

Soil Metal Bioavailability Sample Planning Tool: Read Me File

Version: 1 (1/19/22)

This tool has been developed to inform sample planning for sites with arsenic (As) or lead (Pb) contaminated soils where in vitro bioaccessibility (IVBA) data will be used to estimate soil metal relative bioavailability (RBA) to make site-specific adjustments to measured total soil metal concentrations. The tool 1) simulates the estimation of RBA-adjusted soil arsenic concentration at the scale of a Decision Unit (DU) based on user-specified sampling protocols and assumptions about the DU, and 2) estimates the probability of making a Type 1 or 2 error, where:

- **Type 1 error** = we reject the null hypothesis that the RBA-adjusted mean soil As concentration \geq Action Level (AL), when the null hypothesis is true
- **Type 2 error** = we accept the null hypothesis that the RBA-adjusted mean soil As concentration \geq Action Level (AL), when the null hypothesis is false

In this analysis, a Type 1 error is a false compliance decision error and a Type 2 error is a false exceedance decision error. A false compliance decision error occurs if it is concluded that the RBA-adjusted Exposure Point Concentration (EPC) is less than the AL, when it is actually greater than the AL. This outcome is also referred to as a false rejection error (U.S. EPA, 2006). A false compliance decision error could result in underestimating risk at the site and/or not taking an action when action is needed to reduce risk. A false exceedance decision error occurs if it is concluded that the EPC exceeds the AL, when it is actually less than the AL. This outcome is also referred to as a false acceptance error (U.S. EPA, 2006). A false exceedance decision error could result in overestimating risk at the site and/or taking action at the site to reduce risk when no action is needed.

EPA guidance recommends a target of a < 5% chance of making a Type 1 error and < 20% chance of making a Type 2 error (U.S. EPA 2020).

The bioavailability sampling guidance tool includes several user-specified inputs including, but not limited to:

- The DU-specific Action Level (AL)
- Discrete or composite sampling
 - For composite sampling, the number of sample increments per composite

- The number of samples analyzed for total soil metal concentration and IVBA
- Assumptions about the level of contamination at the DU, specified in terms of how close the DU's true bioavailable soil metal concentration is to the AL
- Assumptions about variability in total metal concentration and % RBA across the DU, specified in terms of the coefficient of variation (CoV), where:
 - $\text{CoV} = \text{standard deviation} / \text{mean}$
- Assumptions about the distribution in total metal concentration and % RBA across the DU, specified in terms of normal, log-normal or uniform distributions
- Assumptions about the mean % RBA for the DU
- Assumptions about analytical error in measurement of total soil metal concentration
- Assumptions about regression model error in predicting RBA from IVBA as described in Diamond et al. 2016.

Table 1. Explanation of user-selected inputs to tool.

Input	Options	Explanation
Contaminant		
Select contaminant	As, Pb	The tool is currently developed for use with Pb or As-contaminated soils where bioavailability will be used to adjust the total soil metal concentration.
Enter site-specific action level (mg/kg)	Custom field	The tool is coded so that the user inputs the Decision Unit's (DU) 'un-adjusted' (for RBA) action level here. The total soil metal concentration is adjusted by %RBA. This value -- the bioavailable soil metal concentration (mg/kg) -- is then compared to the user-input action level.
Sampling Protocol		
Sample aggregation	Discrete, Composite	Select whether a discrete or composite sampling protocol will be used.
Increments per composite	Custom field	If composite is selected for sample aggregation, input the number of sample increments per composite. The tool is currently limited to only simulate a single value here (i.e., it can't simulate scenarios where some composites are made up of X increments, and others Y increments).
# of samples to be analyzed for total metal concentration	Custom field	The number of discrete or composite samples measured for total soil metal concentration, typically by EPA Method 6010 (ICP-OES) or 6020 (ICP-MS).
# of samples to be analyzed for % IVBA	Custom field	The number of discrete or composite samples measured for % IVBA by EPA Method 1340.
Base final bioavailable metal concentration calculation on:	Mean, upper 95% CL of mean	The user selects whether the DU's [bioMetal] (mg/kg) should be calculated using the mean or upper 95% CL in [bioMetal] (mg/kg) across the samples analyzed in the DU.

Decision Unit Assumptions		
Assumed level of soil contamination to simulate (expressed in terms of % below/above the action level)	-25%, 25%, Custom field	In order for the tool to estimate Type 1 or 2 error, the DU's true [bioMetal], while an unknown value, must be assumed. Typically, the closer the DU's assumed true [bioMetal] is to the AL, the more samples that must be analyzed to meet Type 1 or 2 error goals. In order to estimate Type 1 error rates, it must be assumed that the DU's true [biometal] is > AL (a + value must be input here). To estimate Type 2 error rates, it must be assumed that the DU's true [bioMetal] is < AL (a - value must be input here).
Total metal concentration data distribution:	Log-normal, normal	The assumed shape of the distribution in total metal concentration across the DU. The default input is log-normal, but the user can also specify a normal distribution.
Total metal concentration coefficient of variation (CoV):	0.5, 1, 3, Custom field	CoV is used to quantitatively estimate assumed variability in total metal concentration across the DU. The higher the CoV, the greater the assumed variability in totals across the DU.
RBA data distribution:	Normal, uniform, log-normal,	The assumed shape of the distribution in % RBA across the DU. The default input is normal, but the user can also specify a uniform or normal distribution.
RBA CoV:	0.05, 0.15, 0.3, Custom field	CoV is used to quantitatively estimate assumed variability in % RBA across the DU. The higher the CoV, the greater the assumed variability in % RBA across the DU.
RBA mean:	60%, Custom field	As previously noted, the DU's true [bioMetal] must be assumed. The tool calculates the true [bioMetal] based on two inputs: 1) the user input AL, and 2) the assumed level of soil contamination to simulate (expressed in terms of % below/above the action level). For example, when an AL of 400 ppm and assumed level of soil contamination 25% above the AL are input, the tool calculates the assumed [bioMetal] as $400 \times 1.25 = 500$ ppm. Since, $[\text{bioMetal}] = [\text{total soil}] \times \% \text{RBA}$, inputting the assumed %RBA allows the tool to calculate the assumed [total soil] to enable the simulation to be run.
Simulation Parameters		
Simulate measurement error for total concentration	True, false	When 'true' is selected, the tool assumes analytical error in the measurement of total soil metal concentration. The measurement + error values are modeled as a normal distribution centered on the true total soil metal concentrations so that 95% of the probability distribution falls within $\pm 10\%$ of the true value, i.e., with standard deviation = $0.05 \times \text{true mean}$. When 'false' is selected, the tool assumes analytic measurement of total soil concentration = the true total soil metal concentration.
Simulate IVBA model error	True, false	When 'true' is selected, the tool incorporates uncertainty in prediction of % RBA from % IVBA associated with the in-vivo / in-vitro correlation models that have been validated for Pb and As. Prediction error is drawn from a normal distribution with standard deviation calculated from the published 95% prediction interval, i.e., standard deviation = $(0.5 \times 95\% \text{ prediction interval range}) / 1.96$. When 'false' is selected, the tool assumes that the estimated %RBA based on measurement of %IVBA equals the samples true %RBA.
Number of simulations	Custom field	The user can select the number of times the simulation is run. The higher the number of simulations, the more precise Type 1 or 2 error estimates are when the analysis is repeated, but the longer the run time. 5,000 simulations is the default value.
Optional Advanced Tool-Automated Analyses		
Select advanced analysis type	None, Vary Sample Size,	If 'None' is selected, no advanced tool-automated analyses will be simulated.

	Vary Contaminant Level	<p>If 'Vary Sample Size' is selected, the tool will re-simulate the analysis by simulating n+1 samples analyzed for total soil metal concentration and % IVBA, up to the user-input maximum number of samples to simulate -- input via a dropdown menu that will appear when this option is selected. For example, if the user initially selects 3 samples be analyzed for totals and 2 samples be analyzed for % IVBA, the tool will re-simulate the analysis by adding one additional sample to each analysis (i.e., n=4 for totals and 3 for IVBA; n=5 for totals and 4 for IVBA; n= 6 for totals and 5 for IVBA; ...) up to the user-input maximum number of samples to simulate. Type 1 or 2 error estimates will be output to the Sample Simulation Tab on the right half of the screen via a graph that shows how Type 1 or 2 error estimates vary based on the number of samples analyzed for totals and %IVBA.</p> <p>If 'Vary Contaminant Level' is selected, the tool will run a series of simulations where the assumed level of soil contamination (expressed in terms of % below/above the action level) is automatically adjusted based on the user-input minimum (%), maximum (%), and # of simulation intervals – input via dropdown menus that will appear when this option is selected. For example, if the user inputs: minimum (%) = 10%; maximum (%) = 90%; and # simulation intervals = 10, the tool will run a series of simulations starting with an assumption that the DU's true [bioMetal] is 10% above the AL, and then proceeding all the way up to 90% above the AL by adding 10% to the assumed true [bioMetal] (i.e., scenarios of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% and 90% above the AL will be simulated). Type 1 or 2 error estimates will be output to the Sample Simulation Tab on the right half of the screen via a graph that shows how Type 1 or 2 error estimates vary based on the assumed level of soil contamination.</p> <p>Note: The tool is currently coded in a way that requires that the range in contaminant level explored must all be either positive (+) (to evaluate variability in Type 1 error estimates) or negative (-) (to evaluate variability in Type 2 error measurements) relative to the action level. It can't correctly compute inputs that span – and + range.</p>
--	------------------------	---

References

Diamond GL; Bradham KD; Brattin WJ; Burgess M; Griffin S; Hawkins CA; Juhasz AL; Klotzbach JM; Nelson C; Lowney YW; Scheckel KG; Thomas DJ. (2016) Predicting oral relative bioavailability of arsenic in soil from in vitro bioaccessibility, Journal of Toxicology and Environmental Health, Part A, 79:4, 165-173, DOI: [10.1080/15287394.2015.1134038](https://doi.org/10.1080/15287394.2015.1134038)

U.S. EPA (U.S. Environmental Protection Agency). (2006) Guidance on Systematic Planning Using the Data Quality Objectives Process. EPA QA/G-4. U.S. Environmental Protection Agency, Office of Environmental Information: Washington, DC. EPA/240/B-06/001. February. Available online at: <https://www.epa.gov/sites/production/files/2015-06/documents/g4-final.pdf>.

U.S. EPA (U.S. Environmental Protection Agency). 2020. Guidance for Sample Collection for *In Vitro* Bioaccessibility Assay for Arsenic and Lead in Soil and Applications of Relative Bioavailability Data in Human Health Risk Assessment. OSWER 9200.3-100a. U.S. Environmental Protection Agency, Office of Environmental Information: Washington, DC. EPA/240/B-06/001. September. Available online at: <https://www.epa.gov/superfund/soil-bioavailability-superfund-sites-guidance>