489790

EPA RESPONSES TO COMMENTS ON THE DRAFT DATA REPORT

VASQUEZ BOULEVARD & I-70 RESIDENTIAL SOILS SUPPLEMENTAL INVESTIGATION PHYSICO-CHEMICAL CHARACTERIZATION OF SOILS

JANUARY 25, 1999

1.0 INTRODUCTION

The USEPA, Region VIII, is currently performing investigations to characterize the nature and extent of arsenic and lead contamination in soil in a residential area located near Vasquez Boulevard and I-70 (VBI70). A draft report (termed the Phase 1A Draft Report) summarizing some aspects of EPA's investigations into soil samples from this area was provided to members of the VBI70 Workgroup for review and comment on November 12, 1998. This document summarizes review comments received from workgroup members and provides responses to those comments.

2.0 GENERAL COMMENTS

General Comment
Several commentors questioned whether data for metals in addition to arsenic, cadmium, lead and zinc were measured as part of this investigation.

General Response
All analyses performed by the University of Colorado on behalf of the USEPA were performed in accordance with the Quality Assurance Project Plan (QAPP) prepared and approved for this investigation. The QAPP required quantification of concentrations for four metals: arsenic, cadmium, lead and zinc. Geochemical speciation and particle size distribution analysis were stipulated for lead and arsenic only. Data on the concentration and physical form of other metals (e.g., indium, selenium, thallium and mercury) were not collected.

2.0 COMMENTS FROM THE CITY AND COUNT OF DENVER - DEPARTMENT OF ENVIRONMENTAL HEALTH

Comment 1:
Section 2.2. Were the samples analyzed for chemical speciation from the bulk or fine subset? A brief summary of sample preparation would be helpful. It also would be helpful to have a qualitative discussion of the results of speciation. Factors that could be discussed include particle shape (rounded blebs versus blocky particles), whether different phases of metals appear to be co-located, and whether other trace metals were seen in numbers too small to be quantified as ppm.

Response: The samples analyzed for chemical speciation were composed of the fines fraction. As stated in the Quality Assurance Project Plan and the associated Standard Operating Procedures (SOPs), samples were prepared by air drying and then sieving through a 250 um mesh sieve. The initial data report from the geochemical laboratory did not contain any qualitative information on the physical appearance of the particles. Such data have been requested and will be included in the final report. By definition, each particle is assigned a unique phase, so different phases are not co-located in the same grains. As noted above, other trace metals were not quantified.

Comment 2:
Section 2.3. The text would be improved by a discussion of methodologies used that addresses the questions raised in item 3 below. While the earlier narrative concludes that there is little difference between bulk and fine arsenic concentrations, this section concludes that the arsenic-bearing particles are predominately found in particles between <5 to 49 Fm in size. An explanation or discussion regarding this seemingly contradictory situation would be helpful.

Response: The reason that the fine fraction does not become enriched in arsenic (even though essentially all of the arsenic grains are small) is currently being investigated. It is suspected that some of the fine grains of arsenic exist in association with larger particles that do not pass through the 250 um sieve.

Comment 3:
Figures 1A through 1D. These figures graphically represent the comparison of metals concentrations in bulk versus fine samples. The text states that these figures show little difference in concentrations in any of the four chemicals for bulk versus fines. The text also states that the respective 90% confidence intervals are shown. The tables show dotted lines relatively close to the line representing bulk versus fine, which we initially assumed represented the 90% confidence interval. What is the significance of the dotted lines? Our assumption is that the 90% confidence interval is intended to represent 5% chance of error above the line and 5% below. Is this correct? How was the 90%

confidence interval calculated? Were a large number of points eliminated as outliers? What is the standard deviation of the data? It appears that a more broadly defined 90% confidence interval could results in estimations of significant variation between fine and bulk.

Response: The 90% confidence intervals shown in Figures 1A through 1D and Table 2.1 reflect uncertainty in the best-fit linear regression line through the data. These confidence bounds were calculated using a commercial statistical package called TableCurve. The bounds of the confidence interval around the best fit straight line does not imply that 90% of all data points must lie within the confidence interval. No data points were removed as outliers in calculation of the regression line for Figures 1A through 1D.

Comment 4:
Figures 1A through 1D. It would be helpful to provide tables that show the supporting data by bulk and fine concentrations.

Response: A table will be added to the appendix of the final report as suggested.

Comment 5:
Table _-_. Is this table representing bulk or fine concentrations?

Response: The table referred to represents lead and arsenic concentrations in the fines fraction. Please note that this table was provided for information only and will not be included in the final report.

Comment 6:
Table-Supporting Calculations for Relative Arsenic Mass by Phase. The table includes a heading that appears to state that the units are total particle counts. If so, shouldn't the values be whole numbers of particles counted? What are the units in this table and how are these values used? The values are the same as those shown in the next table labeled Number of Particles in Each Phase, which are presented as round numbers. How do these tables differ in information presented?

Response: The title of the tables should not have been labeled as the "Number of Particles in Each Phase". The numbers in each table are intermediate values used in calculation of the metal mass found in each phase. This value is not a whole number, and was rounded for convenience only. Strictly, the units for this value are μg metal per um^2. This number weights the length of particles found for each phase, then multiplies the length by the density of the particles and the mass fraction of arsenic found in that phase. The ratio of this value for each phase to the sum of the values for all phases is a rough estimate of the fraction of the total metal that is present in each phase. Note that this table was provided for information only and will not be included in the final report.

Comment 7:
Table-Supporting Calculations for Relative Arsenic Mass by Phase, Number of Particles in Each Phase. Were other trace particles such as indium, selenium, thallium, or mercury counted? Were particles in small quantities (less than 50 particles) counted? At the Globe site, samples contained small numbers of trace metals that may be indicative of the contaminant source. The presence or absence of trace metals in very small quantities is important information that may be telling in determining contaminant source.

Response: See General Comment.

3.0 COMMENTS FROM THE STATE OF COLORADO

Comment 1:
The scope of this report doesn't include any discussion or proposal of how this data would be applied to the site risk assessment. However, because a subset of 10% of all surface soil samples collected was used to characterize particle size distribution and mineral phase, it will be important to determine how representative this subsample is of the site as a whole when applying this information to a site-wide risk assessment or risk characterization.

Response: The purpose of this report is to present the results from the investigations performed in the Phase 1A - Physico-Chemical Characterization of Soils. Issues regarding the representativeness of the data and the proposed uses of the data in the risk assessment will be discussed with the workgroup as the risk assessment process proceeds.

Comment 2:
Table 1 in section 2.4 of the data report ("In Vitro Bioaccessibility") shows the results of the bioaccessible fraction of each sample tested using an in vitro bioaccessibility test. Because in vitro testing for arsenic lacks method validation or a basis for comparison to other in vivo models, such as swine feeding studies, this data is potentially unreliable and misleading. Table 1 should not be incorporated into a site-specific risk assessment.

Response: EPA agrees that the results of *in vitro* bioaccessibility studies for arsenic are subject to limitations, and does not envision that the *in vitro* data, in and of themselves, will form the basis for a quantitative adjustment to the RBA values for arsenic. However, EPA believes that these data do provide valuable information that must be considered as part of the risk assessment. The proper interpretation and application of these data in the risk assessment will be discussed with the VBI70 Workgroup as the risk assessment process proceeds.

Comment 3:
The state is aware that Dr. John Drexler, who we understand provided much of the supporting data for the report, also analyzed for other metals including indium, selenium, thallium, and mercury. Neither the presence nor numbers of these other metals is documented in this report. The state believes that the presence and numbers of these other metals is critical to determining the source of arsenic contamination at the Vasquez site. Why is this information excluded from this report and when will it be available?

Response: This is not correct. See General Comment.

4.0 COMMENTS FROM THE AGENCY FOR TOXIC SUBSTANCES AND DISEASE REGISTRY (ATSDR)

Comment 1:
The statement under 2.1, last paragraph "This means that there is little difference in concentrations in any of the four chemicals for bulk versus fines" may be true for the higher concentrations of arsenic but not true for arsenic concentrations around 400 ppm, the removal action level.

> *At an XRF bulk sample concentration of 400 ppm arsenic, the corresponding concentration of arsenic in the XRF fine sample is 503 ppm. This is a significant difference when it comes to deciding whether or not a property requires immediate action to stop exposure. If one considers that the removal action level was set for 400 ppm arsenic, the concentration of arsenic in the XRF bulk sample should be 311 ppm when the concentration of arsenic in the XRF fines sample is set at 400 ppm.*
>
> $Y = 1.1685x + 35.697$
> *where y is the XRF measurement for fines and x is the XRF measurement for bulk samples*
> $400 = 1.1685x + 35.697$
> $x = 311$ *ppm*
> *XRF bulk = 311 ppm*
>
> *Therefore, EPA should reconsider the homes where the average concentration of arsenic falls between 311 ppm and 400 ppm for XRF bulk measurements to decide whether or not they should be placed on the list of properties to be remediated under EPA's removal program.*

Response: While the calculation performed by the commentor is mathematically correct, it assumes two things that may not be appropriate. First, the calculation assumes that the small differences between the bulk and fine soil fractions are meaningful. In contrast, EPA interprets the finding that all of the slopes are near 1.0 and all of the slopes are close to zero to indicate that there is little or no authentic difference between bulk and fine. Second (and more importantly), the comment assumes that 400 ppm is some sort of absolute threshold for soil removal to protect human health. In contrast, the range of values EPA identified as warranting removal is 400-900 ppm, and 450 ppm was selected as a "removal action level" simply to be maximally protective. Therefore, even if the true concentration of arsenic in fines is slightly underestimated by measurements based on bulk samples, any remaining soils in the 400-500 ppm range are not believed to pose an acute health hazard to residents. Naturally, any soils in this range may well be subject to remedial actions in the future.

Comment 2:
The following sentence in the same paragraph does not seem consistent with the data presented in the table under 2.1. Specifically, the text states that the intercept for zinc (i.e., 18) is significantly greater than zero and that the intercept for arsenic (i.e., 36) is nearly significant. It seems that the opposite is more logical based on the numbers.

Response: Whether or not a calculated intercept is statistically different from zero depends not only on the value but also on the standard deviation of the value. Thus, it is not illogical or unexpected that a value of 18 (with a relatively small standard deviation) may be significantly different from zero while a value of 36 (with a relatively large standard deviation) is not. For clarity, Table 2.1 has been revised to include the p values associated with each intercept. This change will be included in the final report.

Comment 3:
In the same paragraph, consider using another term rather than "small negative bias." It's a difficult concept to understand. My next point is to consider whether or not it's an important concept to discuss since in the case of lead, cadmium, and zinc it does not affect the comparison of bulk and fine samples. It may be an issue to discuss for arsenic but the text already explains the significance of the slope in comparing the two samples. Maybe the issue is the inherent ability of the XRF to measure arsenic in large particles versus arsenic in small particles. But again, the text already addresses that issue by looking at the slope formula itself.

Response: The XRF method used to measure metal concentrations in soil samples involves grinding the samples to a very fine powder before analysis. Therefore, any difference between bulk and fine samples is not attributable to differences in measurement efficiency due to differences in particle size. The phrase "small negative bias" is a standard statistical phrase. It means that, if the results of the regression analysis are taken literally, then concentrations of metals in bulk samples will tend to slightly underestimate the concentration in fines at low concentration values. Text will be added to make this point more clearly in the final report.

Comment 4:
I would like to see the calibration curve for arsenic to determine the range of arsenic concentrations where the curve is linear. Concentrations of arsenic outside the linear range should not be part of determining the slope of the fines versus bulk comparison. That may be the reason that the intercept is not zero. Of course, another reason might be the ability of the XRF to measure arsenic in different size particles. Also, the 5 times rule should be applied to the MDL and measurements between the MDL and 5 times the MDL should not be part of the determination of the slope because those measurements are estimated values. In looking at Figure 1A, the correlation of arsenic in fines and bulk seems to have a greater variability at the higher levels of arsenic than at the lower levels. This greater variability seems to start around 800 to 900 ppm. If that's so, that the curve is not linear above, for instance 800 ppm, then those higher levels are affecting the

slope and hence the decision about the relationship between the two sample types and whether or not 311 ppm bulk is an appropriate removal action level. Another way to make sure that measurements at either end of the distribution are not affecting the slope of the curve in the middle (for instance at 400 ppm) is to determine the slope in this case for measurements between say 150 ppm and 600 ppm. If the slope is the same then measurements at either end of the curve are not affecting the relationship. If the slope is significantly different then it's a clue to evaluate measurements at both ends of the distribution to determine whether or not they are valid and whether or not they should remain in determining the slope.

Response: At the commentor's suggestion, the sensitivity of the calibration curve was tested using the subset of data for arsenic measurements between 150 ppm and 600 ppm. The resulting slope is 1.16 which is nearly identical to the slope value obtained for the entire data set (1.17). EPA believes that the calibration curve used in quantification of arsenic concentrations is sufficiently accurate over the range of concentrations reported.

Comment 5:
Figure 2A. Consider renaming "Mass of arsenic in each phase (ppm)" to concentration of arsenic in each chemical form (or something to that effect.) I'm concerned that the phrase "mass of arsenic in each phase" is not understandable to most people.

Response: The text has been refined to include a definition and explanation of the term "mass of arsenic in each phase".

Comment 6:
The statement in the paragraph on Quality Control Issues (page 3) about the correlation of 0.95 indicating that there is little difference between the bulk and fine fraction is not the correct interpretation for a correlation coefficient. The high degree of correlation shows that over the range of measurements, the concentrations are highly correlated. That correlation could be that the concentrations are 1:1, 2:1 or 10:1. It is the slope that must be evaluated to determine whether or not there is little difference between the two types of samples.

Response: The comment is correct, and the text has been revised to make clear that inferences regarding the similarity between bulk and fine are based on the fact that the slope is close to 1, rather than the fact that the coefficient of determination (r^2) is near 1.

Comment 7:
The same paragraph makes a statement that the lower concentration of arsenic using ICP may have resulted from insufficient acid digestion of the soil matrix. More discussion is needed by the technical members of the workgroup to determine which measurement (ICP or XRF) provides the more reliable result. This becomes important in deciding which yards require immediate actions to stop exposure and which concentrations should be used to evaluate long-term exposure.

Response: EPA is currently considering ways to further investigate the apparent discrepancy between XRF and ICP, and will discuss the findings and potential implications with the workgroup.

Comment 8:
Under Quality Control Samples, 2.1, the text states that the 30% difference between ICP and XRF measurement is not significant. This conclusion is incorrect. The difference may or may not be significant depending upon which analytical method gives the most reliable measurements. It's not known at this time whether or not the difference is large enough to affect the public health significance of long-term exposure to arsenic in soil because that evaluation has not been conducted yet.

Response: Please note that the text does not state that the 30% difference is not significant. As stated in the report, the basis for this difference is unknown, but that EPA is considering ways to further investigate the apparent discrepancy between XRF and ICP. EPA will discuss the findings and potential implications with the workgroup.

Comment 9:
Here's the justification for why the 30% difference may or may not be significant.

A. If one assumes that the ICP measurement is more reliable, then one must consider the slopes generated for the ICP measurements and for the XRF measurements.

Here's the logic for what needs to be considered when testing whether or not the action level is appropriate because of the method used in analyzing the concentration of arsenic in soil. (This does not address whether or not 400 ppm arsenic is the appropriate removal action level but rather how the slopes affect using 400 ppm.)

Using 400 ppm arsenic as the removal action level, and assuming that ICP measurements are more reliable than XRF measurements, the following conversions are needed to determine the effect of measurement inaccuracies:

As ICP fines " As XRF fines " As XRF bulk

Therefore,
400 ppm As ICP fines " $y = 0.7286x - 27.073$ " $y = 1.1685x + 35.697$

$400 = 0.7286x - 27.073$ and $586.2 = 1.1685x + 35.697$
$x = 586$ ppm $x = 471$ ppm

400 ppm As ICP fines = 586 ppm As XRF fines = 471 ppm As XRF bulk.

The above logic means that the 400 ppm As in the bulk samples as a removal action level remains protective of public health or that EPA could raise the removal action level to 471 ppm.

I have not done the calculations for lead based on the previous procedure. Should the working group decide that the previous procedure is valid, the action level for lead should be reevaluated. It should go up somewhat because of the slope of 0.7126 + 6 for ICP fines and XRF fines. The slope for XRF fines and XRF bulk will not affect the conversion significantly.

Now for the other case.

B. If one assumes that XRF measurements are more reliable, then the only factor that affects using the 400 ppm arsenic as the removal action level is the previously discussed difference in the XRF measurements of bulk samples and fines samples. In this case, XRF bulk ± XRF fines, and when XRF fines = 400 ppm then XRF bulk = 311 ppm.

For lead, the concentration of lead does not change significantly between XRF fines and XRF bulk samples.

Response: With regard to the discrepancy between ICP and XRF, and the question as to which data set to use, EPA feels the following points should be noted: 1) the removal action level is set based on public health concerns and currently available data (the basis for the range of removal action levels is documented in EPA's action memorandum dated September 16, 1998) , 2) steps are being taken to resolve the apparent discrepancy between ICP and XRF, 3) in the interim, the XRF data are being used (because they are higher), 4) the decision as to which data to rely upon in the final analysis will be re-evaluated when the investigation into the discrepancy is complete, and 5) the biological significance of this difference is questionable.

Comment 10:
The technical people on the VBI70 work group should talk about how to use these slope factors and review the logic described previously. They should also discuss the public health significance of the slope for the comparison of XRF measurements of bulk and fines samples and the public health significance of the slope for the comparison of ICP and XRF measurements for the bulk samples, both as it relates to removal activity and future risk assessments for long-term exposure.

Response: EPA intends to continue the discussion about the soil data with all members of the working group, not just the "technical people".

Comment 11:
Please specify in 2.2 Speciation whether the bulk sample or the fines sample was analyzed.

Response: The speciation data were collected using samples of the fine fraction as specified in the Quality Assurance Project Plan (QAPP). The text has been revised to indicate this.

Comment 12:
Please specify in 2.3 Particle Size Distribution whether the bulk sample or the fines sample was analyzed.

Response: See above.

Comment 13:
In Figure 4A, it's unclear why there is a fraction at >250 um. If the fines sample was used, there should not be particle sizes greater than 250 um. If the bulk sample was used, then the percent of As in the >250 um should be larger, although less than 50%.

Response: The length of the particles recorded by the geochemical laboratory is the longest dimension of the particle. Because particles are not usually spherical in nature, a particle with length greater than 250 um may pass the sieve if its smallest dimension is less than 250 um.

Comment 14:
Please specify in 2.4 In Vitro Bioaccessibility whether the bulk sample or the fines sample was analyzed.

Response: The *in vitro* bioaccessability data were collected using samples of the fine fraction as specified in the QAPP. The text has been revised to indicate this.

Comment 15:
At this time, I have no comments on the in vitro bioaccessibility measurements. I will have to investigate this issue further and decide at a later date its applicability to the risk assessment methods ATSDR will use.

Response: Comment noted.