in As RBA and/or to measurement error in in vitro bioaccessibility or
326 RBA.⁵ Therefore, the actual As RBA may be either lower or higher than the best estimate
327 predicted value using IVBA data and the regression model (see 95% predictive intervals,
328 Figure 2). Only one of the 40 observations fell outside of the 95% prediction intervals during
329 the cross validation, indicating that the model provides adequate, and perhaps slightly
330 conservative, uncertainty quantification.

331

332 A desirable property of the in vivo-in vitro relationship is a coefficient of correlation (R)
333 greater than or equal to 0.8 which reflects a strong correlation between As RBA and IVBA

334 data.²⁷ The model presented here, which incorporated As RBA and IVBA data from soils
335 with a wide range of As concentrations derived from a variety of anthropogenic and geogenic
336 sources yielded a strong in vivo-in vitro correlation ($R = 0.91$) that met this criterion. Eleven
337 mining and smelter-impacted soils included in this data set had previously been correlated to
338 As IVBA derived from the SBRC gastric in vitro assay using simple least squares linear
339 regression with a reported R^2 of 0.92.⁷ A strong in vivo-in vitro relationship ($R^2 = 0.90$) for
340 the SRBC method has been reported.²⁶ This study found no significant difference in the
341 slope and y-intercepts ($P > 0.05$) of these relationships illustrating the robustness and
342 reproducibility of SBRC as a predictor of As RBA. These investigators also evaluated other
343 in vitro assays reporting no significant difference in the slopes of in vivo-in vitro correlations
344 when SBRC, IVG, PBET, DIN and UBM gastric and intestinal phases (Solubility and
345 Bioavailability Research Consortium, Deutsches Institut für Normung, In Vitro
346 Gastrointestinal, Physiologically Based Extraction Technique, and the Unified BARGE
347 Method, respectively) were used to derive the in vivo-in vitro relationships.²⁶ However, a
348 significantly ($P < 0.05$) smaller y-intercept was determined for the in vivo-in vitro correlation
349 using SBRC compared to the other in vitro methods. This is important to note as the use of
350 in vivo-in vitro correlations with large y-intercepts may over-predict As adsorption,
351 particularly in soils with low As RBA. Other studies^{20,21} determined that SBRC, IVG, PBET,
352 DIN and UBM assays (including gastric and intestinal phases) all predicted As RBA with
353 varying degrees of confidence ($R^2 = 0.52-0.75$). However, comparison of the in vivo and in
354 vitro results from these studies demonstrated that the SBRC gastric method provided the best
355 prediction of in vivo RBA ($R^2=0.75$).^{20,21} Similarly, a strong correlation has been reported
356 between As RBA determined in juvenile swine and As IVBA determined using SBRC (slope
357 = 0.62, y intercept = 19.68, $R^2 = 0.72$).¹⁹ However, this study included soils spiked with
358 exogenous As, which strongly affected the overall R^2 value.

359

360 The approach to measuring As RBA (single versus multiple doses; Area Under the Curve
361 (AUC) versus Steady State Urinary Excretion (SSUE)) may influence in vivo outcomes in
362 terms of whether single versus multiple As doses are administered and whether absorption is
363 determined using AUC or SSUE.²⁶ The USEPA noted that an advantage of steady state
364 models is that they more closely mimic the status of receptors who receive continuous daily
365 exposure to contaminated soil and dust.¹⁵ In addition, under steady state conditions, urinary
366 As excretion is constant so that urinary excretion factors can be estimated by averaging As
367 concentrations from multiple samples over time. Although As RBA comparisons have been
368 made between mouse and swine models using the SSUE approach,⁸ it is unknown to what
369 extent these conditions influence As RBA measurement. Variability in bioaccessibility
370 measurements from in vitro analyses may result from subtle differences in the conduct of
371 assays.¹⁸ To address the uncertainty associated with in vivo-in vitro correlation variability,
372 comparative studies of As RBA with different animal models and endpoints would be
373 advantageous. In addition, assessment of sources of inter-laboratory variability associated
374 with both in vitro and in vivo measurements could be beneficial.

375

376 Oral ingestion of metal contaminated soil and dust is often a “risk driver” for human
377 exposure at contaminated sites, resulting in remedial action. Even a small bioavailability
378 adjustment to site-specific RBA may result in significant remediation cost savings⁵.
379 Therefore, reliable, quick, and inexpensive methods for assessing As RBA in soil are needed
380 to reduce exposure estimate uncertainties in human health risk assessment and reduce clean-
381 up costs. The in vivo-in vitro correlation and independent data validation presented here for
382 the SBRC method provides critical supporting information for use in human health risk
383 assessment.

384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408

Acknowledgements

The authors would like to acknowledge the support of the Centre for Environmental Risk Assessment and Remediation, University of South Australia and the U.S. Environmental Protection Agency, Office of Research and Development. MRCAT operations are supported by the Department of Energy and the MRCAT member institutions. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. This article was reviewed in accordance with the policy of the National Exposure Research Laboratory, U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the view and policies of the Agency.

Supporting Information Available

Soil properties and speciation, bayesian model formulation, equations used for model formulation, references for model formulation. This material is available free of charge via the Internet at <http://pubs.acs.org>.

409 **References**

- 410 (1) ATSDR. Toxicological profile for arsenic; Agency for Toxic Substances and Disease
411 Registry, Division of Toxicology and Environmental Medicine/Applied Toxicology
412 Branch: Atlanta, GA, 2007.
413
- 414 (2) Calabrese, E. J.; Stanek, E.; Barnes, R. Methodology to estimate the amount and
415 particle size of soil ingested by children: implications for exposure assessment at
416 waste sites. *Regul. Toxicol. Pharmacol.* **1996**, 24:264-268.
417
- 418 (3) Davis, A.; Ruby, M. V.; Bergstrom, P. D. Bioavailability of arsenic and lead in soils
419 from the Butte, Montana mining district. *Environ. Sci. Technol.* **1991**, 26:461-468.
420
- 421 (4) Dudka, S.; Miller, W. P. Permissible concentrations of arsenic and lead in soils based
422 on risk assessment. *Water Air Soil Pollut.* **1999**, 113:127-132.
423
- 424 (5) USEPA. *Guidance for Evaluating the Oral Bioavailability of Metals in Soils for Use*
425 *in Human Health Risk Assessment*, OSWER 9285.7-80; U.S. Environmental
426 Protection Agency, 2007.
427
- 428 (6) Stanek, E. J.; Calabrese, E. J.; Barnes, R. M.; Danku, J. M.; Zhou, Y.; Kostecki, P. T.;
429 Zillioux, E. Bioavailability of arsenic in soil: Pilot study results and design
430 considerations. *Hum. Exp. Toxicol.* **2010**, 29:945-960.
431
- 432 (7) Bradham, K. D.; Scheckel, K. G.; Nelson, C. M.; Seales, P. E.; Lee, G. E.; Hughes,
433 M. F.; Miller, B. W.; Yeow, A.; Gilmore, T.; Harper, S.; Thomas, D. J. Relative
434 bioavailability and bioaccessibility and speciation of arsenic in contaminated soils.
435 *Environ. Health Perspect.* **2011**, 119: 1629-1634.
436
- 437 (8) Bradham, K. D.; Diamond, G. L.; Scheckel, K. G.; Hughes, M. F.; Casteel, S. W.;
438 Miller, B. W.; Klotzbach, J. M.; Thayer, W. C.; Thomas, D. J. Mouse assay for
439 determination of arsenic bioavailability in contaminated soil. *J. Tox. Environ. Health,*
440 *Part A* **2013**, 76: 815-826.
441
- 442 (9) Basta, N. T.; Foster, J. N.; Dayton, E. A.; Rodriguez, R. R.; Casteel, S. W. The effect
443 of dosing vehicle on arsenic bioaccessibility in smelter-contaminated soil. *J. Environ.*
444 *Sci. Health, Part A* **2007**, 42: 1275-1281.
445
- 446 (10) Brattin, W. J.; Casteel, S. W. Measurement of arsenic relative bioavailability in swine.
447 *J. Toxicol. Environ. Health A* **2013**, 76: 449-457.
448
- 449 (11) Freeman, G. B.; Schoof, R. A.; Ruby, M. V.; Davis, A. O.; Dill, J. A.; Liao, S. C.;
450 Lapin, C. A.; Bergstrom, P. D. Bioavailability of arsenic in soil and house dust
451 impacted by smelter activities following oral administration in cynomolgus monkeys.
452 *Fundam. Appl. Toxicol.* **1995**, 28: 215-222.
453
- 454 (12) Juhasz, A. L.; Smith, E.; Weber, J.; Rees, M.; Rofe, A.; Kuchel, T.; Sansom, L.;
455 Naidu, R. Comparison of in vivo and in vitro methodologies for the assessment of
456 arsenic bioavailability in contaminated soils. *Chemosphere* **2007**, 69: 961-966.

- 457 (13) Roberts, S. M.; Weimar, W. R.; Vinson, J. R.; Munson, J. W.; Bergeron, R. J.
458 Measurement of arsenic bioavailability in soil using a primate model. *Toxicol. Sci.*
459 **2002**, 67 (2): 303–310.
460
- 461 (14) Roberts, S. M.; Munson, J. W.; Lowney, Y. W.; Ruby, M. V. Relative oral
462 bioavailability of arsenic from contaminated soils measured in the cynomolgus
463 monkey. *Toxicol. Sci.* **2007**, 95 (1): 281–288.
464
- 465 (15) USEPA. *Compilation and review of data on relative bioavailability of arsenic in soil.*
466 OSWER 9200.1-113, U.S. Environmental Protection Agency, 2012.
467
- 468 (16) Rodriguez, R. R.; Basta, N. T.; Casteel, S. W.; Pace, L. W. An in vitro gastrointestinal
469 method to estimate bioavailable arsenic in contaminated soils and solid media.
470 *Environ. Sci. Technol.* **1999**, 33: 642-649.
471
- 472 (17) Ruby, M. V.; Davis, A.; Schoof, R.; Eberle, S.; Sellstone, C. M. Estimation of lead
473 and arsenic bioavailability using a physiologically based extraction test. *Environ. Sci.*
474 *Technol.* **1996**, 30:422-430.
475
- 476 (18) Wragg, J.; Cave, M.; Basta, N.; Brandon, E.; Casteel S.; Denys, S.; et al. An inter-
477 laboratory trial of the unified BARGE bioaccessibility method for arsenic, cadmium
478 and lead in soil. *Sci. Total Environ.* **2011**, 409: 4016-4030.
479
- 480 (19) Brattin, W.; Drexler, J.; Lowney, Y.; Griffin, S.; Diamond, G.; Woodbury, L. An in
481 vitro method for estimation of arsenic relative bioavailability in soil. *Toxicol. Environ.*
482 *Health Part A* **2013**, 76: 458-478.
483
- 484 (20) Juhasz, A. L.; Smith, E.; Weber, J.; Naidu, R.; Rees, M.; Rofe, A.; Kuchel, T.;
485 Sansom, L. Assessment of four commonly employed in vitro arsenic bioaccessibility
486 assays for predicting in vivo arsenic relative bioavailability in contaminated soils.
487 *Environ. Sci. Technol.* **2009**, 43: 9887-9894.
488
- 489 (21) Juhasz, A. L.; Weber, J.; Smith, E. Influence of saliva, gastric and intestinal phases on
490 the prediction of As relative bioavailability using the unified Bioaccessibility
491 Research Group of Europe method. *J Hazard. Materials* **2011**, 197: 161-168.
492
- 493 (22) Ruby M. V.; Schoof R.; Brattin W.; Goldade M.; Post G.; Harnois M.; et al. Advances
494 in evaluating the oral bioavailability of inorganics in soil for use in human health risk
495 assessment. *Environ. Sci. Technol.* **1999**, 33:3697–3705.
496
- 497 (23) Denys, S.; Caboche, J.; Tack, K.; Rychen, G.; Wragg, J.; Cave, M.; Jondreville, C.;
498 Feidt, C. In vivo validation of the unified BARGE method to assess the
499 bioaccessibility of arsenic, antimony, cadmium, and lead in soils. *Environ. Sci.*
500 *Technol.* **2012**, 46: 6252-6260.
501
- 502 (24) Kelley, M. E.; Brauning, S. E.; Schoof, R. A.; Ruby, M. V. *Assessing Oral*
503 *Bioavailability of Metals in Soil.* Battelle Press: Columbus, 2002.
504
- 505 (25) DIN, Deutsches Institut für Normung e. V Soil Quality - Absorption availability of
506 organic and inorganic pollutants from contaminated soil material. DIN E 19738, 2000.

- 507
508 (26) Juhasz, A.L.; Smith, E.; Nelson, C.; Thomas, D.; Bradham, K. Variability Associated
509 with As in Vivo-in Vitro Correlations When using Different Bioaccessibility
510 Methodologies. *Environ. Sci. Technol.* **2014**, 48: 11646-11653.
511
- 512 (27) Juhasz, A. L.; Basta, N. T.; Smith, E. What is required for the validation of in vitro
513 assays for predicting contaminant relative bioavailability? Considerations and criteria.
514 *Environ. Pollut.* **2013**, 180: 372-375.
515
- 516 (28) Kutner, M. H.; Nachtsheim, C. J.; Neter, J., Li, W. *Applied Linear Statistical Models*
517 McGraw-Hill Irwin: Burr Ridge, IL, 2005.
518
- 519 (29) Blume, L. J.; Schumacher, B. A.; Schaffer, P. W.; Cappo, L. A.; Papp, M. L.; Van
520 Remortel R. D.; et al. *Handbook of Methods for Acid Deposition Studies, Laboratory*
521 *Analyses for Soil Chemistry*. EPA/600/S4-90/023; U.S. EPA: Las Vegas, NV, 1991.
522
- 523 (30) Gelman, A.; Hill, J. *Data Analysis Using Regression and Multilevel/ Hierarchical*
524 *Models*. Cambridge University Press: New York, 2007.
525
- 526 (31) Chatfield, C. Model Uncertainty. *Encyclopedia of Environmetrics*; John
527 Wiley & Sons: Hoboken, NJ, 2006.
528
- 529 (32) Casteel, S. W.; Cowart, R. P.; Weis, C. P.; Henningsen, G. M.; Hoffman, E.; Brattin,
530 W. J.; et al. Bioavailability of lead to juvenile swine dosed with soil from the
531 Smuggler Mountain, NPL Site of Aspen, Colorado. *Fundam. Appl. Toxicol.* **1997**,
532 36:177–187.
533
- 534 (33) Rees, M.; Sansom, L.; Rofe, A.; Juhasz, A.; Smith, E.; Weber, J.; Naidu, R.; Kuchel,
535 T. Principles and application of an in vivo swine assay for the determination of
536 arsenic bioavailability in contaminated matrices. *Environ. Geochem. Health* **2009**,
537 31:167–177.
538
- 539 (34) Howell, R.J. Sorption of arsenic by iron oxides and oxyhydroxides in soils. *Appl.*
540 *Geochem.* 1994, 9:279–286.
541
- 542 (35) Jenne, E.A. Controls on Mn, Fe, Co, Ni, Cu, and Zn concentrations in soils and water:
543 the significant role of hydrous Mn and Fe oxides. *Trace Inorg. Water* **1968**, 73:337–
544 387.
545
- 546 (36) McKeague, J. A.; Brydon, J. E.; Miles, M. N. Differentiation of forms of extractable
547 iron and aluminum in soils 1. *Soil Sci. Soc. Am. J.* **1971**, 35:33–38.
548
- 549 (37) Koch, I.; Reimer, K. J.; Bakker, M. I.; Basta, N. T.; Cave, M. R.; Denys, S.; et al.
550 Variability of bioaccessibility results using seventeen different methods on a standard
551 reference material, NIST 2710. *J Environ. Sci. Health Part A* **2013**, 48: 641-655.
552
553
554

555

556 Table 1. Soil source, arsenic concentration, IVBA and RBA values for the 40 soils included
 557 in this study.

Soil ID	Source	Soil [As] (mg/kg) ^a	RBA (%) ± SD	IVBA (%) ± SD	Soil ID	Source	Soil [As] (mg/kg) ^a	RBA (%) ± SD	IVBA (%) ± SD
1	mining	244	15.3 ± 1.7	18.1 ± 0.4	21	orchard	396	46.0 ± 1.9	48.1 ± 0.8
2	mining	173	13.9 ± 1.6	6.8 ± 0.8	22	mining	197	28.7 ± 4.2	22.0 ± 0.2
3	mining	6900	14.5 ± 1.3	17.5 ± 0.6	23	mining	884	22.9 ± 5.3	17.0 ± 0.4
4	mining	280	39.5 ± 2.5	53.6 ± 0.2	24	mining	293	17.8 ± 0.8	12.3 ± 0.3
5	mining	4490	14.3 ± 1.4	8.8 ± 0.1	25	mining	223	19.6 ± 2.6	17.3 ± 0.1
6	mining	491	17.0 ± 0.7	22.8 ± 0.6	26	mining	494	17.8 ± 2.5	15.5 ± 0.1
7	mining	207	18.6 ± 4.2	25.7 ± 0.4	27	mining	738	11.1 ± 1.2	13.4 ± 3.5
8	mining	182	26.4 ± 2.6	32.9 ± 0.2	28	mining	777	4.3 ± 0.9	0.0 ± 0.0
9	mining	990	48.2 ± 3.6	73.1 ± 0.6	29	mining	943	2.9 ± 0.3	0.1 ± 0.0
10	mining	829	49.2 ± 3.1	74.3 ± 1.3	30	mining	898	1.9 ± 0.3	0.1 ± 0.0
11	mining	379	51.1 ± 3.2	53.2 ± 0.5	31	mining	668	3.5 ± 0.4	0.0 ± 0.0
12	mining	837	11.4 ± 0.5	18.2 ± 2.7	32	railway corridor	981	35.9 ± 1.9	54.3 ± 2.5
13	SRM	601	42.3 ± 2.1	53.9 ± 4.1	33	railway corridor	246	44.6 ± 4.2	47.0 ± 2.1
14	SRM	1510	41.5 ± 2.4	41.8 ± 1.7	34	railway corridor	108	23.5 ± 2.6	27.0 ± 0.8
15	SRM	879	16.2 ± 0.6	14.5 ± 0.2	35	railway corridor	184	22.8 ± 2.5	11.9 ± 0.1
16	orchard	322	26.1 ± 2.0	18.8 ± 0.3	36	cattle dip	965	21.5 ± 2.1	9.0 ± 0.4
17	orchard	462	34.9 ± 3.0	16.1 ± 0.4	37	mining	573	6.4 ± 0.4	3.5 ± 0.4
18	orchard	401	20.7 ± 3.2	18.0 ± 0.2	38	mining	583	14.0 ± 0.3	21.3 ± 0.2
19	orchard	422	34.7 ± 2.6	27.9 ± 0.8	39	gossan	239	20.2 ± 2.6	12.4 ± 0.6
20	orchard	340	32.8 ± 3.5	35.4 ± 1.9	40	cattle dip	313	28.8 ± 2.4	36.5 ± 1.3

558 ^a Determined by Instrument Neutron Activation Analysis; SD = standard deviation; SRM = Standard Reference Material

559

560

561 Table 2. Model parameters (slope, intercept and R²) and As RBA prediction accuracy under
 562 the various regression model development scenarios.

Model	Model Parameters			RBA Prediction Accuracy			
	# soils	Slope (± SE)	Intercept (% ± SE)	R ²	Mean AE (%)	Median AE (%)	Range AE (%)
Training data	31	0.67 ± 0.06	7.1 ± 1.8	0.83	4.9	4.6	0.3 – 17.0
Validation data	9	0.56 ± 0.15	10.3 ± 4.5	0.73	4.9	5.4	0.1 – 15.7
All data	40	0.65 ± 0.05	7.8 ± 1.6	0.81	4.9	4.8	0.0 – 16.7

563 **SE = standard error; AE = absolute error**

564

565

566

567 Table 3. Results of independent model validation.

Soil ID	IVBA (%)	Predicted RBA - model (%)	Observed RBA - mice (%)
32	54.3	43.5	35.9
33	47.0	38.6	44.6
34	27.0	25.2	23.5
35	11.9	15.1	22.8
36	9.0	13.1	21.5
37	3.5	9.4	6.4
38	21.3	21.4	14.0
39	12.4	15.4	20.2
40	36.5	31.6	28.8
		Mean AE (%)	5.5

568

569

570

571 **LEGENDS**

572 **Figure 1.** Comparison of between-laboratory reproducibility for in vitro bioaccessibility
573 method.

574 **Figure 2.** Results of fitting the linear regression model for the prediction of As relative
575 bioavailability using SBRC in vitro assay. The 31 training data points are shown as circles
576 and the 9 validation data points are shown as triangles. The overall fitted model was RBA
577 (%) = 0.65 * IVBA (%) + 7.8, with dotted lines representing the 95% prediction intervals.

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

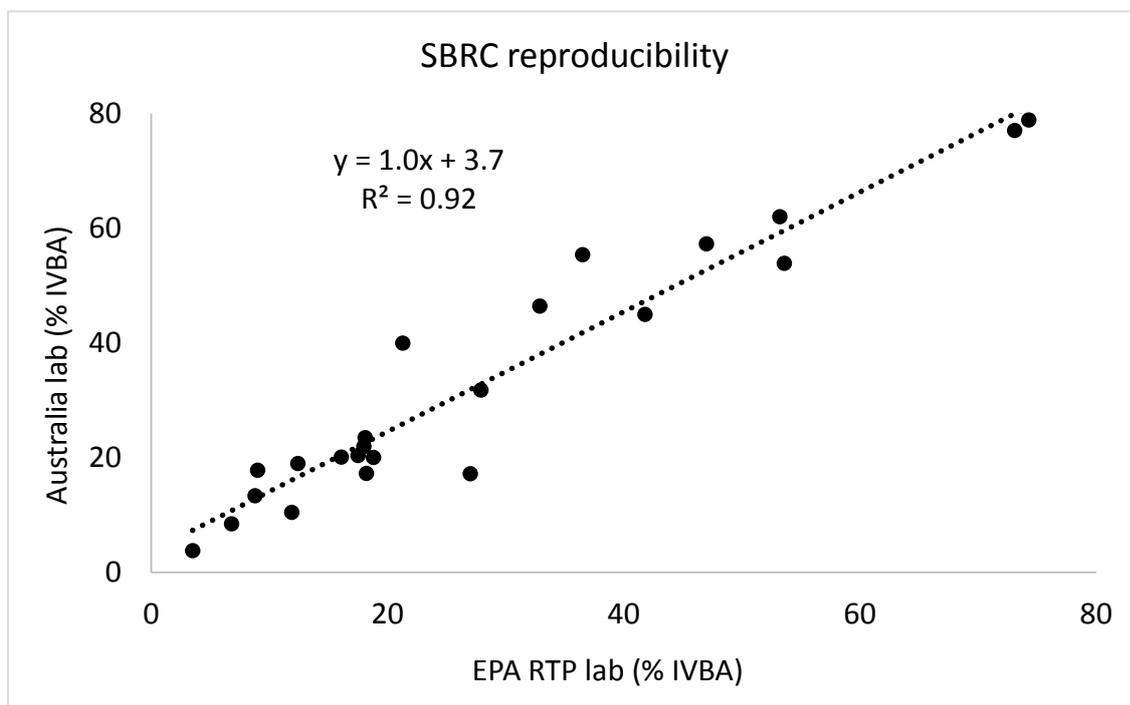
593

594

595

596 **Figure 1.**

597



606

607

608

609

610

611

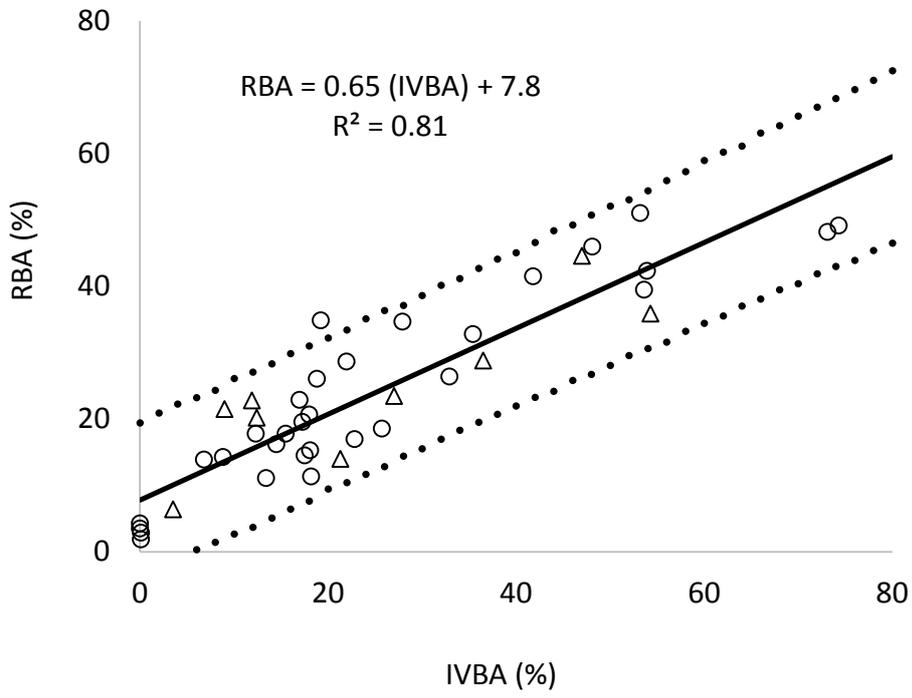
612

613

614

615

616 **Figure 2.**



617

618

619

620

621

622

623

624

625

626

627

628

629

630

631 **TOC ART**

632



633

634

635

636

637

Supporting Information

Independent data validation of an in vitro method for prediction of relative bioavailability of arsenic in contaminated soils

Karen D. Bradham^{1*}, Clay Nelson¹, Albert L. Juhasz², Euan Smith², Kirk Scheckel³, Daniel R. Obenour⁴, Bradley W. Miller³, David J. Thomas¹

¹ United States Environmental Protection Agency, Office of Research and Development, Research Triangle Park, NC 27711, USA.

² Centre for Environmental Risk Assessment and Remediation, University of South Australia, Mawson Lakes, SA 5095, Australia.

³ United States Environmental Protection Agency, Office of Research and Development, Cincinnati, OH 45224, USA.

⁴ North Carolina State University, Department of Civil, Construction, and Environmental Engineering, Raleigh, NC 27695, USA.

*Corresponding Author: bradham.karen@epa.gov, 109 TW Alexander Drive, MD-D205-05, Research Triangle Park, North Carolina; phone: (919) 541-9414; fax: (919) 541-3527.

Supporting Information Contents (7 pages)

	Page
Soil arsenic speciation determination	S2
Mouse assay methodology	S2
As IVBA determination	S3
Bayesian Model formulation	S3
Table SI-1 Test soil elemental concentrations and soil properties	S5
Table SI-2 Test soil speciation data	S6
References	S7

Soil arsenic speciation determination

A subset of test soils (soils 1-27) were characterized for As speciation using the Materials Research Collaborative Access Team's (MRCAT) beamline 10-ID, Sector 10 at the Advanced Photon Source (APS), Argonne National Laboratory (ANL), Argonne, IL. The electron storage ring operated at 7 GeV in top-up mode. A liquid nitrogen-cooled double-crystal Si(111) monochromator was used to select incident photon energies and a platinum-coated glass mirror was used for harmonic rejection. The beam energy was calibrated by assigning the first derivative inflection point of the K-edge of sodium arsenate (11874 eV). Three As K-edge (11867 eV) X-ray absorption spectroscopy (XAS) spectra were collected in lifetime corrected fluorescence mode (16-element solid state Ge detector, Canberra) and transmission mode with an ionization chamber at room temperature for every soil and reference sample. Data analysis was conducted using IFEFFIT software.¹ Replicate scans for each sample were merged, then normalized, and converted into k space. A principal component analysis coupled with linear-combination fitting (LCF) was used to identify the major arsenic species in the samples. LCFs were performed using the first derivative of the XANES spectra from reference standards to arsenic phases in the soil samples over a fit range of -20 to 130 eV. Reference materials for LCF, based on principal component analysis, included arsenate (As(V)) and arsenite (As(III)) sorbed to ferrihydrite, scorodite [Fe(As^VO₄)], schneiderhöhnite [Fe²⁺Fe³⁺₃(As^{III}₅O₁₃)], realgar (AsS), lollingite (FeAs₂) and arsenopyrite (FeAsS). Data for LCF fits reveal As speciation in each soil as ratios of these mineral forms.

Mouse assay methodology

Arsenic RBA was determined using an in vivo mouse model performed in four- to six-week-old female C57BL/6 mice. The basal diet for mouse assays was powdered AIN-93G purified rodent diet² obtained from Dyets (Bethlehem, PA). The As concentration in the basal diet was below the INAA detection limit. Based on this detection limit and measured diet consumption, As dosage from ingestion of basal diet was less than 30 µg/kg/day. Amended diets were prepared by blending of test soils or SRMs with basal diet. For test soils or SRMs, the soil:diet ratio was typically 1% (w/w). Arsenate (As^V)-amended diets were prepared by addition of sodium arsenate

heptahydrate (Sigma, St. Louis, MO) to powdered AIN-93G purified rodent diet. Mice were housed in an American Association for the Accreditation of Laboratory Animal Care-accredited facility and animal procedures were approved by the Institutional Animal Care and Use Committee of the National Health and Environmental Effects Research Laboratory.

Mice had free access to amended diet and tap water for 9 days, with urine and feces collection and measurement of food consumption performed daily for 10 days. For sample collection and data analysis, the unit of observation was the cage (i.e., combined excreta of three mice). Typically, an assay included four cages of animals (12 mice) that received the same amended diet. Urine and feces from each individual cage were pooled over the course of the assay and processed for arsenic analysis by INAA.

As IVBA determination

Arsenic IVBA was determined using the SBRC in vitro method. In vitro assays included addition of 1 g test soil to 100 mL gastric fluid consisting of 0.4 M glycine at pH 1.5 in a 125-mL high-density polyethylene bottle and rotated end-over-end in a water bath at 37°C for 1 hr. All samples were extracted in duplicate or triplicate. Quality control standards, including reagent blanks, blank spikes, sample matrix spikes, and NIST 2710A SRM, were included with each batch (12 bottles per batch). All QC sample observations were within allowable quality control limits as defined by USEPA Method 9200.2-86.^{3,4} In vitro extraction solutions were refrigerated at 4°C for preservation and analyzed by Inductively Coupled Plasma-Mass Spectrometry (USEPA SW-846 Method 6020).⁴ Quality control standards for the ICP-MS were included with all analyses as described in USEPA SW-846 Method 6020. All soils tested in the IVBA protocol were identical to those administered to mice in the in vivo studies and used in the mineralogy studies described above.

Bayesian model formulation

In the Bayesian model formulation, each RBA replicate $y_{i,j}$ of soil type i and replicate number j is assumed to come from a normal distribution:

$$y_{i,j} \sim N(\tilde{y}_i, \tilde{\sigma}_i) \quad (1)$$

where \tilde{y}_i is the true (unknown) RBA value for soil type i , and $\tilde{\sigma}_i$ is the standard deviation of the replicates. Based on a preliminary analysis of the replicates, $\tilde{\sigma}_i$ appears to be proportional to \tilde{y}_i , such that it can be approximated as:

$$\tilde{\sigma}_i = \bar{y}_i \beta_\sigma \quad (2)$$

where \bar{y}_i is the mean RBA of the replicates of soil type i , and β_σ is a scaling parameter determined by the model. The IVBA replicates, $x_{i,j}$, are modeled in the same fashion as the RBA replicates. RBA and IVBA can then be related to each other as follows:

$$\tilde{y}_i = f(\tilde{x}_i) + \epsilon \quad (3)$$

where $f(\tilde{x}_i)$ is the prediction of RBA using IVBA; and ϵ is the normally-distributed prediction error, i.e., $\epsilon \sim N(0, \sigma_\epsilon)$. In this study, we considered the model form of $f(\tilde{x}_i)$:

$$f(\tilde{x}_i) = a + b\tilde{x}_i \quad (4)$$

In eq 4, the relationship between RBA and IVBA is a linear function with intercept parameter a and slope parameter b .

Model parameters were estimated through Bayesian inference, using WinBUGS software⁵ called from R⁶ via R2WinBUGS.^{7,8} The MCMC sampling was performed in three parallel ‘chains’ of up to 25,000 samples each, and the first half of each chain was removed as a ‘burn-in period’.⁹ The remaining chain portions were then thinned to 2500 samples each to reduce autocorrelation, and checked to ensure that they had converged on equivalent posterior parameter distributions. Convergence was evaluated using the \hat{R} statistic,¹⁰ and when \hat{R} is less than 1.1 one for all model parameters, convergence is considered achieved.

Test soil elemental concentrations and properties

Table SI-1. Select physico-chemical properties of soils (< 250 µm particle size fraction) in this study.

Soil ID	pH	Soil totals (mg kg ⁻¹)				
		As ^a	Al ^b	Fe ^b	Mn ^b	P ^b
1	7.5	244	24040	42580	812	1690
2	6.4	173	19970	60650	688	1400
3	5.5	6900	15970	139410	966	1090
4	2.3	280	4482	77880	17	62
5	2.7	4490	14400	140130	435	948
6	6.9	491	12950	42830	548	1800
7	6.7	207	21520	39380	1150	2320
8	6.9	182	21930	22940	441	780
9	6.4	990	12520	17650	517	1060
10	6.5	829	10220	15670	619	631
11	5.1	379	10660	14380	264	178
12	7.1	837	14740	276280	2750	1310
13	5.0	601	19090	31250	9320	1020
14	4.0	1510	10860	38590	1820	989
15	5.9	879	4136	36050	2600	6740
16	6.2	322	37690	26560	232	1170
17	6.2	462	66850	46650	854	1410
18	5.6	401	53480	40960	1350	1440
19	5.9	422	47000	35210	1540	1770
20	5.7	340	20930	23170	1670	1640
21	5.6	396	12750	20350	466	1280
22	5.2	197	27360	23440	303	50
23	6.4	884	28440	29960	738	51
24	6.5	293	40120	35510	407	34
25	5.4	223	34150	28330	350	45
26	6.0	494	28010	37840	567	36
27	6.6	738	19740	31190	877	43
28	2.9	885	2740	174230	0.5	194
29	3.1	566	678	106440	6	46
30	3.3	802	1700	199070	2	159
31	3.5	552	3680	179000	13	184
32	8.3	981	11350	31950	766	861
33	8.8	246	22980	20480	373	197
34	7.8	108	26450	33660	279	157
35	6.4	184	20710	31240	487	16
36	5.7	965	72080	97660	1520	24
37	6.6	573	9060	30350	421	66
38	7.6	583	4700	25490	378	25
39	8.6	239	10700	23860	232	4
40	5.2	313	18860	19150	1000	16

Table SI-2. Speciation data for soils used in this study. Values are given in percent (%).

Soil ID	As(V)-Oxygen Coordination		As(III)-Oxygen Coordination		As-Sulfide Coordination			R-factor ¹
	As(V) Sorbed to Oxides	Scorodite	As(III) Sorbed to Oxides	Schneiderhöhnite	Realgar	Loellingite	Arsenopyrite	
1	97			3				0.005
2	87		13					0.001
3	51	16	13				21	0.003
4	65	19		6		11		0.004
5	61	29		5		5		0.003
8	58		20	23				0.003
9	54	22	24					0.001
10	80	5	15					0.001
11	55	10	13		22			0.002
12	14		14	12	38	9	13	0.006
13	95	5						0.001
14	75	11	14					0.001
15	77		12				11	0.002
16	100							0.001
17	100							0.001
18	100							0.002
19	100							0.003
20	100							0.002
21	90		10					0.002
22	44		4	16	36			0.005
23	77				23			0.001
24	54			23	23			0.002
25	79			10	11			0.001
26	71			14	14			0.001
27	78				22			0.001

¹ R-factor = $[\sum((\text{data-fit})^2)]/[\sum(\text{data}^2)]$

References

- (1) Ravel, B.; Newville, M. ATHENA, ARTEMIS, HEPHAESTUS: Data analysis for X-ray absorption spectroscopy using IFEFFIT. *Journal of Synchrotron Radiation*. **2005**, 12: 537-541.
- (2) Reeves, P. G.; Nielsen, F. H.; Fahey, G. C. AIN-93 purified diets for laboratory rodents: Final report of the American Institute of Nutrition *ad hoc* writing committee on the reformulation of the AIN-76A rodent diet. *J. Nutr.* **1993**, 123: 1939–1951
- (3) Kelley, M. E.; Brauning, S. E.; Schoof, R. A.; Ruby, M. V. *Assessing Oral Bioavailability of Metals in Soil*; Battelle Press: Columbus, 2002.
- (4) USEPA. *Compilation and review of data on relative bioavailability of arsenic in soil*. OSWER 9200.1-113, U.S. Environmental Protection Agency, 2012.
- (5) Lunn, D. J.; Thomas, A.; Best, N.; and Spiegelhalter, D. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing*. **2000**, 10:325-337.
- (6) R Development Core Team *R: A language and environment for statistical computing*; R Foundation for Statistical Computing: Vienna, Austria, 2008. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- (7) Sturtz, S.; Ligges, U.; and Gelman, A.E. R2WinBUGS: a package for running WinBUGS from R. *Journal of Statistical Software*. **2005**, 12(3): 1-16.
- (8) Gelman, A.; Hill, J. *Data analysis using regression and multilevel/hierarchical models*; Cambridge University Press: Cambridge, U.K., 2007.
- (9) Gelman, A., Carlin, J.B.; Stern, H.S.; and Rubin, D.B. *Bayesian data analysis*; Chapman & Hall/CRC: Boca Raton, Florida, 2004.
- (10) Brooks, S. P.; Gelman A.E. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*. **1998**, 7(4): 434-455.