Slide 1

The Triad Approach to Better Cleanup Projects: Illustrated with the Tree Fruit Case Study

Deana M. Crumbling, M.S.
Technology Innovation Office
U.S. Environmental Protection Agency
and
Kira P. Lynch, M.S.
Innovative Technology Advocate
Seattle District, US Army Corps of Engineers

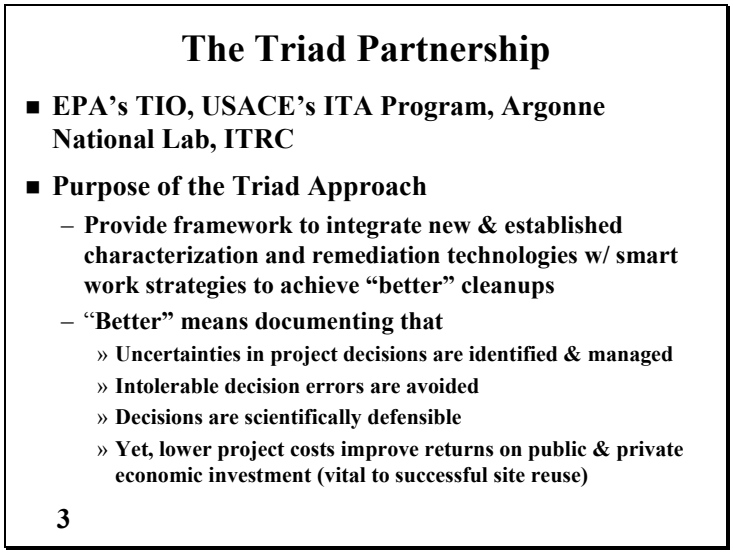
1

Slide 2

Seminar Outline

- Overview of the Triad Approach
- Managing uncertainty means documenting "Why's"
- Updating the environmental data quality model
- Suggested terminology for communicating data quality concepts
- Illustrating the Triad with the Tree Fruit Case Study

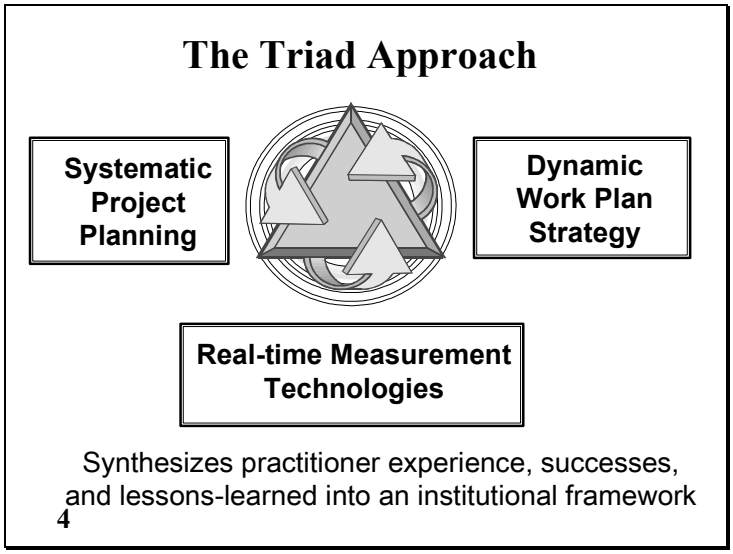
Slide 3



Acronyms: EPA TIO = U.S. Environmental Protection Agency Technology Innovation Office
USACE ITA = U.S. Army Corps of Engineers Innovative Technology Advocates
ITRC = Interstate Technology and Regulatory Council

Based on the current work strategies, low end estimates for cleaning up the thousands of contaminated sites in the U.S. runs to at least $200 billion dollars (in 1996), not even counting Brownfields sites. The Triad approach is concerned with making the characterization and cleanup of contaminated sites more affordable and more effective.

Slide 4



The Triad approach is a work strategy framework for economically managing project decision uncertainties by drawing on the accumulated technical knowledge and experience gained from the past 20-30 years of hazardous site cleanup. The Triad approach proactively exploits new characterization and treatment tools. The Triad serves as a synthesis of various, yet conceptually similar, work strategies developed by innovative and successful site professionals from the U.S. Department of Energy, Tufts University, the U.S. Army Corp of Engineers, the U.S. Environmental Protection Agency, other federal and state agencies, and practitioners in the private sector.

The 3 Legs of the Triad Approach:
Systematic Project Planning

- Take the time to clarify project-specific and decision-specific issues with all stakeholders
- Articulate clear project goals and the decisions (and the tolerable uncertainties) that must be made to bring the project to a satisfactory resolution
- Evaluate potential causes for making decision errors; identify uncertainties
- Develop strategies to manage uncertainties so that decision errors can be avoided
    - Chart best course to reach project goals using conceptual site models (CSMs) that help identify information gaps (i.e., uncertainties) and clarify goals
    - Use multi-disciplinary technical team for project planning and implementation

Dynamic Work Plan Strategy

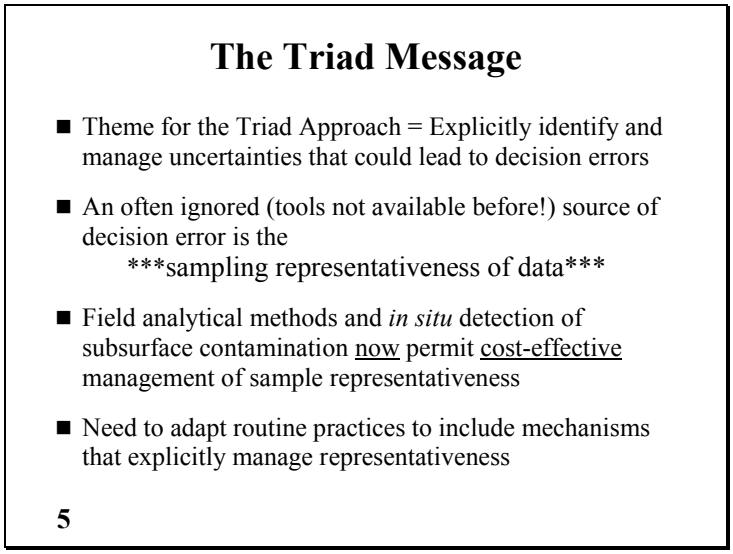
- Real-time, decision-making in the field by experienced technical personnel allows for a seamless flow of site activities = fewer mobilizations
- Regulator-approved decision trees guide data gathering to support rapidly and efficiently evolving the CSM to maturity

Real-time (or near real-time) Data Availability

- Generally will mean on-site analyses
- Support implementation of dynamic work plans
- Permit management of sampling uncertainty
- Method/technology selection and QC design based on integrating the intended data uses with available technologies that can meet the turn-around time and "field-friendliness" needed to support the dynamic work plan.
- Mix-and-match analytical techniques according to specific needs (e.g., field and traditional lab methods; direct push in situ detections and an on-site lab; etc.)

For more details, see the EPA issue paper "Using the Triad Approach to Improve the Cost-effectiveness of Hazardous Waste Site Cleanups " (EPA 542-R-01-016), available at http://cluin.org/triad/ . Additional information and supporting materials are available at http://cluin.org/tiopersp/#curr_pers

Slide 5



The concepts embodied in the Triad are not new. The Triad is a re-articulation and broadening of the original DQO concepts for contaminated site cleanups, just as the DQO process was an articulation of the scientific method customized for the environmental regulatory arena. The Triad adds emphasis, however, on "recognizing, identifying, and managing uncertainty" as the mechanism through which good science is practiced and defensible decisions are made within the environmental cleanup context. A very important source of project decision uncertainty is the representativeness of the data sets upon which project decisions are based. Data uncertainty, when unresolved, can lead to decision errors that affect the protectiveness of redevelopment and the extent and efficiency of cleanup activities. To ensure protectiveness in the face of uncertainty, regulators are often forced to order cleanups that may be much more extensive than actually necessary, incurring unnecessary cleanup and insurance costs. Uncertainty about cleanup costs are a frequent barrier to Brownfields redevelopment and urban renewal.

The Triad approach is being developed as the technical foundation for the next generation of site characterization and cleanup practice and as the technical underpinnings of the one-cleanup-program envisioned by OSWER upper management. The Triad approach was articulated to serve as a technically sound and internally consistent scientifically-based technical paradigm around which successful sampling, analytical, and remedial strategies could be integrated, while welcoming future innovations and cleanup program evolution.

Slide 6

Using the Triad approach requires
Systematic Project Planning.

Systematic project planning
means always being able to explain WHY!!

Systematic project planning
means never having to say,
"Because that's the way we've always done it."

6

Managing uncertainty and scientific defensibility involves two primary aspects:

- Being able to explain the "why" behind every choice that potentially impacts the correctness of project decisions.
- Acknowledging and making allowances for what is unknown and impossible (or too expensive) to find out.

Examples of the kinds of "why's" that must be known to manage data representativeness under the Triad approach (i.e., there should be a technical rationale behind each choice that links the data produced to the intended project decisions):

- Why are we taking x number of samples?
- Why are we collecting the samples with this type of collection tool?
- Why are we collecting samples at these locations?
- Why are we homogenizing this volume of sample?
- Why are we using this sample prep method?
- Why are we using Method X as the determinative method (i.e., to generate the analytical result)?
- Why are we using these particular QC checks?
- Why are we selecting these particular samples for splitting between 2 different analytical methods?
- Why are we analyzing the data using this statistical method?

Under the Triad approach (which strives for scientific defensibility), the answer to these types of questions can never by "Because this is the way we've always done it." or "Because the regulator said to do it that way." This is not to say there will not be times where regulations or other circumstances force one to do things that appear to have no scientific justification. However, these instances should be documented as unavoidable deviations from the Triad approach.

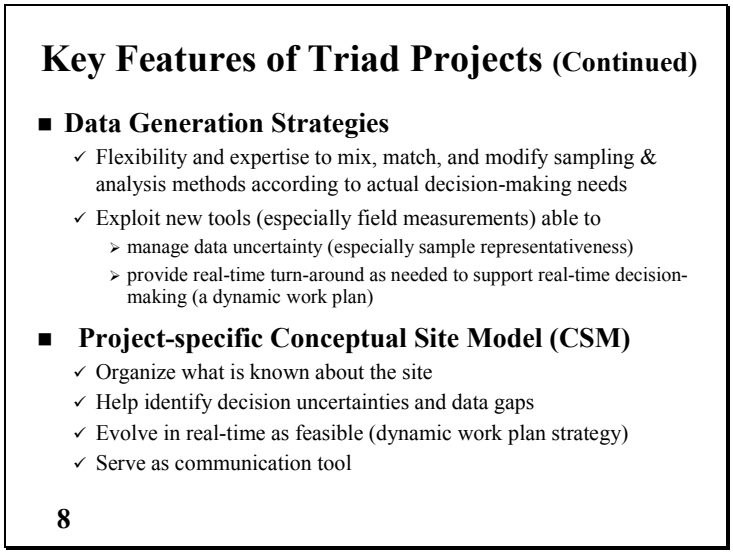
Slide 7

Key features of what works to make for an efficient, well-run project are

- Detailed, thorough project-specific systematic planning that can get you to site closeout in the most resource effective way possible.
- Use existing site-specific information and knowledge/experience based on similar site types as much as possible to identify or at least anticipate the ultimate goals for a site. Include stakeholders in the planning process and seek consensus. Who are stakeholders? One definition is that a "stakeholder" is 'anyone who can drive a stake through the heart of your project.' The local community is a very important stakeholder for most site cleanup projects.
- Articulate project goals clearly, so they can be used as guideposts to focus site work and keep it on track, minimizing decision errors that could jeopardize achieving the goals. This ensures that uncertainty is actively managed through planning, not passively managed by repeating the investigation until resources run out, and then taking "your best guess." If the planning process shows that reducing uncertainty to a desired level is impractical given the logistical or resource constraints associated with the project, an alternative plan can be considered before funds are wasted on fruitless field activities.
- Speed up projects by making decisions in the field as much as possible. Shorter project lifetimes save money, minimize staff turnover and disruption, and make sure data do not go out of date. Real-time decisions need real-time data. Use newly available technologies for improved access to environmental media for sample collection, for detecting and analyzing for contaminants and media properties, and for imaging surface and subsurface features.

Under the Triad approach, it is imperative that work planning documents clearly explain what decisions the project will be making and tie data collection activities directly to those decisions. Report documents should discuss the actual outcomes of data generation (including an interpretation of method QC checks for assessing data quality) and interpretation and how the data supported the project decisions that were actually made.

Slide 8



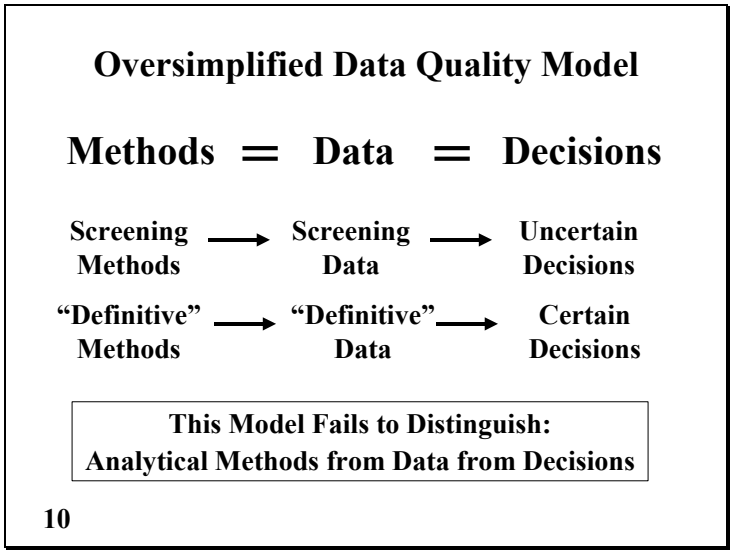
- Use flexible strategies to gather the data needed to make the intended project decisions at the agreed upon level of decision confidence. All factors contributing to the uncertainty in the data must be managed, not just analytical quality. For heterogeneous matrices, all aspects of sample collection and their relevance for sample representativeness are of critical concern.
- Use conceptual site models (CSMs) of various configurations (maps, exposure models, hydrogeologic models, fate and transport models, etc., as appropriate) as tools to organize information about the site and identify data gaps that must be filled in order to make decisions. Use visualization and decision support software to manage data, plot contamination, and estimate statistical uncertainty.
- Use modern communication technology (Internet, cell phones, etc.) to share data, seek regulator approval for deviations from prepared work plans, and keep all parties up-to-date with the latest developments in the project.

This kind of approach was generally not practical in the early years of site investigations. Neither the tools nor the experience existed. Plotting an intelligent course through the maze of site characterization and cleanup requires a level of scientific and engineering knowledge, practical field experience, and regulatory and programmatic frameworks that have taken years to develop. Now that much knowledge and experience are in place, it is theoretically possible (although institutionally difficult) to mature the field by replacing one-size-fits-all assumptions with pertinent, knowledge specific to the variables impacting project decision-making. It is important to convince both regulators and practitioners that the site cleanup industry can begin to discard the process-driven, one-size-fits-all approach (where practitioners are held accountable only for following the process) in favor of intelligent project design (where practitioners are held accountable for demonstrating that the desired outcome is achieved).

Slide 9

Updating the
Environmental Data Quality Model
to Manage
Data Uncertainties

Slide 10



Example of language defining data quality according to classification of the analytical method: Data Quality Objectives Process for Superfund: Interim Final Guidance (Sept. 1993), page 42: "Screening data are generated by rapid, less precise methods of analysis with less rigorous sample preparation."
Page 43: "Definitive data are generated using rigorous analytical methods, such as approved EPA reference methods. Data are analyte-specific, with confirmation of analyte identity and concentration. Methods produce tangible raw data (e.g., chromatograms, spectra, digital values) in the form of paper printouts or computer-generated electronic files. Data may be generated at the site or at an off-site location, as long as the QA/QC requirements are satisfied. For the data to be definitive, either analytical or total measurement error must be determined."

Yet other language from this guidance acknowledges the importance of factors other than the analytical method (from page 43):
"Definitive Data QA/QC Elements

- Sample documentation (location, date and time collected, batch, etc.);
- Chain of custody (when appropriate);
- Sampling design approach (systematic, simple or stratified random, judgmental, etc.);
- Initial and continuing calibration;
- Determination and documentation of detection limits;
- Analyte(s) identification;
- Analyte(s) quantification;
- QC blanks (trip, method, rinsate);
- Matrix spike recoveries;
- Performance Evaluation (PE) samples (when specified);
- (page 44) Analytical error determination (measures precision of analytical method): An appropriate number of replicate aliquots, as specified in the QAPP, are taken from at least one thoroughly homogenized sample, the replicate aliquots are analyzed, and standard laboratory QC parameters (such as variance, mean, and coefficient of variation) are calculated and compared to method-specific performance requirements defined in the QAPP;
- Total measurement error determination (measures overall precision of measurement system, from sample acquisition through analysis): An appropriate number of co-located samples as determined by the QAPP are independently collected from the same location and analyzed following standard operating procedures. Based on these analytical results, standard laboratory QC parameters (such as variance, mean, and coefficient of variation) are calculated and compared to established measurement error goals. This procedure may be required for each matrix under investigation, and may be repeated for a given matrix at more than one location at the site."

Also note the mention of sampling concerns on page 41: "Design elements that must be documented in the Sampling and Analysis Plan include: sampling types (e.g., composite vs grab samples); general collection techniques (e.g., split spoon vs. core drill…); sample support (i.e., the amount of material to be collected for each sample); …"

But the regulatory and engineering models actually used in the site cleanup arena view environmental data quality using the very simple model where data quality is equivalent to the nature of the analytical method:

- Definitive analytical methods automatically produce definitive quality data.
- Screening analytical methods automatically produce screening quality data.

A discussion of whether this view is correct requires that we define the terms we will use to discuss this topic.

What is the difference between a definitive analytical method and a screening analytical method?

- The key is a difference in the perceived amount of uncertainty in analyte ID or in analyte quantification. Screening methods have (or are perceived to have) more uncertainty in one or both tasks than definitive methods.

What is the difference between definitive data and screening data?

- The goal of generating data is to support making a decision. Therefore, "definitive" data are seen as supporting a defensible decision; whereas screening data can not support a defensible, at least, not with that data only. So you could say that the difference between definitive data and screening data is the amount of uncertainty in the data set with respect to the decision to be made.

The common thread in contrasting the word "definitive" vs. the word "screening" is the degree of uncertainty in whatever we are talking about.

What is a "method"?

- An analytical method is the general description of the procedures used to operate an analytical technique. For example, GC-MS is a technique that can be used to detect and quantify chemicals that have properties that allow them to be volatilized in a gas stream through the GC, and are of sufficient molecular mass to be detected by the MS. Usually, a method is developed that is generally applicable to certain groups of contaminants. So there is a method for more volatile compounds (with boiling points less than about 200 C) and a different method for less volatile compounds (with boiling points greater than about 200 C). The behavior of a specific analyte within an analytical system such as a GC-MS depends on a number of properties other than just boiling point. A single generalized method cannot account for all these other properties. Therefore, the "analytical quality" for different analytes varies when all analytes are detected and quantitated from a single run. Analytes that happen to have properties that match well with the operating conditions of the generalized method will have better analytical quality than analytes whose properties are not as well matched to that set of operating conditions.

There are many specific details about running a method that are not covered in the method. These specific details will vary depending on the construction of the particular instrument or equipment used, and the application for which the method is being used. Highly detailed descriptions of how methods are adapted to specific equipment and applications are called Standard Operating Procedures (SOPs). Optimizing analytical performance (such as precision, bias, and detection capability) for a specific analyte requires optimizing the specific procedures and operating conditions to the properties of the analyte in question. Analytes with similar properties (such as hydrophobic compounds that tend not to dissolve well in water) will have similar optimal operating conditions, whereas analytes with different properties (such as more hydrophilic compounds that dissolve better in water) would require a different set of optimal operating conditions. It is impossible with current technology to operate a GC-MS such that both hydrophobic and hydrophilic analyte performance is optimized in a single run. However, economics have encouraged the environmental community to expect a single analytical method to produce data for hundreds of analytes in a single analysis, yet all data is expected to be of the same analytical quality.

For example, using an "approved" or published method (such as EPA SW-846 Method 8270) to operate a definitive technique (GC-MS) perfectly in compliance with the laboratory's audit-compliant SOP, will generate high analytical quality for some analytes and poor analytical quality for other analytes in the very same sample, even if all other variables are controlled. (But under the current environmental paradigm, many other variables are neither controlled nor documented.)

So does the idea hold that "definitive" methods always generate "definitive" data? Here are cases where a "definitive" method would produce data that do not support good decisions (i.e., would not be "definitive"):

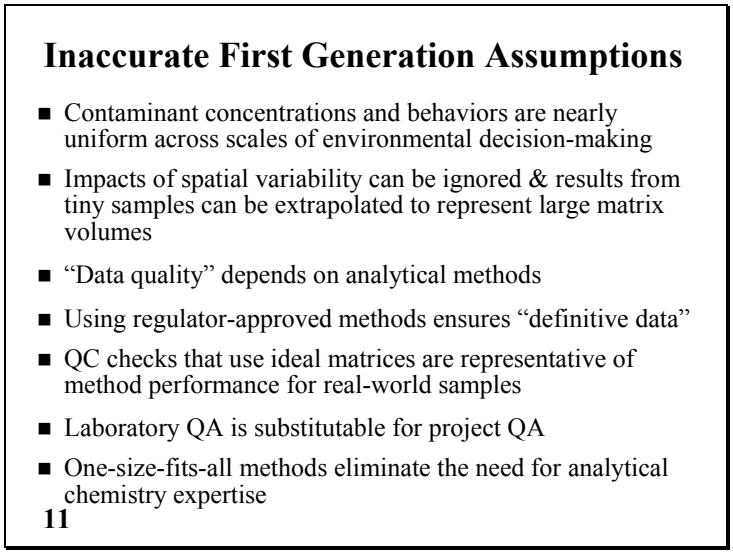
- The sampling procedures or design used were not representative of the decision.
- Even if the sampling scheme was representative, a "definitive" method can easily produce "bad data" because definitive methods are still subject to interferences that compromise performance (although usually definitive methods are less susceptible to interferences than might be expected for screening methods).

It is possible that data generated by a screening method can be used to make a perfectly defensible and confident decisions. For example: the Dexsil L2000 instrument detects all organochlorine compounds that can be extracted with the extraction solvent used, that will react with elemental sodium to strip off chlorines and convert them to chloride, to be detected by a chloride-specific electrode. But the instrument does not distinguish which organochlorine compounds are present in its response. If there is careful control over extraction efficiencies, detection limit concerns, and other data comparability issues, and the Dexsil instrument found no organochlorine contaminants in any samples from a dense sampling scheme across a site, there is high confidence that a decision of "clean" (of organochlorine contaminants) is correct. This data set (in conjunction with the QC data that demonstrates adequate method performance, which may include some split samples for more analyte-specific analysis to establish data comparability) may be considered "definitive" as far as that decision is concerned.

On the other hand, if there were detections of organochlorine contaminants by the Dexsil instrument above the lowest possible regulatory threshold, this is preliminary indication that contamination of concern may be present. More work would need to be done to identify exactly which organochlorine compounds are present, and whether they are present at concentrations above their respective regulatory thresholds. By itself, the Dexsil data set (which includes the QC data that demonstrates adequate method performance) would be considered "screening" with respect to decisions about whether specific contaminants of concern are present above their respective threshold over some designated exposure or remediation (decision) units (specific areas or volumes of the site). In that circumstance, the Dexsil method is best used to establish the spatial distribution and general concentration range and variability of contamination (i.e., manage sampling uncertainties). Based on the Dexsil data set, another sampling design would generate a analyte-specific data set(s) to fully meet risk assessment or remediation decision-making needs.

So deciding whether data are “definitive” or “screening” depends on what the data results are and how they are being interpreted and used. For example, uncertainties in the identity of analytes detected by a screening analytical method can be managed by means external to the method—historical knowledge about the site or the waste, or prior data.

Slide 11



The "data" referred to here is analytical chemistry data for pollutant/contaminant concentrations in the environmental media encountered with cleanup of hazardous waste sites.

The current data quality model functions as if the following assumptions were true (but they are not!):

- The variability present in environmental systems (i.e., variability in contaminant concentrations or variability in contaminant behavior (such as in analytical measurement systems or toxicity toward receptors) is so minimal that it can be safely ignored
- "Data quality" is determined by the documentation and accuracy of the laboratory analytical method procedures
- Analytical accuracy for environmental samples can be ensured by using one-size-fits-all regulator-approved methods
- QC checks using ideal matrices (reagent water, clean sand) are representative of method performance for real-world samples
- Laboratory QA is substitutable for project QA (i.e., if method performance is in control, project decisions are trustworthy)
- After the selection, performance, and interpretation of analytical methods has been "standardized," analytical chemistry expertise is no longer needed either at the project or lab level since all potential variables that could affect the usefulness of data have already been accounted for

Slide 12

The Foundation of a Better Data Quality Model

Data Quality = Should be assessed according to the ability of data to provide information that meets user needs

- Users need to make correct decisions
- Therefore, data quality is a function of data's…
  - Ability to represent the "true state" (of the decision unit) in the context of the decision the data user wants to make

12

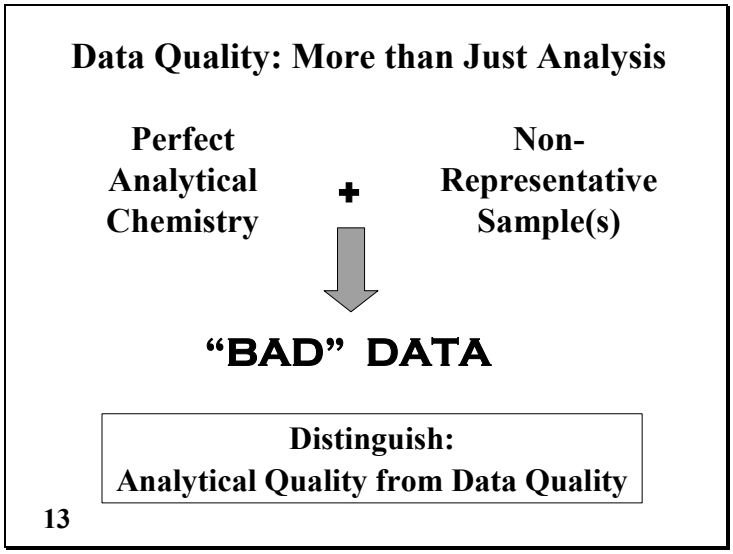
First of all, define key terms:
"Data" = analytical results for chemical contaminants generated on environmental samples; used for supporting environmental decisions.
Quality = "the totality of *features* and *characteristics* of a product or service that bear on its ability to meet the stated or implied needs and expectations of the user" (USEPA OEI QMP 2000, v. 1.3) = "fitness for use" (Jeff Worthington, USEPA OEI, "Information Quality Systems" presentation at the May 2001 EPA Conference on Environmental Statistics and Information)
Data quality = "the totality of features and characteristics of data that bear on its ability to meet the stated or implied needs and expectations of the user/customer" (USEPA OEI QMP 2000, v. 1.3).
Data quality = "degree to which data satisfies stated or implied needs"… "High quality data is sufficiently trustworthy to meet the needs of the business purpose for which it was intended." (Oracle Data Quality Inspector software literature)
"…data quality, as a concept, is meaningful only when it relates to the intended use of the data. Data quality does not exist in a vacuum; one must know in what context a data set is to be used in order to establish a relevant yardstick for judging whether or not the data set is adequate." (USEPA QA/G-9; July 2000 version, page 0-1).

"Quality" is an emergent property arising from the interaction of several characteristics of the element in question (such as "data," as in data quality, or "information," as in "information quality") with various aspects of the intended use of that element (i.e., how the data or information are to be used). For that reason, terms involving "quality" (data quality, water quality, air quality, etc.) have always been very broad and ambiguous. If "quality" is "good," the implication is that all characteristics of that element are in line with the intended use. If "quality" is "poor," additional explanation is needed to know which characteristic(s) is(are) not synchronized with the intended use. For example, "poor water quality" could be due to high levels of bacteria or nutrients or pollutants, or low dissolved oxygen or low pH, etc.,etc. More information is needed to understand the reason for the "poor quality."
It is also possible for identical values of the same characteristic to be considered "good quality" under one intended use scenario, yet constitute "poor quality" under a different intended use scenario. Therefore, any efforts to correct "quality" problems must become very specific about what the factors producing "quality" actually are and how they interact with intended use before the source of a problem can be identified and addressed. What are the factors that contribute to good quality in environmental data? (Will be discussed in subsequent slides.)

Representativeness = the data generation process produces information that accurately reflects the site feature(s) that is(are) being subjected to the decision-making process. "Representing the 'true state'" for environmental contamination must be linked solidly to "the context of the decision to be made." Why? Because in contaminated site projects, the true state (such as the concentrations of contaminants or the stratigraphy controlling contaminant movement) can easily vary markedly over scales as small as feet to inches. It is not resource-feasible to characterize the "true state" for the entire site (vertically and horizontally) at a scale that small. Yet, an accurate understanding of contamination at a small scale for certain portions of a site MAY be important to an accurate conceptual model that can support cost-effective cleanup. So how does a project manager choose the scale (or scales for different portions of the site) over which to characterize the "nature and extent of contamination"? Remember that we can't afford to take an infinite number of samples, so data points must be used to extrapolate or "represent" the true state in order to make environmental decisions. Therefore, the scale over which to characterize must be matched precisely to the scale over which decision-making will occur. Therefore, data representativeness can only be assessed by explicitly linking the all aspects of sample collection, preparation and analysis with the specific decisions for which the data are intended to support.

If linkage is simply assumed, for example, when method checklists or sampling 'rules-of-thumb' are used without verifying their appropriateness, there is a strong likelihood that a mismatch between sample representativeness and the intended decision will occur. Why? Because most of these checklists and rules-of-thumb were developed based on simplifying assumptions used before scientists had a chance to develop more complete understanding of the heterogeneity of environmental systems and how that impacts sampling and analysis choices. Now we have hard evidence that these simplifying assumptions often do not hold true. Conclusions based on inaccurate assumptions risk being wrong, and decisions based on erroneous assumptions risk being wrong. The good news is that we now have knowledge and technologies to replace the need to rely on over-simplified assumptions. Thus the field can begin to mature beyond the first approximations it had to begin with.

Slide 13



If assessment of data quality is grounded in defensible decisions, then the concept of data quality must extend beyond just the quality of the analysis. If a sample is not representative of the feature under investigation, BAD data is produced even if the analysis is perfectly accurate. It is "bad data" because data generated on non-representative samples will be misleading (i.e., will lead to erroneous conclusions).

The issue of sampling representativeness and its relationship to the ability of data to support decision-making (and the challenges posed when sampling heterogeneous environmental media) have been discussed for years in many different forums. But because of the complexities of sampling and analyzing environmental materials, these issues and their ramifications are not well understood by many important environmental decision-makers.

Unfortunately, there is a widespread misconception that "highly accurate analyses automatically produce accurate data." This has caused a great deal of energy to be focused on trying to create one-size-fits-all lists that prescribe exactly what analytical methods should be used for evaluating contaminated sites. This thinking fails to consider the many variables that impact the adequacy of environmental data for decision-making purposes, so it ultimately fails to achieve its purpose. In addition, the terminology commonly used over the years has become ingrained with unspoken assumptions that reinforce this misconception.

Slide 14

Representative Data - Key to Sound Science

Using good science
in the cleanup of contaminated sites means that
the scale of data generation and interpretation
must closely "match" (i.e., represent) the scale of
project decisions being based on that data.

"Sound science" also means that uncertainty must be acknowledged and managed since an exact match is not usually feasible for complex, heterogeneous systems.

What types of things must be considered when developing a representative data set?

14

The scales of data generation and interpretation (i.e., data representativeness and uncertainty both on the sampling side and on the analytical side) must be explicitly matched to the scale of decision-making.

Potential scales include spatial (inches, yards, miles); temporal (minutes, days, years); chemical identity [analyses that are specific for a single target analyte, analyses that respond to an analyte class (target analyte + degradation products or compounds with similar structures), surrogate marker compounds]; and bioavailability [elemental oxidation states (e.g., Cr(III) vs. Cr(VI), chemical speciation (e.g., inorganic mercury vs. organic mercury), incorporation into mineral structure or chelation by organic matrices).

Uncertainty is always present since our ability to exactly match the scale of data generation with the scale of decision-making will always be limited by funding constraints and technological inadequacies. Questions that must be grappled with on a project-specific basis include: "What exactly are the decisions being made?" "What level of decision uncertainty can be tolerated in those decisions?" "How does uncertainty in the data interact with decision uncertainty?" "How does the project manager know how much uncertainty is present in the data?" Although the knowledge and experience needed to tackle those questions now exists within the environmental science arena, doing so requires multi-disciplinary collaboration and expertise. "Sound science" cannot be conducted without including a range of diverse disciplines and technical expertise in the decision-making process.

Slide 15

The concept of "representativeness" must be grounded in the decision context

Different decisions require different representativeness. For example:

- A data set representative of a risk assessment decision usually needs to estimate the average concentration over a fairly large decision unit (called an "exposure unit")
- A data set representative of a cost-effective remedial design must provide information about concentration extremes at a scale specific to the remedial option considered. Remedial scales are nearly always different from risk assessment scales.

It is impossible to specify a one-size-fits-all data set that could be representative of all potential site decisions!

Therefore, the first step of ensuring data quality is to clearly understand to what decisions the data will be applied.

15

The concept of "representativeness" is very vague and ill-defined for many working in the environmental field. This is because standard or regulatory definitions for "representativeness" also tend to be rather vague.

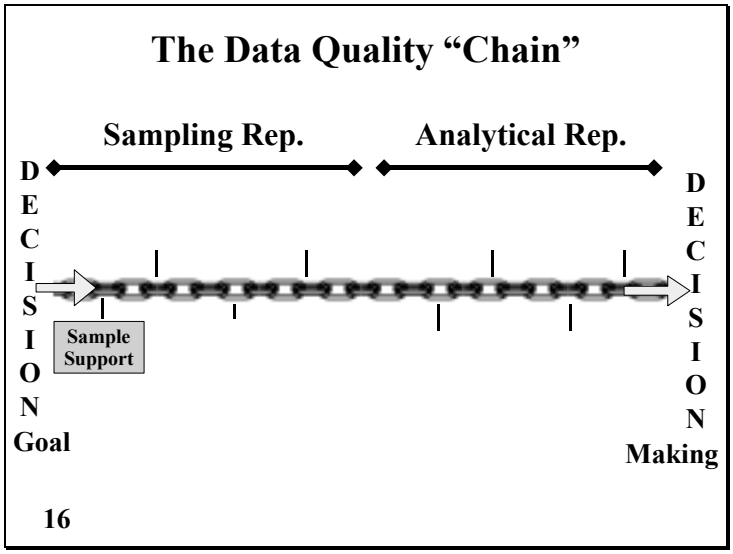
- The RCRA solid waste regulations at 40 CFR §260.10 define a **representative sample** as: "a sample of a universe or whole (e.g., waste pile, lagoon, ground water) which can be expected to exhibit the average properties of the universe or whole."
- ASTM (consensus standard D 6044-96) defines a representative sample as "a sample collected in such a manner that it reflects one or more characteristics of interest (as defined by the project objectives) of a population from which it was collected."

It is not clear from these definitions whether the term "sample" refers to a statistical sample (made up of a number of individual specimens) or to a single sample, or whether the authors intended to allow either interpretation. A critical issue with the RCRA regulatory definition is that representativeness is defined in terms of an "average." Operationalizing this definition for contaminated site cleanup poses problems. First, the extreme heterogeneity of environmental matrices and contaminants makes determination of a statistical "average" difficult and expensive. Second, some environmental decisions (notably, those decisions involved with selecting and designing remedial systems) should not be made based on an "average," if that average encompasses wide variation.

In order to be useful for managing projects in the environmental field, the concept of representativeness must be made more concrete and meaningful. This can be done by simply adding the word "of." This adjusts the terminology and people's thinking to make it clear that data or other information must be representative of the intended decision or specific property under investigation. In this way, "representativeness" becomes linked to a concrete decision and decision unit rather than just an abstract "average." The ASTM definition seems to reflect this same kind of approach ("…reflects one or more characteristics of interest (as defined by the project objectives)…"

One of the reasons cleanups become inefficient and costly is that the decisions to be made and the representativeness of data sets is not thought through in advance of field work. Decisions are articulated after data are collected. Then it is discovered that the data are not really representative of the decisions you want to make. The option of returning to the field to collect more data is time consuming and costly, but usually necessary. The other option, force-fitting non-representative data to the decision-making process constitutes very poor science. There is significant risk that decisions will be made in error (e.g., remedial designs will not be effective), and that there will be legal and public-relations battles over the interpretation of the data and implementation of the decisions. For data collection to be cost-effective, each decision must be known in advance so that data collection can be tailored to be representative of that decision.

Slide 16



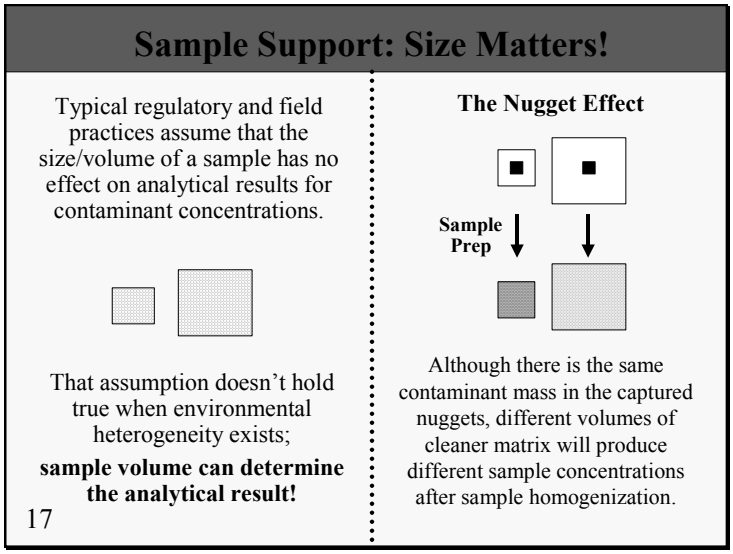
Various sources of uncertainty that impact the representativeness of data (i.e., "data quality"). The chain of activities that produces data of adequate quality must be anchored in the decision-making process. First it must be known what decision the data are to be representative of. Of course, data generation involves 2 main elements: sampling and analysis.

Sampling procedures must be selected to be representative of site features in the context of the decision to be made. The representativeness of sampling can be further broken into discrete components.

A very important component is the concept of sample support:

- Sample support = volume, dimensions, physical orientation and characteristics of the specimen being removed from the parent matrix. How should a sample/specimen be removed so that it retains the characteristics of the parent matrix that is under investigation? This is explained more clearly on the next slide.

Slide 17

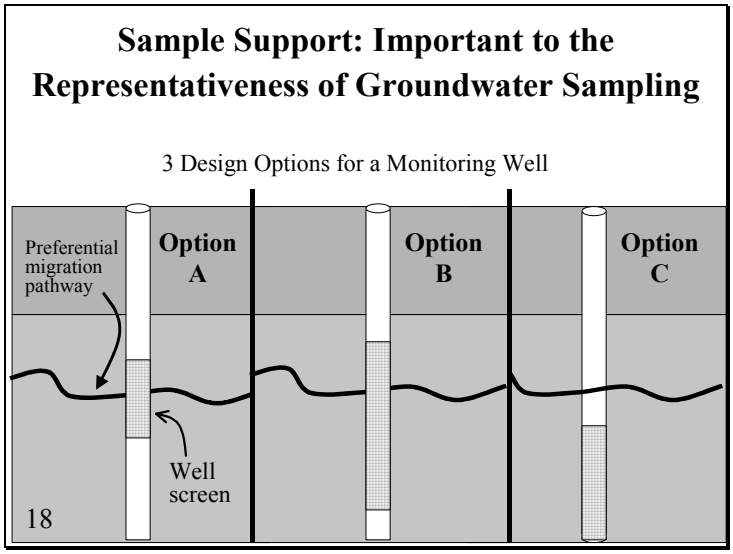


The issue of "sample support" for heterogeneous environmental and waste matrices should make us reconsider the common (and usually unacknowledged) assumption that the reported concentration of an environmental sample should be the same no matter what volume of sample is collected.

The volume of the sample is an important factor that influences the reported concentration for a sample, especially when contaminants are heterogeneously distributed throughout the parent matrix. All samples must be homogenized (through physical or chemical means) prior to analysis. For heterogeneous samples (that are affected by the nugget effect to a lesser or greater degree), the analytical result for a sample is determined by how much contaminant is captured in that sample, and how much cleaner matrix is contained in the sample (that serves to dilute the contaminant during homogenization). The nature of contaminant release to the environment (such as release to ground surface in the form of a powder or particulate) increases the probability of heterogeneity, as does contaminant solubility, mobility, and the age of the release. Obviously, environmental variables (such as precipitation, wind erosion, temperature, matrix composition) interact with the contaminants' properties to mitigate or aggravate heterogeneity. Contaminants that may at first have been more homogeneously released may become heterogeneously distributed throughout a matrix if their chemical properties cause them to preferentially partition onto mineral surfaces or into organic carbon that are themselves heterogeneously distributed, or inclusion of those matrix components into the analytical sample is variable or unpredictable.

The issue of sample support is becoming an increasingly important determinant of analytical result as more sophisticated analytical technologies require smaller and smaller volumes of sample. At one extreme, sensors technologies currently under development will have miniscule sample supports, and data interpretation will be extremely difficult unless there is much greater awareness and management of sample support concepts.

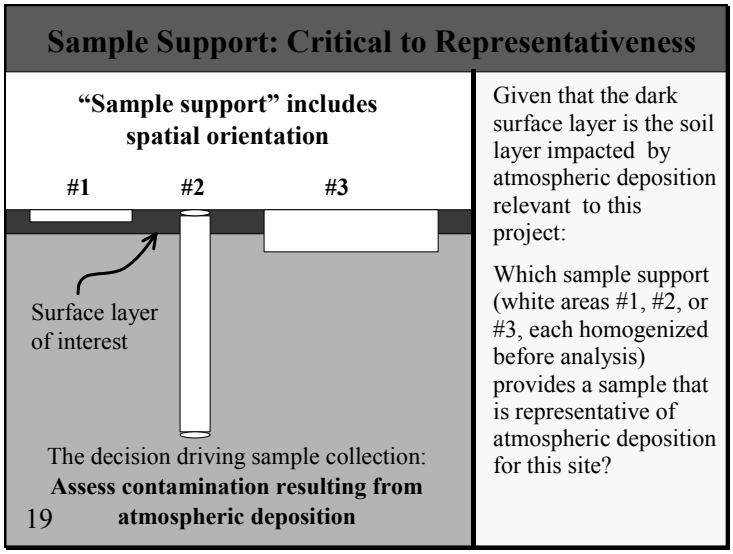
Slide 18



When contaminants preferentially partition into, or migrate through, one stratigraphic layer (such as a clay layer or a fracture), the volume of cleaner water mixed with water from the contaminated layer (influenced by factors such as screen length; purge intensity; hydraulic conductivity of layers) will determine the contaminant concentration of the sample. If this source of variability is not understood and controlled, it is not unexpected that laboratory results will vary widely, because the samples truly contain different concentrations.

Vertical heterogeneity of aquifers can be understood through using new characterization tools, such as direct push in situ detection systems [e.g., membrane-interface probes (MIP) with PID, FID, or ECD detectors; laser-induced florescence (LIF) detectors; and others] and by vertical profiling with passive diffusion bag samples. This knowledge is then used to decide how best to consistently collect samples that are representative of the intended decision.

Slide 19



This series of slides will illustrate concepts related to "sample support." These concepts are presented in a simplified form and do not attempt to portray the more exacting aspects of this topic.

The panel on the left illustrates how sample volume and orientation must be selected to be representative of the decision to be made. Any of the 3 samples might be argued to represent true site conditions, but only one can be argued to be representative of site conditions in the context of the decision (atmospheric deposition).

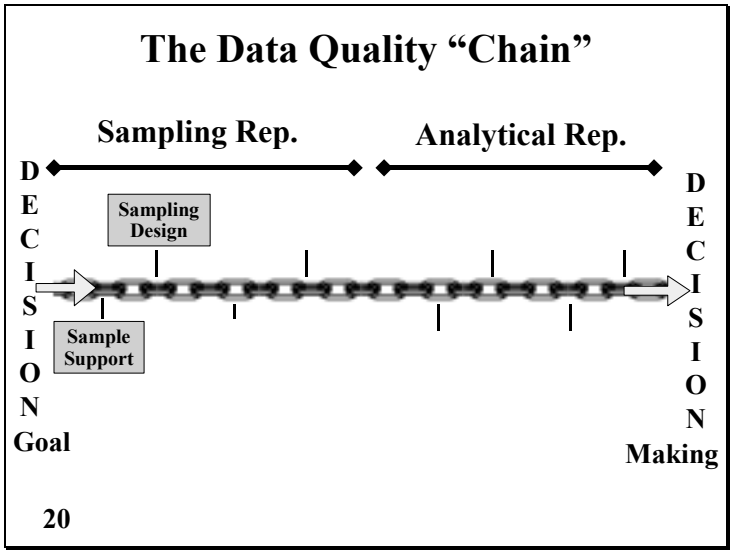
Color Key for left panel:

- Dark brown depicts surface soil impacted by surface deposition of lead from the atmosphere.
- Light brown depicts soil that would not be expected to be impacted by this atmospheric deposition.
- White areas depict the volume and orientation of material removed that becomes the "sample."

Keep in mind that the entire sample is homogenized prior to subsampling for analysis.

The sample support (the physical dimensions of the sample) for Sample #1 would be representative of the matrix impacted by atmospheric deposition, but the sample supports of samples #2 and #3 would not be. Sample support #3 illustrates the importance of strict control over sample support in scenarios where careful stratification of populations is required to avoid biasing results by including non-representative sample. Even though the general orientation of sample collection in #3 is similar to #1, the concentration of lead in sample #3 would be expected to be "diluted" by the inclusion of "cleaner" soil from a non-representative layer into the sample.

Slide 20

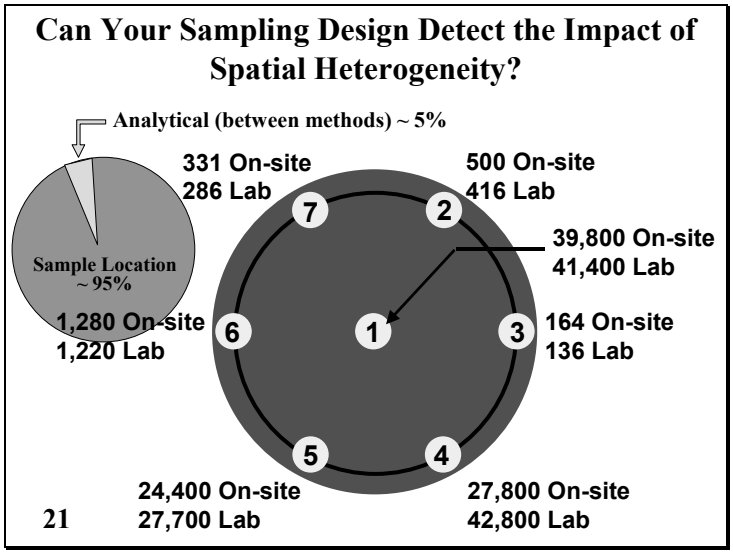


Various sources of uncertainty that impact "data quality," continued.

Another component of sampling is:

- Sampling design = sample numbers, locations, and timing. How many, where, and when should samples be collected so that the data set will give an accurate representation of the question being asked about a site?

Slide 21



Example of characterizing sampling variability from USACE/CRREL work (Tom Jenkins). This example is from the Monite installation, which is contaminated with explosives residues due to facility operations that reclaimed explosives from out-of-date munitions.

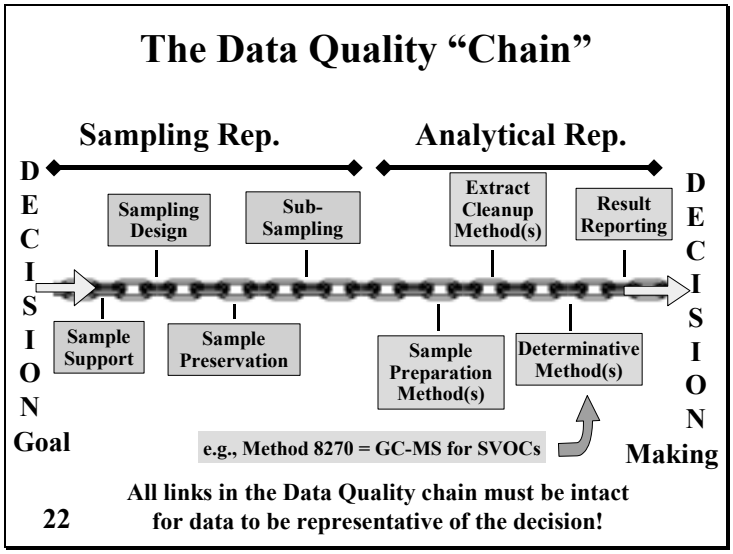
The wheel presented the results of the analysis of the set of 7 discrete samples. The analyte = TNT; units = ppm. Diameter of wheel = 122 cm (4 ft). Surface samples were taken from 0 cm to 15 cm depth (0 – 6 inches), by a stainless steel auger with diameter = 5 cm (2 inches).

Each soil core (one from each of the 7 locations) was thoroughly homogenized. Subsamples from the homogenized sample were analyzed by both an on-site analytical method (EnSys Colorimetric Test Kits; EPA SW-846 Method 8515) and in a traditional laboratory (EPA SW-846 HPLC Method 8330). Note the general agreement of the on-site colorimetric results with the off-site HPLC results. Note the differences in the results among the seven sampling locations. A very different decision regarding the need for remediation might be made if the location for sample collection was at position number 1 or position number 7 (although they are only 2 ft apart)!

An analysis of variance (ANOVA) was used to partition the variability in results between 1) the variability due to the position of the sample (sample location) and 2) differences between the field vs. lab analyses. [ANOVA is based on the assumption that the data are normally distributed, which is not the case with this data set. Therefore, the values obtained are only approximations.] It found that at least 95% of the total variability was due to position (that is, due to matrix heterogeneity) and no more than 5% was due to the difference between analytical method. Another way to state this is that: In this example, matrix heterogeneity caused 19 times more uncertainty in the data results than did the choice of analytical method, over a distance of only about 2 feet.

Conclusion: Spatially, the matrix was very heterogeneous with respect to its concentration of TNT and since any one of these discrete samples would be a legitimate sample by the traditional approach, the traditional approach would not provide representative samples to characterize this site.

Slide 22



Various sources of uncertainty that impact "data quality," continued.

Other components of representative sampling and analysis are:

- Pre-analysis Sample Processing which consists of sample preservation (if required) and transport to the analytical facility. Once in the analytical facility, the bulk sample is usually subsampled to take an aliquot that will be further processed for the actual analysis. Significant loss of sample integrity (and erosion of sample representativeness) can occur during both sample transport and subsampling.
- Sample preservation between collection and analysis = Are the analytes that were originally present in the sample when it was part of the parent matrix still present in the same concentrations by the time the sample is prepared for analysis?
- Sample subsampling for analysis = Will the subsample actually subjected to analysis accurately reflect the properties of the bulk sample?

Sample Analysis must also be representative of the site conditions in the context of the decision to be made. Sample Analysis is composed of several components:

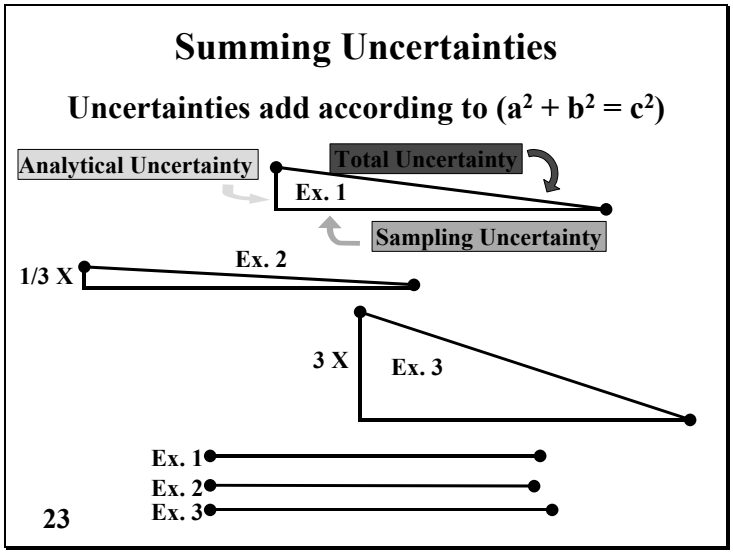
- Sample preparation method(s) (drying, grinding, heating, purging, extraction with solvent, digestion with acid) = Will the sample preparation step accurately transfer the target analytes into a form that can be introduced into the determinative method? Or could losses or contamination occur?
- Extract cleanup method(s) (the sample preparation step can bring along other chemicals that interfere with accurate determination of the target analyte) = Is a cleanup method needed to remove interfering species from the extract so that target analyte results will be truly representative of the target analyte and not biased by other compounds?
- Sample analysis (determinative) method = Will the determinative method accurately measure the analyte presented to it? Or will interferences dampen or enhance the signal of the target analyte? Are the operating conditions of the method adequately optimized to respond to the target analyte to the desired degree of accuracy?
- Result reporting = Are the results transmitted to the data user representative of the results that were obtained by the analytical instrument?

Note: This is intended to be an illustration of the conceptual steps involved in sampling and analysis. Not every sampling or analytical activity is physically composed of separate steps. For some analyses, various steps are combined or are unnecessary. For example, using purge-&-trap for VOC analysis of water samples will not usually require subsampling in the laboratory, and the purge-&-trap device effectively combines sample preparation and cleanup into a single step.

What step of this chain currently gets all the attention and discussion?
But where does most of the error lie?

The total uncertainty in data is the compounding of the uncertainties from each step of the data quality chain. Overall "data quality" includes all these uncertainties, and all these uncertainties ultimately impact decision quality. If there are regulatory or other requirements for the use of Method 8260 as an "approved method" for analyzing for VOCs, but there is no consideration of the rest of the data quality chain, how much confidence can truly be had in overall data quality? Similarly, you can extensively audit Method 8260 or any other SW-846 determinative method, but if other sources of uncertainty are ignored, how far does it get you in trying to establish decision quality?

Slide 23



Understanding the importance of managing sampling uncertainty. (Although presenting uncertainty this way oversimplifies the mathematics involved to compress the individual sources of uncertainty into only 2 major components, this legitimately illustrates the basic concept.)

Uncertainties (when expressed as statistical standard deviations) add as orthogonal vectors, that is, the sum of 2 uncertainty components (represented by the sides of a right triangle) is represented by the hypotenuse. The heterogeneity of environmental materials, especially solids (waste materials, soils, the subsurface) is very high. The vast majority of result uncertainty in environmental samples is due to sampling considerations. Attempts to quantify the relative contributions of sampling and analytical variabilities to the environmental measurement process have "estimated that up to 90 percent of all environmental measurement variability can be attributed to the sampling process." (Reference: Homsher et al, 1991, see Environmental Lab articles in Resources/Links section). It is reasonable to expect that the actual value would vary greatly from project to project and analyte to analyte, depending upon the environmental matrix and the concentrations of the contaminants, the mechanism by which contaminants were introduced into the environment, the fate and transport of the contaminants, as well as how the partitioning of variability was derived and calculated.

The Example 1 figure illustrates a ratio of sampling uncertainty to analytical uncertainty (when expressed as standard deviation) in soil of about 9 to 1 ratio (a convenient ratio for displaying the figure on this slide; compare this ratio to the ratios calculated for the Brownfields case study discussed below). As illustrated in Example 2, decreasing the analytical uncertainty to 1/3rd of the original without addressing sampling uncertainty will no doubt add to the analytical costs, but will not meaningfully decrease the overall uncertainty in the data. Alternatively, allowing the analytical uncertainty to increase to 3 times the original without changing the sampling uncertainty does not significantly increase the overall uncertainty in the data (Example 3).

The overall uncertainty is what impacts the decision-making process (i.e., the overall data quality impacts the decision quality). Therefore, both analytical and sampling uncertainties must be managed. Minimizing one without addressing the other is pointless.

The Example 1 figure illustrates a ratio of sampling uncertainty to analytical uncertainty in soil of about 9 to 1 ratio. The selection of this ratio was rather arbitrary, largely to accommodate the creation of the presentation graphics. Such a ratio corresponds to 98.8% of the variability in the data being attributable to the sampling side, as opposed to variability in the analytical method. Is this a reasonable choice?

To evaluate this question, a soil data set from the investigation of a former scrap yard site (where 291 samples were analyzed for metals, and over 550 samples were analyzed for PAHs) was used to prepare estimates of sampling vs. analytical variability. Analytical variability was estimated by obtaining laboratory control sample (LCS, solid matrix) from the same laboratory that performed the sample analyses. (The LCS data was averaged over a 9-month period and involved n = about 400 for the metals, and n = about 70 for the PAHs.) Total variability was estimated by computing the standard deviation on an analyte by analyte basis for the entire sample set. The sampling standard deviation is calculated by solving the Pythagorean Theorum. The following sampling to analytical ratios were found for selected metals with 90 to 100% detections: arsenic—3 to 1; cadmium—55 to 1; mercury—480 to 1; chromium—661 to 1; and lead—1085 to 1. The following sampling to analytical ratios were found for selected PAHs with greater than 80% detections: benzo(a)pyrene—1464 to 1; phenanthrene—2375 to 1; and fluoranthene—2514 to 1. If matrix spike/matrix spike duplicates (MS/MSDs) data (n = 24 to 31 pairs) were used instead of LCS data to assess analytical performance for the PAHs, the sampling to analytical ratios dropped to benzo(a)pyrene—501 to 1; phenanthrene—528 to 1; and fluoranthene—517 to 1. Obviously, the choice of a sampling to analytical ratio of 9 to 1 for the purpose of a conceptual illustration is not excessive. (See "Supplemental Material--Partitioning Data Variability" in "Links to Additional Resources" section of this seminar's materials.)

The use of the triangle for representing environmental data variabilities has been in the published literature since at least 1977. [Reference: Rhodes, Raymond, C., "Components of Variation in Chemical Analysis" in *Validation of the Measurement Process,* James R. DeVoe, ed., American Chemical Society Symposium Series 63, Washington, DC, 1977, pp. 176-199.] The length of the line segments correspond to the standard deviation, therefore the units for the lines are whatever units are being used for the data itself, such as ppm. Discussion of the triangle representation is included in EPA's new G-5i document (Guidance on Data Quality Indicators), which is currently in peer-review draft and should soon be accessible from the following URL: http://www.epa.gov/quality/qs_docs/.

Slide 24

Partitioning Data Uncertainty

Example using a Brownfields project data set (scrap yard site with contaminated soil)

Std Dev $_{Sampling}$: Std Dev $_{Analytical}$ = Samp:Anal Ratio

(Total variability determined from entire data set. LCS data used to estimate analytical variability. Sampling variability calculated by subtraction.)

As (natural background present): 22.4 : 7 = 3 : 1
Pb (high spatial variability): 3255 : 3 = 1085 : 1

A 3:1 ratio for sampling-to-analytical Std Dev = 90% of statistical variance due to non-method considerations

A 1000:1 ratio for sampling-to-analytical Std Dev = 99.999% of statistical variance due to non-method considerations

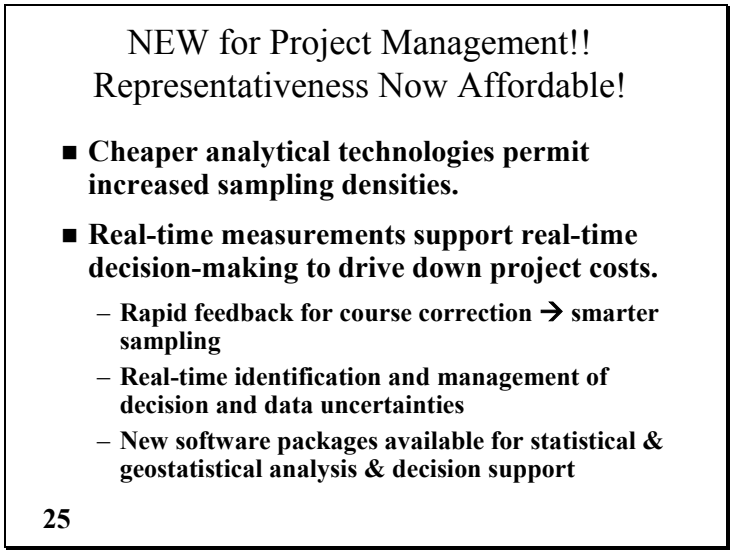
24

This example site was a former scrap yard with soils contaminated with metals and petroleum hydrocarbons. See "Supplemental Material--Partitioning Data Variability" in "Links to Additional Resources" section of this seminar's materials for a more detailed discussion of this example and calculations for additional metal and PAH analytes. A summary of this write-up appears below.

The ratio of "sampling" vs. "analytical" variability (expressed as a standard deviation) can be coarsely partitioned using the following procedure (or a variation therefore): Analytical variability, as method (i.e., not just instrument) variability, can be estimated from the precision of the laboratory control sample (LCS) or matrix spike/matrix spike duplicate (MS/MSD) results.

- An LCS should go through the same sample prep, cleanup, and determinative/instrument methods as the real samples. The more similar the LCS matrix is to the real-world matrix under consideration, the more representative this estimate of "pure" method variability will be of the method variability for the real-world samples. However, LCS matrices are currently seldom representative of real world samples.
- MS/MSD pair precision provides an estimate of the analytical/method variability experienced when real-world matrix effects are factored in.

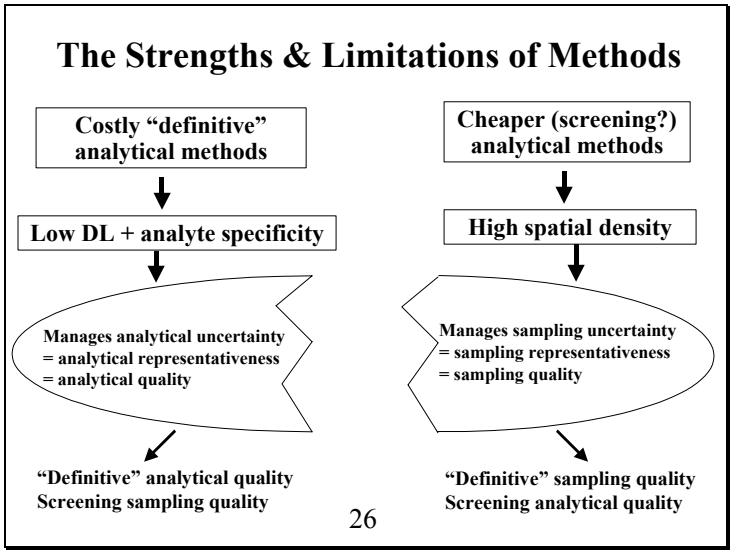
Slide 25



Thanks to technologies that are relatively new to the environmental field, we can begin to address the problem of sample representativeness. One aspect of these technologies that allows management of sampling uncertainty is the ability to run many more samples because per test costs are lower. Another aspect is that many, although not all, of these technologies can be run in the field. This saves sample preservation, transportation, and storage costs. But most importantly, real-time testing results support real-time decision-making, which offers a whole host of benefits. Real-time management and interpretation of data is supported by a variety of new decision-support software tools.

- VSP software pkg FREE: http://dqo.pnl.gov/VSP/index.htm
- SADA software pkg FREE: http://www.tiem.utk.edu/~sada/
- FIELDS/SADA software: http://www.epa.gov/region5fields/static/pages/index.html

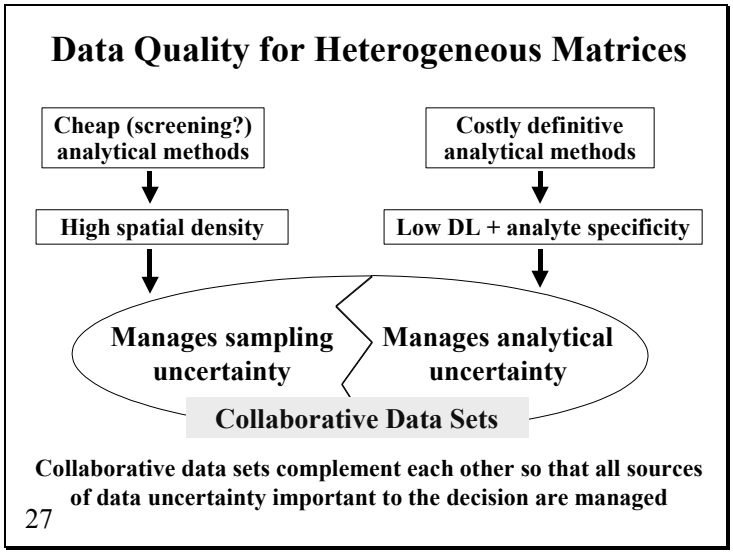
Slide 26



Costly traditional laboratory methods, if used properly, may produce very high analytical quality on the samples received by the laboratory. However, use of these methods does not ensure DATA quality. The cost of traditional laboratory analysis discourages project managers from collecting very many samples. Therefore the sampling representativeness of the laboratory results is frequently unknown, and as a consequence, the data is of screening quality at best.

Screening analytical methods may permit the greater sampling densities needed to establish a good picture of contaminant distributions. Depending on the nature of the decision for which the data might be used, a particular screening method may be sufficient to support decision-making on its own. However, the analytical uncertainty in screening methods is generally too high to support risk assessment decisions and other decisions that require analyte-specific, quantitative data.

Slide 27

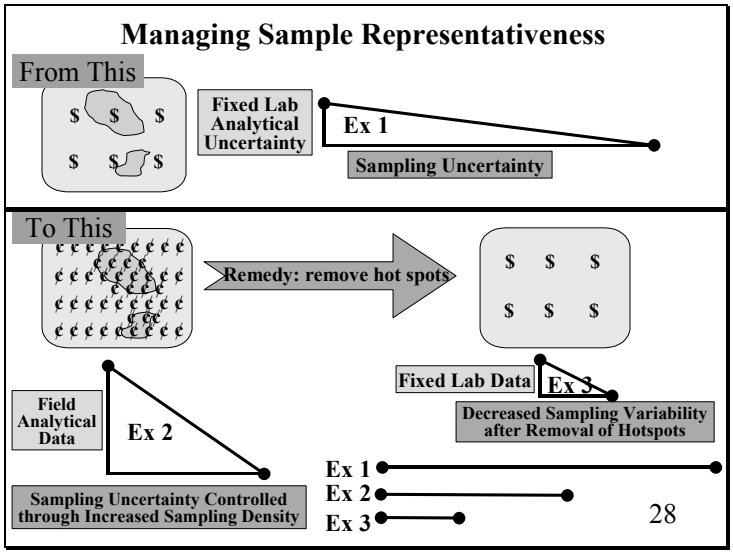


Good data quality at an affordable cost is generated by using both screening and more definitive methods in conjunction with each other. Because of their lower cost, screening methods are best for generating higher data densities that can manage for uncertainty due to environmental heterogeneity (sampling variability). Representative samples can then be selected for more rigorous analysis as needed to manage for remaining analytical uncertainty.

Collaborative data sets complement each other in that uncertainty in one data set is managed by the information in the other. The data sets must be used together to manage all major sources of potential error in the data sets. This is similar to a "weight of evidence" approach.

High density sampling is performed by using cheaper methods (which may be run in the field, although don't have to be). The cheaper methods are often screening methods, but they may be definitive analytical methods (such as field-portable GC-MS for VOCs). After sampling uncertainty is managed, any residual analytical uncertainty needed to meet the desired decision certainty is managed using more rigorous methods (which may be run in a fixed lab, but don't have to be). If there is no residual analytical uncertainty after sampling uncertainty is managed, no more analyses are required, and the second column is not needed. It depends on the nature of the method, the performance of the method with the site-specific matrix, and the nature of the decision to be made on the basis of the data.

Slide 28

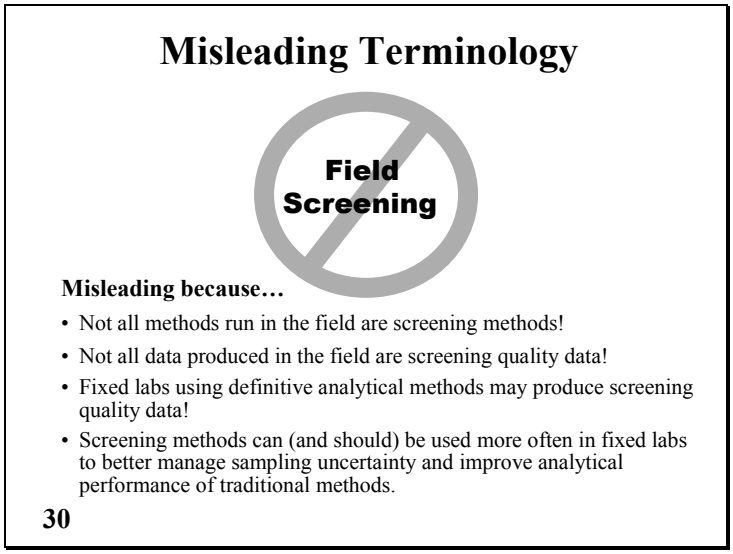


In contrast to the way definitive methods are conventionally used (upper panel), field analytical methods can be used to increase the sampling density, which permits rigorous management of sampling uncertainty (middle panel). Reliable site decisions can be then be made (such as whether to rigorously delineate and remove hotspots of contamination). If needed to meet regulatory requirements for final site closure, follow-on analysis of samples can be performed by definitive, analyte-specific methods. The selection of samples for final closure decisions builds on the previous characterization decisions or cleanup actions to markedly decrease sampling variability in the data set used to support site closure or decisions about regulatory compliance.

Slide 29

Examples of Terminology to Anchor Data Quality Concepts in Uncertainty Management

Slide 30



The term "field screening" is discouraged because it is ambiguous, and it implies things that are not true. The facts are that:

- Definitive methods can be run in the field, and some field technologies are based on definitive methods (such as a field-portable GC-MS).
- Even when screening methods are used, the data may be completely capable of supporting defensible decisions. This is the antithesis of "screening quality data," which indicates there is too much uncertainty to support defensible project decision-making.
- Note that data produced by fixed labs using definitive analytical methods may be of screening quality data if sampling uncertainty is not controlled, if generalized methods are used to report analytes that behave poorly in that generalized method, or if matrix interferences compromise method performance.
- The use of screening methods in fixed laboratories would be highly cost-effective means of increasing sampling density and selecting representative samples for follow-up analysis by more definitive methods. If a dynamic work plan is not being used, rapid turn-around of results would not be needed. However, close coordination with the laboratory to develop and implement an analytical decision tree is advisable.

Slide 31

Proposed Clarification of Terms Data Quality

- **Collaborative data sets =** distinctly different data sets (i.e., produced by different methods that might not be statistically comparable) used in concert with each other to co-manage sampling and analytical uncertainties to an acceptable level. Usually this is the most cost-effective way to generate decision quality data.
- **Decision quality data* = Effective data* =** data shown to be effective for decision-making (see extended definition, slide 32)
- **Screening quality data* =** some useful information is provided; but too much uncertainty present to support decision-making if used alone. [Note: Applies to <u>both</u> excessive analytical or sampling uncertainties. Applies to data produced by definitive analytical methods if the sampling representativeness is not known.]

* Includes sampling uncertainty. Nature of the analytical method irrelevant.

31

Terminology to express data quality concepts should focus on the ability of data to meet project decision-making activities, encouraging explicit identification and management of uncertainties in the data that could lead to decision errors:

1) Collaborative data sets = It is possible that data sets (that by themselves would be considered screening quality) may become part of an effective data set if other data or information is available to manage residual uncertainty to the point where decision-making is defensible when this information is combined. This may sometimes be considered a type of "weight of evidence" approach. Using different techniques to manage various aspects of analytical or sampling uncertainty is often more cost-effectively than trying to manage all relevant data uncertainties using a single technique.
2) Decision quality data = Effective data = data of known quality that can be logically shown to be effective for making defensible project decisions (because BOTH sampling and analytical uncertainties have been controlled to the degree necessary to meet clearly defined project goals). The nature of the analytical method (screening method vs. definitive method) is irrelevant.
3) Screening quality data = Data that provide some useful information, but sampling and/or analytical uncertainties about the data set limit the ability of those data to support defensible project decision-making on their own merits. Again, the nature of analytical method (screening vs. definitive) is irrelevant.

Slide 32

"Effective Data"
"Decision Quality Data"

Data of
known quality
that can be logically demonstrated to be
effective for making the specified decision
because both the
sampling and analytical uncertainties
are managed to the degree necessary to meet clearly
defined (and stated) decision confidence goals

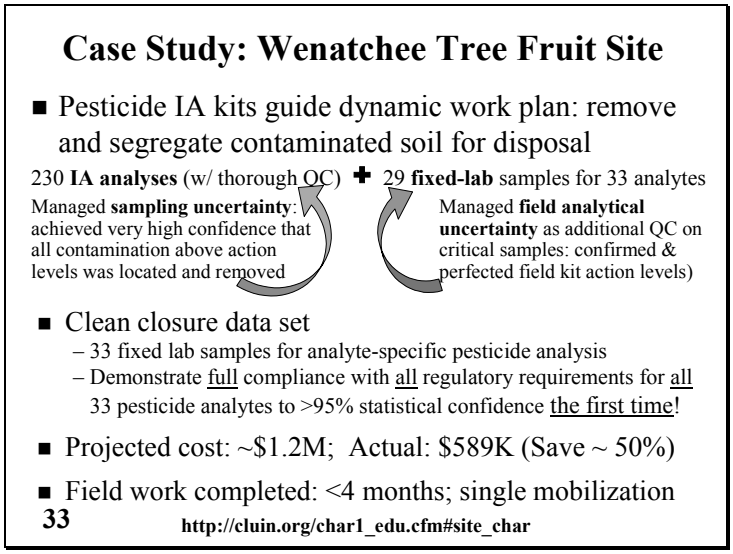
32

The decision(s) that the data are to support must be clearly articulated!! If "decision quality data" or "effective data" (= data effective for decision-making) are claimed on a project-specific basis, the person making the claim must be prepared to document the following: 1) what the decisions are that the data are claimed to support; 2) how good those decisions are supposed to be (the decision confidence goals); and 3) establish that the sampling and analytical quality (i.e., data quality) is known and appropriate to the data use (i.e., making the decision).

- Sample representativeness (location, timing, sample support, subsampling, sample integrity, etc.) must be explicitly considered in the context of the decision to be made.
- The impact of sampling variability must be balanced against the impact of analytical variability.
- Analytical uncertainties (analyte identification, sensitivity, variability in quantitation, the influence of interferences, etc.) must be acknowledged, understood, and managed to the degree needed to achieve the stated decision goals.
- Data of "known quality" means that QC performance and documentation is adequate to meet project-specific needs (QC that is relevant to addressing the analytical uncertainties that bear on the decision).

If this chain of scientific evidence is not built (during planning, implementation, and data interpretation) you run the risk that decisions based on the data will be indefensible if challenged (even if the decisions were actually correct), or the decisions will be erroneous, resulting in wasted effort and expense or failure to protect receptors or both.

Slide 33



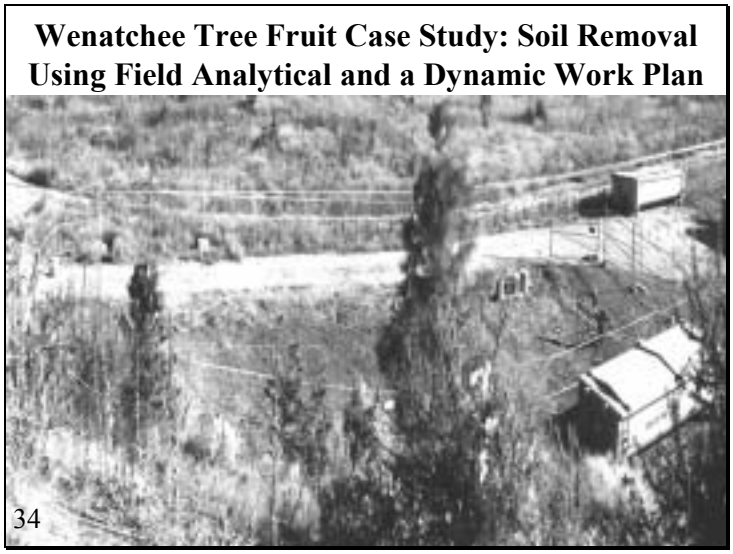
The key features of the project that contributed to its success included:

- Systematic planning accomplished by a team representing the USACE, EPA, the site owners, and state regulators with the appropriate mix of skills and decision-making authority.
- An initial conceptual site model based on a review of historical records from the site. The CSM is refined over the course of the project.
- A dynamic work plan that permitted the field team to make real-time decisions on the basis of data generated in the field.
- A pilot study demonstrated the utility of the field analyses and provided information used to establish site-specific action levels.
- An adaptive sampling and remediation strategy that relied on a combination of field analyses and fixed laboratory data.

The combined benefits of this approach facilitated the "surgical" removal and segregation of contaminated materials and ensured that closure testing would demonstrate regulatory compliance to a high degree of certainty. Significant time and cost savings over the life of the project were possible by making field activities such as sample collection, sample analysis, soil removal, soil segregation, and final disposal of soil and wastewater highly efficient and effective.

The case study report and supporting materials (USACE work plans) can be found at http://cluin.org/char1_edu.cfm#site_char (See entry for "Pesticide Site Cleanup Using a Dynamic Work Plan and Immunoassays")

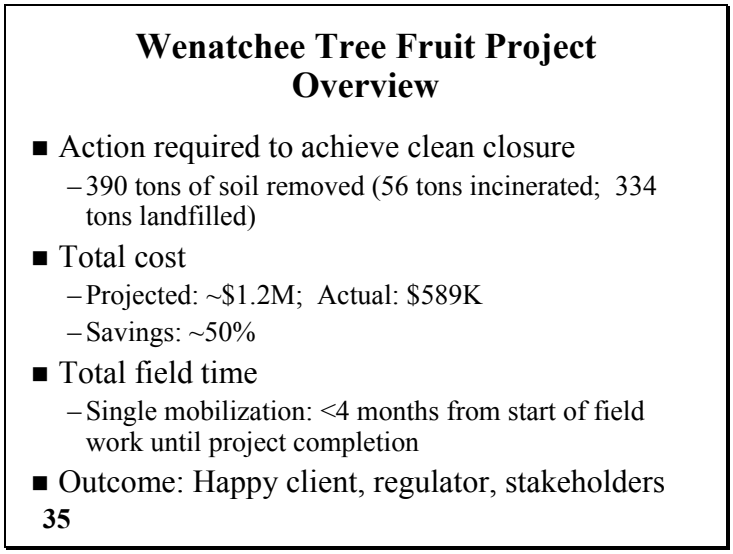
Slide 34



Case Study Example: Project of the USACE Seattle District.

Size of plot roughly 90 ft X 40 ft.

Slide 35

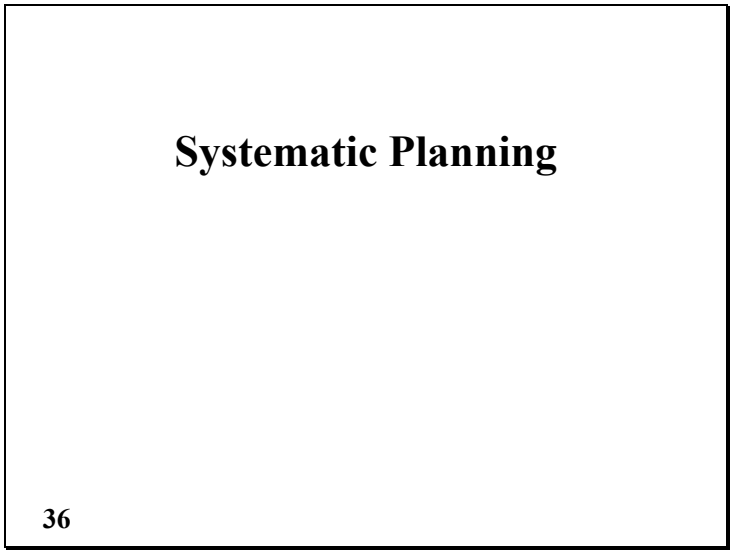


The key features of the project that contributed to its success included:

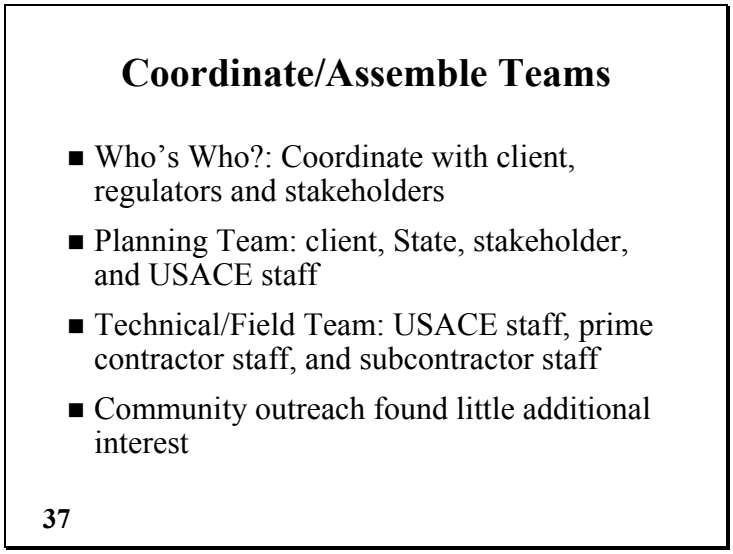
- Systematic planning accomplished by a team representing the USACE, EPA, the site owners, and state regulators with the appropriate mix of skills and decision-making authority.
- An initial conceptual site model based on a review of historical records from the site. The CSM is refined over the course of the project.
- A dynamic work plan that permitted the field team to make real-time decisions on the basis of data generated in the field.
- A pilot study demonstrated the utility of the field analyses and provided information used to establish site-specific action levels.
- An adaptive sampling and remediation strategy that relied on a combination of field analyses and fixed laboratory data.

The combined benefits of this approach facilitated the "surgical" removal and segregation of contaminated materials and ensured that closure testing would demonstrate regulatory compliance to a high degree of certainty. Significant time and cost savings over the life of the project were possible by making field activities such as sample collection, sample analysis, soil removal, soil segregation, and final disposal of soil and wastewater highly efficient and effective.

Slide 36



Slide 37

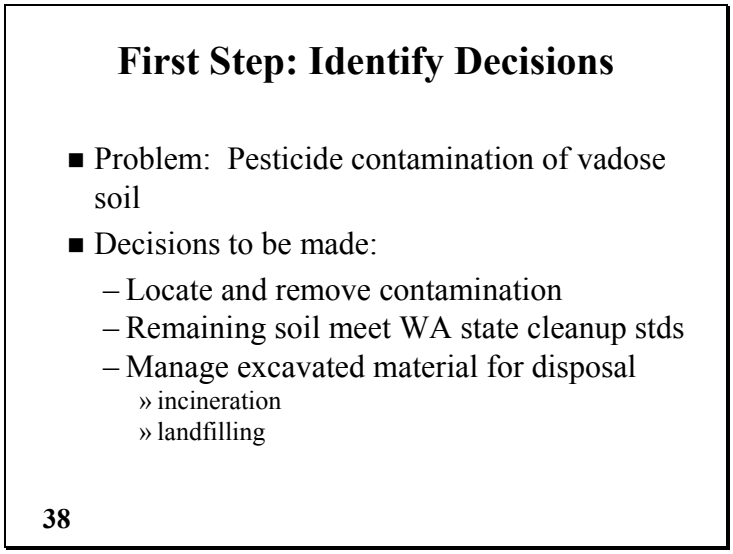


USACE's Client/Customer = EPA Office of Research & Development (ORD)
Regulator = State of Washington Dept. of Ecology's Voluntary Cleanup Program
Stakeholders = 1) Washington State University (WSU, current property owner), represented by 3 people (the WSU Environmental Manager, the WSU Facility Manager, and an Environmental Scientist in charge of the cleanup issue); and 2) the developer for a nearby housing tract

- Planning Team = USACE staff (Project Manager/Team Leader, Project Chemist/Scientist, Project Engineer, Health & Safety Industrial Hygienist, and a Construction Engineer) + 1 rep from EPA ORD + 1 rep from Ecology + 3 reps from WSU

- Field Team = USACE staff (Project Manager/Team Leader, Project Chemist/Scientist, Field QA Officer, Construction/Project Engineer, and H&S) + USACE's Prime Contractor staff (Project Manager, Field Engineer, and Project Chemist/QC Officer) + Subcontractors for a chemistry technician, Geoprobe operator, excavation, and waste disposal

Slide 38



The Test Plot was a former PHS/EPA pesticide test facility, used to examine potential land disposal methods for OC and OP pesticides. It was used from mid-1960's to early 1980's when the property was transferred to WA State University. Contamination from research remained after property transfer. WA State University wanted to use the land for residential development.

WSU and EPA's Emergency Response Team performed investigations of the site in the 1980s and 90s. Several rounds of sampling and fixed laboratory analysis found that high levels of chlorinated pesticides were present amid clean areas. But sampling was always insufficient to delineate the size and location of contaminated areas.

EPA asked the USACE in 1996 to design and implement both the site characterization and remedial action. Funding delays caused the Corp's contractors not to be tasked until mid-1997. Cleanup was performed under the WA State Voluntary Cleanup Program.

Slide 39

Desired Decision Confidence

- Detect contamination
  - Grid size set to detect a 5 ft. x 10 ft. elliptical hotspot
- Remove contamination so that remaining soil meets stringent WA state regulatory cleanup standards:
  - for 33 individual pesticide analytes
  - to a 95% statistical confidence

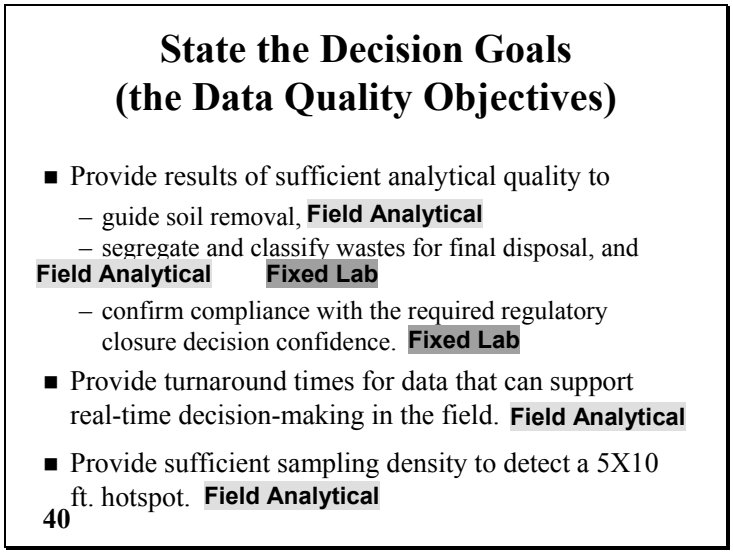
39

Model Toxics Control Act Cleanup (MTCA) Regulation for Washington State governed requirements for cleanup standards.

The final closure confirmation data set (for 33 target pesticide analytes) must pass 3 conditions:

1) The 95% (upper confidence limit) UCL for each analyte must be statistically shown to be less than the cleanup criteria for that analyte. The UCL is determined as the mean of a lognormal distribution if appropriate, or is determined by another statistical measure as instructed by the MTCA requirements (based on the statistical distribution of the data and the size of the sample population). At the same time,
2) On an analyte by analyte basis, the analyte concentration for no more than 10% of the closure samples can exceed the Cleanup Standard for that particular analyte; and
3) No single sample concentration can be greater than 2 X the cleanup standard for any particular analyte.

Slide 40

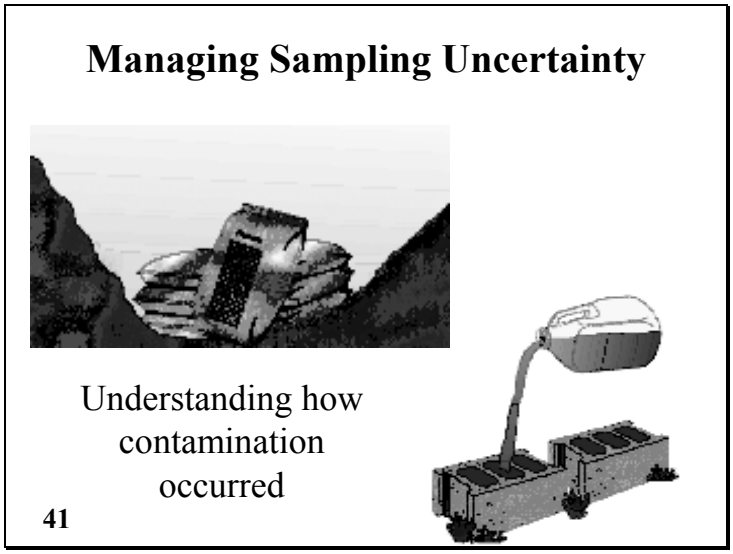


What are the project's decision objectives that the data will be used to support? Only when you know that can you decide what quality of data will be needed to achieve the project goals.

For this project, the DQOs for the overall project were as follows:

- Provide analytical results for DDT, cyclodienes (especially dieldrin and endrin), and other contaminants of concern (see Table 1) with quantitation limits that are less than the field/operational action levels in order to guide the removal of contaminated soil from each defined "column" of soil at the site such that final cleanup goals will be met within a single field mobilization.
- Collect sufficient soil data to confirm that the soil left in place meets the MTCA cleanup standards such that:
  - no more than 10 percent of samples exceed the cleanup standard,
  - no sample can exceed two times the cleanup standard, and
  - the true mean concentration must be below the cleanup standard as measured by a 95% upper confidence limit on the mean.
- Ensure that the turnaround time (TAT) for the field-generated data supports the real-time decision-making needs of the dynamic work plan.
- Provide analytical results that can be used to segregate and classify excavated soil and other remediation wastes for management as solid, hazardous, or dangerous waste according to RCRA and the Washington State Dangerous Waste Regulations.

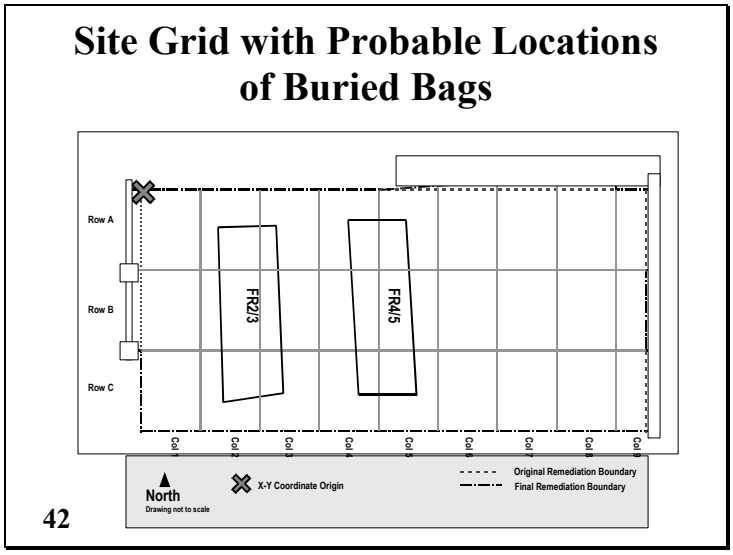
Slide 41



Bags were buried in the Focused Removal (FR) areas. Markers for these 2 areas still existed. These areas were dug up first to remove the bags of neat pesticides while protecting the health of site workers.

To locate contamination that occurred from deliberate spills of pesticide solutions, systematic sampling would be performed. Markers for these areas were no longer present, or were unreliable.

Slide 42



The location of the 6 ft. deep Direct Push (DP) soil core was selected at random within each grid.

Slide 43

Optimize the Data Collection Design

- Use a Dynamic Work Plan
- Use immunoassay (IA) field kits for on-site analysis to guide DWP
- Perform pre-field work pilot study to
  - Assess IA kit suitability
  - Estimate field/IA kit decision/action levels
  - Evaluate Geoprobe performance
  - Prepare SOPs and contingency plans
- Use fixed lab analyses to generate site closure confirmation data sets

43

Develop a dynamic work plan (DWP) using on-site analysis to guide the progressive "surgical" removal of contaminated soils, segregate wastes, and mature (evolve) the CSM by determining the exact boundaries of contamination.

The use of immunoassay (IA) kits provided inexpensive high density sampling to detect and delineate hot spots, while providing TATs of 1 hour or less to guide soil removal and segregation.

- The IA kit used for "DDT" is sensitive to DDT + DDE + DDD (all 3 present at the site), and potentially to other related pesticides and daughter products. DDT was the most persistent of the pesticides disposed at the site, and so was a good indicator of any remaining pesticide contamination.
- The IA kit used for cyclodienes is sensitive to chlordane, dieldrin, and endrin (endrin was known to be present at the site; dieldrin was possibly present), and potentially to other related pesticides and daughter products (such as endosulfan). Endrin was the primary driver at this site for waste soil that needed to be incinerated (as "dangerous waste") vs. lesser contaminated soil that could be landfilled.

A pilot study (i.e., a "demonstration of applicability," per SW-846 recommended terminology) was performed about a year before actual field activities began. This pilot study was used to evaluate several aspects of the project at the same time:

1) Select the most appropriate kit for the site-specific soils. Several IA kits were tested and compared to fixed lab results for thoroughly homogenized split site samples. The kits with the most predictable performance and greatest correlation to fixed lab result sums were selected for use during actual field work. A "perfect" correlation was not expected: not only does the kit respond to several distinct analytes, the relative degree of response to each of those analytes will vary. In addition, no matter how thorough, homogenization will never be perfect.
2) Derive initial IA field action/decision levels. See graphs. The kit results that corresponded to the regulatory standards (or the sum of individual standards) is selected, and then a "safety factor" (to compensate for the degree of uncertainty in interpreting the results) was incorporated to create the field action level. On-going evaluation of the performance of the kits during early stages of the project would allow revision of the field action levels (if needed) in later stages of the project.
3) Evaluate opportunities to optimize kit performance and streamline the sample collection/preparation workload (apply principles of PBMS). Develop kit-specific QA/QC procedures. Prepare project-specific SOPs to document and guide all procedures to be implemented during the project.

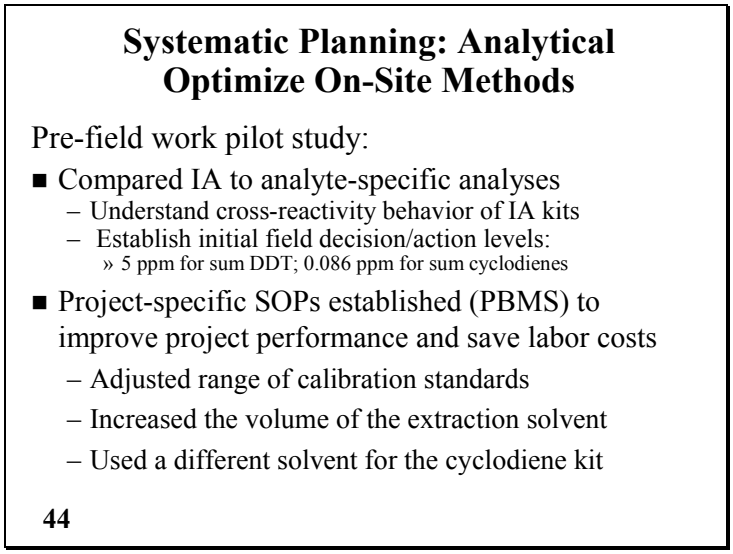
4) Collect site-specific samples for the demonstration of analytical applicability and concurrently evaluate Geoprobe (direct push technology) performance at the site to evaluate the likelihood of refusal.
5) Develop contingency plans in case of equipment breakage or failure during sample collection, preparation, and analysis.

The selective use of very expensive fixed lab analyses for pesticides will be reserved for

- building confidence in the field action levels used with each IA kit (originally developed from the results of the pilot study) for field decision-making (accomplished by splitting critical decision samples for confirmatory analysis); and
- establishing regulatory compliance by creating a clean closure data set (quantitative data for all 33 target analytes needed) after IA testing determined that compliant soil had been reached.

Contingency plans for sampling and analysis were developed in case Geoprobe or IA performance was found to be unacceptable, or to accommodate equipment breakage/failure.

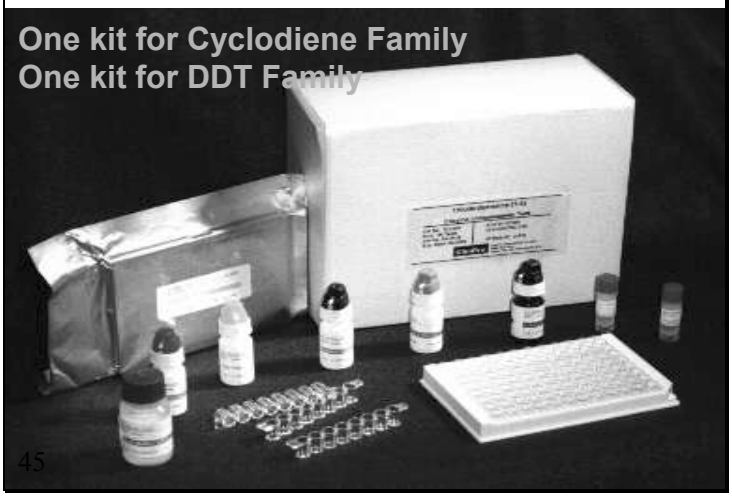
Slide 44



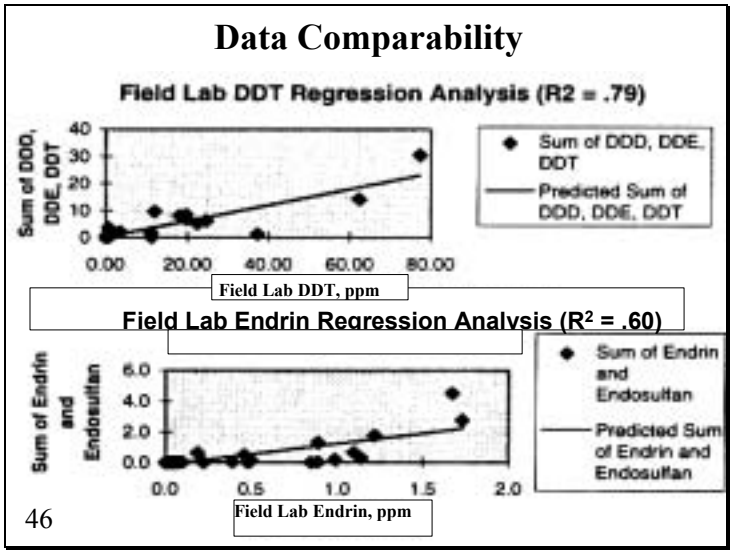
- The pilot study confirmed that IA kits are intentionally biased 100% high by the manufacturer. Comparison of the IA results with fixed lab results allowed a prediction of site-specific, decision-specific action levels for the IA kits to use to guide the DWP in the field.
- Project-specific Standard Operating Procedures (SOPs) were developed for the site-specific IA kits using a Performance-Based Measurement System (PBMS) approach. This was done by
  - adjusting the calibration procedures so that the most accurate part of the 3-point IA calibration curve (the center of the curve) fell on the field action level established for each kit;
  - doubling the extraction volume to 20 mL to better bracket the field action levels;
  - modifying the IA sample extraction method for the cyclodiene test kit from the package insert recommendation (for a water/methanol mix) to pure methanol. This made it possible to run both the DDT and the cyclodiene IAs off the same soil extract, decreasing the amount of labor and sample prep materials involved without sacrificing data quality.

Slide 45

Immunoassay Kits



Slide 46



Fixed lab (definitive) data is plotted on the vertical axis (the y-axis), and IA data is plotted on the horizontal axis (the x-axis). Points that fall below the regression line are instances where the IA kit has a higher than predicted result. Points that fall above the regression line are instances where the IA kit has a lower than predicted result. Note that for DDT, an IA value of 20 ppm predicts a fixed lab result for the sum(DDT, DDD, and DDE) of about 8 ppm.

Recall that the field action level for the DDT kit was initially 5 ppm (later raised to 10 ppm); and the field action level for the cyclodiene kit was 0.086 ppm. MTCA action levels: DDT—2.9 ppm; DDD—4.2 ppm; DDE—2.9 ppm; endrin—0.4 ppm; endosulfan—480 ppm.

On-going comparison between the IA results and fixed lab results revealed:

1) To be meaningful for regression analysis, the IA results must be compared to fixed lab results that sum all of the cross-reacting analytes that are present in the sample.
2) The cyclodiene kit correlated fairly well with the sum of endrin and a closely related pesticide, endosulfan, which was also consistently present in the samples. The points along the x-axis that fall below the regression line were due to the fact that other cross-reacting (by the IA analysis) pesticides other than endrin and endosulfan were sometimes present in the samples. Points that fall above the line are likely due to an inability to completely homogenize the sample prior to splitting it.
3) The DDT IA results below 10 ppm correlated well with the mix of individual DDT, DDE and DDD concentrations that did not exceed their respective MTCA standards. As a result, the IA field action level for DDT was refined to 10 ppm. The 10 ppm DDT action level was used during the gross removal phase (Phase 4) to determine the need for further excavation.

Slide 47

QA/QC for IA Kits

- 3-point calibration & CCV w/ each batch (12 samples)
- Reagent blank
- Matrix duplicate (intra-laboratory sample split)
- LCS [prepared from a purchased (soil) PE sample]
- Split sample confirmation analysis (by fixed lab analysis) for samples representing critical decision points
  - Excavation boundaries
  - Clean closure data set for regulatory compliance

47

Slide 48

Systematic Planning: Analytical Optimize Off-site Methods

Certain fixed lab methods for pesticides were optimized using PBMS principles:

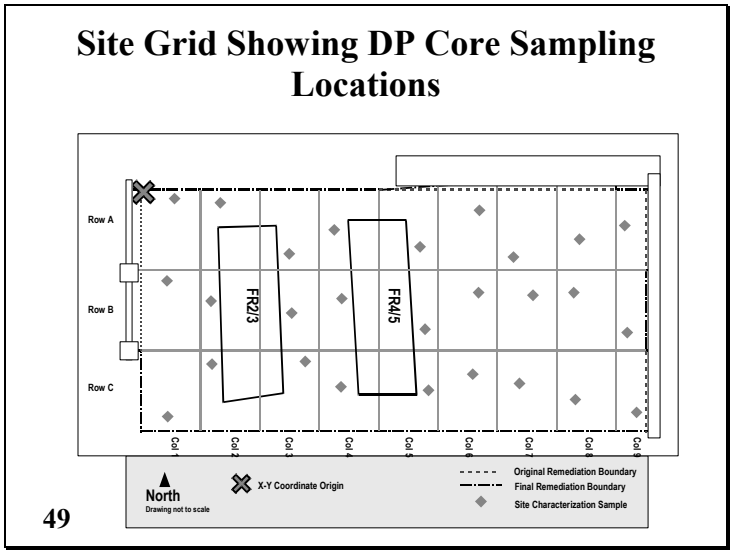
- Organophosphorus (OP) pesticides:
  - SW-846 Method 8141 (GC/NPD) was changed to SW-846 Method 8270 (GC/MS)
- Carbamates by GC:
  - A blend of EPA Water Method 632 and SW-846 Method 8141 (GC/NPD) was used
- Paraquat in soil by spectrophotometry:
  - An industry developed method was used

48

PBMS (Performance-Based Measurement System) optimization was done for some fixed lab methods to better serve specific project needs:

- OP Pesticides: The detector was changed from a nitrogen-phosphorus detector (NPD) to a mass spectrometer (MS). This improved the selectivity of the method for the OP pesticides in the site-specific samples.
- Carbamates by GC: The project-specific detection limit requirements and restricted analyte list allowed the less sensitive but more selective GC/NPD technique to be used (rather than GC/MS or HPLC). Analytical performance was thus improved due to the reduction of site-specific matrix interferences. Also, a different surrogate compound was selected for the method to accommodate the pesticides routinely used in the geographical area.
- Paraquat: A method developed by Chevron Oil (a manufacturer of paraquat) was used.

Slide 49



The location of the 6 ft. deep Direct Push (DP) soil core was selected at random within each grid.

Slide 50

Dynamic Work Plan Decision Matrix for Characterization (Geoprobe Core) Testing

Scenario #	0-12"	12 to 24"	24 to 36"	36 to 48"	48 to 60"	60 to 72"	Action
1	No	n/a	n/a	n/a	n/a	n/a	Confirmation Sampling
2	Yes	No	n/a	n/a	n/a	n/a	Find contamination in 0-12" sample, field sample 12-24" sample. Find no contamination in 12-24" sample above MTCA: Remove 0-12" of soil. Confirmation Sampling. No Further Action.
3	Yes	Yes	No	n/a	n/a	n/a	Find contamination in 0-12" sample, field sample 12-24" sample. Find contamination in 12-24" sample, field sample 24-36" soil sample. Find no contamination in 24-36" sample above MTCA: Remove 0-24" of soil. Confirmation Sampling. No Further Action.
4	Yes	Yes	Yes	No	n/a	n/a	Find contamination in 0-12" sample, field sample 12-24" sample. Find contamination in 12-24" sample, field sample 24-36" soil sample. Find contamination in 24-36" sample, field sample 36-48" soil sample. Find no contamination in 36-48" sample above MTCA: Remove 0-36" of soil. Confirmation Sampling. No Further Action.
5	Yes	Yes	Yes	Yes	No	n/a	Find contamination in 0-12" sample, field sample 12-24" sample. Find contamination in 12-24" sample, field sample 24-36" soil sample. Find contamination in 24-36" sample, field sample 36-48" soil sample. Find contamination in 36-48" sample, field sample 48-60" soil sample. Find no contamination in 48-60" sample above MTCA: Remove 0-48" of soil. Confirmation Sampling. No Further Action.
6	Yes	Yes	Yes	Yes	Yes	No	Find contamination in 0-12" sample, field sample 12-24" sample. Find contamination in 12-24" sample, field sample 24-36" soil sample. Find contamination in 24-36" sample, field sample 36-48" soil sample. Find contamination in 36-48" sample, field sample 48-60" soil sample. Find contamination in 48-60" sample, field sample 60-72" soil sample. Find no contamination in 48-60" sample above MTCA: Remove 0-60" of soil. Confirmation Sampling. No Further Action.

50

This is one of the two decision logic matrices for sampling the plot. One decision matrix was developed for the Focused Removal area. The matrix displayed here is the one that dealt with characterizing all the other exposure columns OUTSIDE of the Focused Removal areas, where surface disposal or incidental contamination (such as from accidental spills) could have occurred. (This characterization would occur after the bags of pesticide product had been removed.) This is a tabular representation of a series of "if/then" statements used to guide decision-making related to the removal and segregation of soil based on real-time IA results.

Since contamination in this area would have occurred from the surface down, it is reasonable to assume that clean overburden is a strong predictor that deeper layers would be clean as well. Geoprobe cores to 6 ft. depth (3-ft. X2) were to be collected. The 6 1-ft lifts would be prepared as samples at the same time, but would be analyzed by IA in sequential order, starting with the surface (0-12") sample. Based on the presence of contamination, the designated action would be taken.

Read chart as (Scenario 1) "If no contamination above the field action level was detected in the top 1-ft. lift, then confirmation sampling was performed at the surface and no further action was taken." (The samples taken from the deeper lifts would be archived so they would be available for later testing, just in case.)

Scenario 2: "If contamination was detected in the first lift of the Geoprobe core, then the 2nd 1-ft lift sample was analyzed.
If the 2nd 1-ft lift sample is not contaminated, then no deeper lift samples are analyzed. (Those samples will be archived.) A confirmation sample will be taken from the surface of the remaining soil.

Confirmation sampling consisted of analyzing the sample by IA. If IA results above the field action levels were detected, another lift of soil was removed, and confirmation testing was repeated. If IA results were below the action level, a sample was sent off to the fixed lab for analysis for all COCs. For any column, the grid (recall there are 3 grids in each column) with the highest IA result was selected to be sampled for confirmation testing.

Slide 51

Dynamic Work Plan for Removing Contaminated Soil

- IA results on DP core samples used to develop backhoe excavation profile
  - Profile correctness later confirmed by fixed lab results
- After backhoe excavation complete, floor analyzed by IA
  - If IA results > field action level, more soil removed by hand.
- New floor tested by IA.
  - When IA results < field action level, sample for fixed lab analysis collected (for clean closure data set)

51

After the designated amount of soil was removed by the backhoe from each removal column, samples of the floor of the excavation were taken and analyzed by IA. At least one sample was taken from the floor of each grid (recall that a column contained 3 grids). Judgmental sampling was used if there appeared to be any indication of contamination. If none was visible, the sampling location within the grid was selected at random. If an IA field action level was exceeded in a grid, another layer of soil was removed from that grid by hand (shovel), and another IA sample analyzed, until compliant soil was reached.

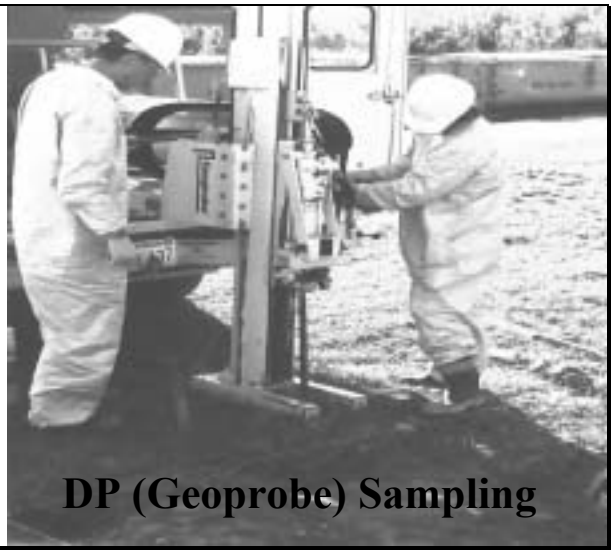
Slide 52

Implementing the Work Plan

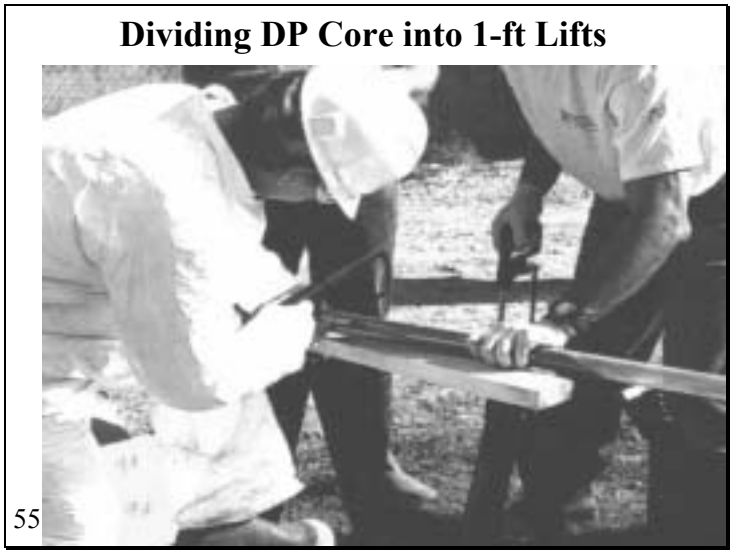
Slide 53

Excavating to remove buried bags of pesticide pure product from FR4/5.

Slide 54

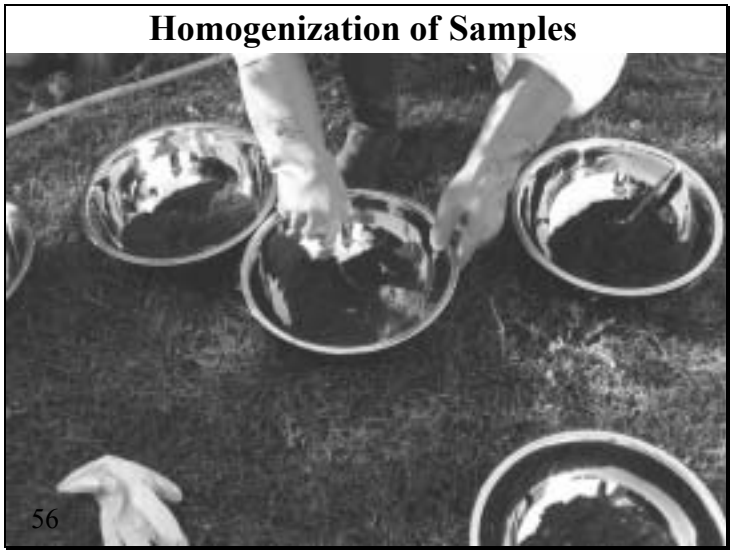


Slide 55



Each 1-ft lift was homogenized as a single sample.

Slide 56



Preliminary homogenization of a sample involved debris (sticks, roots, stones larger than 3/8-inch diameter) removal, crushing and mixing.
Final homogenization involved coning and quartering:

- soil drawn up into a cone,
- then split into 4 separate cones (1 cone → 4 cones)
- each small cone stirred with stainless steel spoon to even consistency
- 2 cones mixed back together into one larger cone X2 (4 cones → 2 cones)
- resulting 2 cones mixed back together (2 cones → 1 cone)

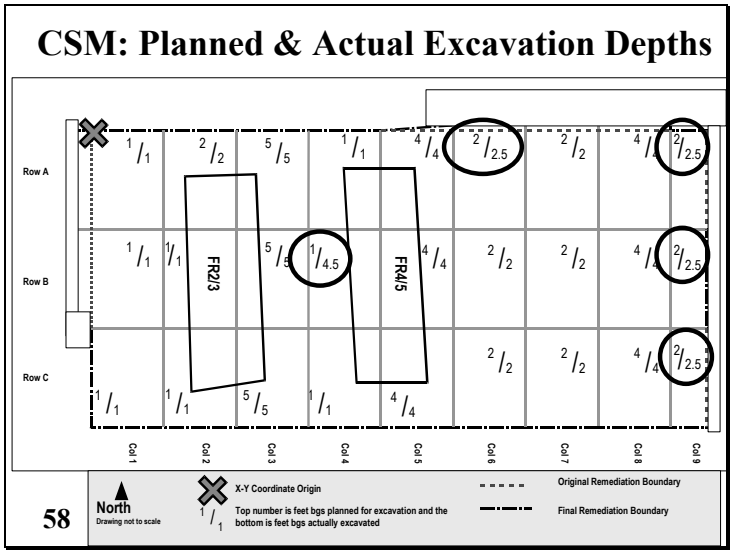
The sample was then split into 3 4-oz. jars.
Care was taken throughout the process to keep samples from being exposed to direct sunlight.

Slide 57

Collecting a Clean Closure Sample

57

Slide 58



Partially mature CSM, showing extent of vertical excavation. The first number (X/X) in each grid represents the predicted (planned) excavation depth based on the contamination profile developed from IA analysis of the DP cores. The second number represents the final (actual) excavation depth. For all but 5 grids (those circled in pink), the predicted depth was the same as the final depth.

Process: After the profiled depth (equal to the first number) was excavated by backhoe, the floor of the excavation was tested using IA. If kit response was found higher than the kit's action level, more soil was removed in that grid using a hand shovel. IA testing and removal continued until compliant soil by IA testing was reached. The second number represents this final excavation depth. Clean closure confirmation samples (for fixed lab analysis of all 33 target analytes) were collected after all non-compliant soil was found and removed per IA.

Slide 59

Unexpected Contamination? No Problem!

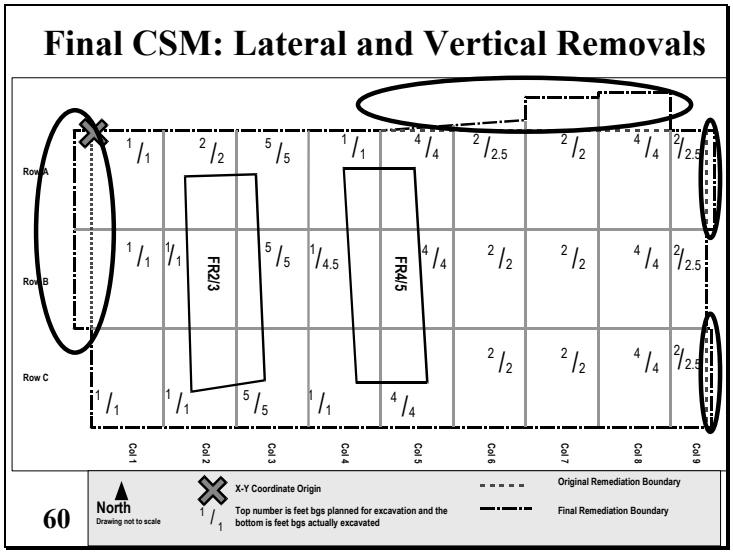
- After floor excavation completed, State regulator asked for sidewall testing (beyond scope of original work plan)
- Unexpected contamination found in some pit sidewalls (outside of original site boundaries)
- IA guided sidewall delineation and excavation
  - IA results accepted as sole data to establish clean closure for 80% of the sidewall area

59

Immediately after excavation and sampling of the floor was completed, closure confirmation sampling of the sidewalls by IA was begun. Contamination of the sidewalls that extended beyond the original boundaries of the site was discovered, and IA was used to delineate the extent of this contamination.

Efforts to obtain approvals to modify the scope of the work plan to accommodate the additional sampling plan and excavation of the sidewalls were immediately begun. Beginning 10 days after the existence of sidewall contamination was discovered, excavation of the delineated sidewall contamination was completed in 3 days. 60 tons of soil were removed from the sidewalls along 3 sides of the test plot. Confidence in the reliability of decisions based on the IA tests was now high enough (by the end of the project) that only 4 of the 20 IA closure confirmation samples (refer to Figure 5 of EPA Case Study Report) were sent for fixed lab analyses to ensure that compliant soil had indeed been reached.

Slide 60



Completely mature CSM, showing extent of sidewall excavation required after IA testing of sidewalls discovered contamination beyond the site boundaries. Sidewall excavations were done to the same depths as the grids adjacent to them.

Slide 61

Data Effective for Making Decisions

- Decision: Locate and remove contaminated soil
  - 230 IA analyses; 29 fixed lab samples as confirmatory QC
  - Outcome: Very high degree of certainty that all significant contamination located and removed
- Decision: Demonstrate clean closure
  - 33 fixed lab samples for analyte-specific pesticide analysis
  - Outcome: Demonstrated full compliance with all regulatory requirements for all 33 analytes to a 95% (or better) confidence
  - 16 IA for sidewalls (kit decision level accepted as "clean")
- Decision: Dispose of soil (RCRA Subtitle C requirements)
  - Fixed lab analyses (TCLP OC pesticides, total OP and OC pesticides, TCLP metals) demonstrate compliance

Slide 62

Wenatchee Tree Fruit Project Triad Approach Successes

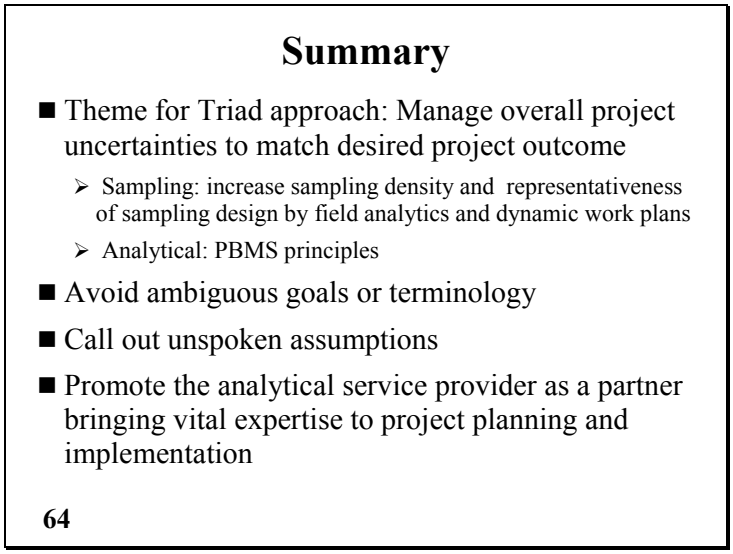
- Systematic planning focused on end-use of data
- State regulator focused on project outcome/performance, permitting flexibility to maximize innovation and resource savings.
- Demonstration of method applicability (a pilot study) determined:
  - appropriate field sampling & measurement tools
  - project-specific field action levels for decisions
  - project-specific SOPs and QC for field and fixed lab analyses

Slide 63

Wenatchee Tree Fruit Project Triad Successes (continued)

- IA analysis increased the number, density, and information value of samples.
- CSM refined/matured in the field; specific sampling strategies selected to match specific decisions
- Final Cost: Site characterized (incl. analytical PBMS), remediated, closed (very high confidence), restored, and waste disposed for < ½ projected cost
  - Waste disposal: 56 tons incinerated; 334 tons landfilled
- Time: < 4 months of field work

Slide 64

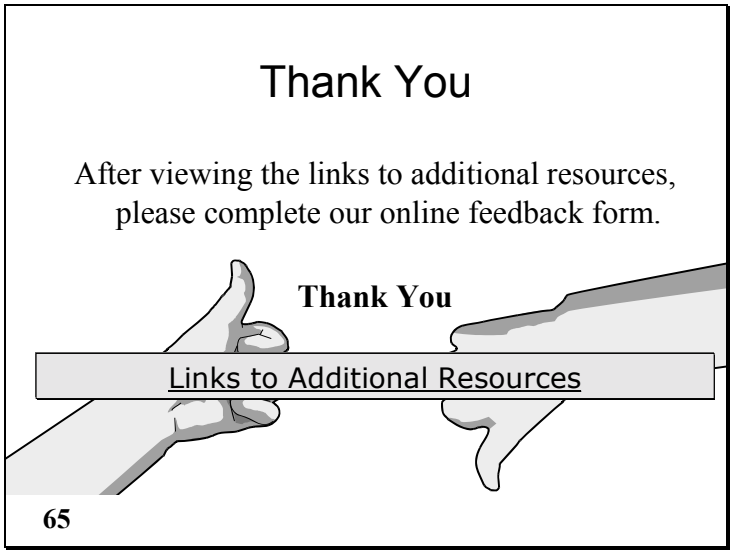


Manage overall project uncertainty to match the desired project outcome

- ID causes of uncertainty, and assess their impact on decision quality
- Develop strategies to manage uncertainties due to
  - Sampling: increase sampling density and improve representativeness of sampling design
  - Analytical: PBMS

Even if field kits are operated by a technician, a qualified chemist is needed to select appropriate kits, to design SOP and QC, and to ensure proper interpretation of the results.

Slide 65


Thank You
After viewing the links to additional resources, please complete our online feedback form.
Thank You
Links to Additional Resources
65