Data Review, Verification, and Validation

Data Review or Verification

- Data Review - has many meanings in environmental context. For today's purpose, focus will be on quality control (QC) data generated from field and laboratory QC samples and external samples (such as performance evaluation samples).

- Verification - confirmation by examination and provision of objective evidence that specified requirements have been fulfilled.

Preliminary Questions to Ask

- Are the identifications and numerical results correct?
- Are there sufficient, documented results to determine accuracy and precision?
- Was laboratory performance acceptable?
- Are there any significant technical, documentation, or other problems which may invalidate the results?

Quality Control Checks

- Contamination (blanks)
- Accuracy (spikes, performance samples)
- Precision (replicates)
- Sensitivity (detection limits)

Equipment or Rinsate Blank

- A sample created by rinsing sampling equipment after it has been cleaned.
- Frequency: usually 1:10 or each day.
- Purpose: help identify contamination due to decontamination procedures, ambient field conditions, storage conditions, or laboratory problems.
- Action(s) to consider: discount (do not correct) positives; fix decon procedures; check w/lab; possible resample.

Field Blank

- Sample created by adding distilled or deionized water to a container in field. Used when using dedicated or disposable equipment.
- Frequency: usually 1:10 or each day.
- Purpose: help identify contamination due to ambient field conditions, bottles/storage conditions, or laboratory problems.
- Action(s) to consider: discount (but don't correct) positives, check bottles, check w/lab; possible resample.

Trip Blank

- ✓ Volatile free water placed in VOA vial by lab and sent to field with bottles.
- ✓ Frequency: One per shipping container.
- ✓ Purpose: Identify contamination from transit, bottles, or laboratory conditions.
- ✓ Action(s) to consider: Re-evaluate shipping protocols, check w/lab.

Reagent Blank

- Sample generated by laboratory to demonstrate reagents are free of contamination.
- Frequency: whenever new batch of reagents received; not all labs run, few report to clients.
- Purpose: Identify contamination in common chemicals used in laboratory.
- Action(s) to consider: laboratory should take action with suppliers; reagents should not be used.

Laboratory or Method Blank

- ✓ Sample generated by laboratory and introduced at beginning of sample processing (digestion, extraction, etc.).
- ✓ Frequency: 1:batch or 1:20 samples
- ✓ Purpose: Identify contamination introduced within laboratory.
- ✓ Action(s) to consider: discount (but do not correct). positives; check w/lab; redo analysis; resample.

Field Matrix Spike

- ✓ Known amounts of an analyte or representative compounds are added to sample in field. Sample submitted blind. This is effectively a PE sample. Uncommon QC sample.
- ✓ Frequency: if run, once per sampling event.
- ✓ Purpose: test laboratory performance and ability to obtain correct results.
- ✓ Action(s) to consider: check w/lab to assess whether can perform method, look at other QC (lab MS, LCS).

Laboratory Matrix Spike (MS)

- ✓ Known amounts of an analyte or representative compounds are added to sample(s) in laboratory.
- ✓ Frequency: 1:20 or 1:batch
- ✓ Purpose: identify whether lab has performed method properly or if sample matrix is introducing a positive or negative bias.

Laboratory Matrix Spike (MS) *(continued)*

- Action(s) to consider: check w/lab; determine whether result due to matrix problem or lab problem (look at LCS results, if OK = matrix; see whether a 2nd sample was prepared and run, if 2nd result out=matrix, problem, if in=lab problem). Make sure not someone else's sample due to batch QC. Monitor future site results for pattern. Be aware of matrix bias in results. Determine if spiked sample representative of all samples.

Laboratory Control Sample (LCS) aka: Blank Spike or Laboratory Fortified Blank

- ✓ Known amounts of an analyte or representative compounds are added to a "clean" matrix (lab water or clean sand) in laboratory.
- ✓ Frequency: 1:20 or 1:batch
- ✓ Purpose: identify whether lab has performed method properly.
- ✓ Action(s) to consider: request lab reanalyze all samples in batch associated with LCS if haven't already; possible resample at lab cost; use results w/caution.

Instrument Spike

- ✓ Known amounts of an analyte or representative compounds are injected directly in instrument
- ✓ Frequency: as needed when contamination suspected
- ✓ Purpose: determine losses of material due to instrument
- ✓ Action(s) to take: Nothing. Typically not reported to client.

Post Digestion Spike

- ✓ Metals spike made after digestion procedure. Used in method of standard additions to correct for matrix effects.
- ✓ Frequency: usually as needed.
- ✓ Purpose: permits calculation of results for metals although a matrix effect exists.
- ✓ Action(s) to consider: not a QC sample per se, used for quantitation.

Surrogate Spike

- Known amounts of organic compounds, similar in behavior to target analytes, are added to samples before processing.
- Frequency: in every sample.
- Purpose: mimic behavior of target compounds and identify either matrix or extraction problems.
- Action(s) to consider: if all surrogates out, require re-extraction. If some out, look at similarities to targets. Re-extraction is possible option. If sample all gone, may need to resample.

Co-Located Sample

- ✓ Second sample collected at same location (water, air) or nearby (soil, sediment). Sent blind to laboratory.
- ✓ Frequency: usually 1:10, may not collect if collecting replicates.
- ✓ Purpose: determine heterogeneity of matrix, reproducibility of sample technique and laboratory performance.
- ✓ Action(s) to consider: expand number of samples or area sampled in future events or resample. Check laboratory duplicates or matrix spike duplicates to make sure looking at field variability, not laboratory.

Field Replicate (duplicate)

- A sample divided into two homogeneous parts.
- Frequency: 1:10
- Purpose: determine reproducibility of sub-sampling technique and laboratory performance
- Action(s) to consider: check laboratory duplicates or matrix spike duplicates to make sure looking at field variability, not laboratory. Check field sampling procedures. In extreme cases, resample.

Matrix Spike Duplicate (MSD)

- A known amounts of an analyte or representative compounds are added in the laboratory to a second aliquot of sample used for matrix spike.
- Frequency: 1:20 or 1:batch.
- Purpose: determine laboratory reproducibility or precision. MSD is used because often samples do not contain organic compounds so no results are available on which to do precision calculations.
- Action(s) to consider: check LCSD results. View results with caution and be sensitive to upper and lower range of concentrations. Check whether a batch QC, although this less critical to

Laboratory Control Sample Duplicate (LCSD)

- ✓ Known amounts of an analyte or representative compounds are added to a second "clean" matrix (lab water or clean sand) in laboratory. Duplicate of LCS.
- ✓ Frequency: 1:20 or 1:batch.
- ✓ Purpose: determine laboratory precision without matrix effects.
- ✓ Action(s) to consider: reanalysis of all samples in batch. Resample at lab cost.

Laboratory Duplicate

- ✓ Second processing and analysis of sample. Usually for general chemistry or organic analyses.
- ✓ Frequency: 1:20 or 1:batch.
- ✓ Purpose: determine laboratory precision.
- ✓ Action(s) to consider: check w/lab. Check LCSD results (may not be available for inorganics). View results with caution and be sensitive to upper and lower range of concentrations. Check whether batch QC, although this less critical to precision than to bias due to recovery problem.

Field Split

- A field replicate/duplicate that is sent to a second laboratory.
- Frequency: seldom, usually only if problems develop in previous work.
- Purpose: used as a check on laboratories.
- Action(s) to consider: check laboratory QC results. Consider PE samples. Attempt to determine which lab accurate. Determine which lab to be kept.

Laboratory Split

- A laboratory created replicate/duplicate that is sent to a second laboratory.
- Frequency: seldom, mainly when problem suspected.
- Purpose: determine inter-laboratory precision. Independent assessment of laboratory problems in primary laboratory.
- Action(s) to consider: check laboratory QC results. Consider PE samples. Attempt to determine which lab accurate and should be kept.

Single Blind Performance Evaluation (PE) Sample

- ✓ Known amounts of an analyte or organic compounds are provided to lab in a labeled vial or bottle.
- ✓ Frequency: once a quarter, once a sample shipment, or not at all. Depends on a number of factors.
- ✓ Purpose: check laboratory's ability to perform analysis under optimum conditions.

Single Blind Performance Evaluation (PE) Sample

- Action(s) to consider: lab should pass when it knows it is being tested. Consider suspension of work if doesn't pass. At minimum, lab should demonstrate how it will address problem.

Double Blind Performance Evaluation (PE) Sample

- ✓ Known amounts of an analyte or organic compounds are provided to lab, but are introduced with samples so lab is not aware of presence.
- ✓ Frequency: once a quarter, once a sample shipment, or not at all. Depends on a number of factors.
- ✓ Purpose: check laboratory's ability to perform analysis without its know it is being tested.

Double Blind Performance Evaluation (PE) Sample

- ✓ Action(s) to consider if QC criteria not met: lab should pass when it knows it is being tested. Consider suspension of work if doesn't pass. At minimum, lab should demonstrate how it will address problem.

Instrument Detection Limit (IDL)

- An analyte injected directly into instrument. Either calculation of the standard deviation of 7 replicates or a signal/noise determination made.
- Frequency: usually once a year, can be more frequent.
- Purpose: determine lowest concentration of an analyte that an instrument can detect.
- Action(s) to consider: since only determines instrument capability, usually not used by client, method detection limit or reporting limit more useful.

Method Detection Limit (MDL)

- An analyte or organic compounds injected into a clean matrix and processed like a sample. Either calculation of the standard deviation of 7 replicates or a signal/noise determination made.
- Frequency: Usually once a year.
- Purpose: determine lowest concentration of an analyte a laboratory can detect.
- Action(s) to consider: compare MDL to action levels or regulatory standard to ensure will be able to make required decisions. Consider alternative methods or laboratory if unable to reach objectives.

Quantitation Limit (QL)
aka: Reporting Limit

- MDL "bumped" up to a level where lab feels confident all positives are real. Usually a factor of 2 to 10 times MDL.
- Frequency: calculated value after MDL study.
- Purpose: insurance that laboratory is reporting only analytes it detects with confidence.
- Action(s) to consider: compare QL to action levels or regulatory standard to ensure will be able to make required decisions. Consider alternative methods or laboratory if unable to reach objectives. Consider having laboratory

Data Validation

What Is Data Validation?

- Data Validation is a review of analytical data based on the analytical method which may result in qualification of results.
- Data validation, from EPA's perspective, is a systematic, rigorous, thorough, and defined process, usually conducted by a third party, independent of the data generator.
- Data validation reports present the data user with annotated results which may be used to determine if results are of sufficient quality to meet project objectives.
- Data validation results assist the decision maker or data user in determining the usability of data.

Why Perform Data Validation?

- ✓ Data validation helps verify that data are of known and documented quality.
- ✓ Data validation is a tool to help in determining whether results are appropriate for the intended use as defined by the data quality objectives in QA Plans or Sampling Plans.

When Should Data Validation Be Considered?

- ✓ Data to be used in a risk assessment
- ✓ Data to be used for Superfund site scoring
- ✓ Data which might be subjected to a higher level of scrutiny or challenge such as in support on an enforcement situation

What Is Data Validation Based On?

- ✓ National Functional Guidelines (NFGs) (Contract Laboratory Program (CLP) analyses).
- ✓ Procedures developed consistent with NFGs for methods not covered by NFGs.
- ✓ Procedures developed independently by validating organization.
- ✓ Professional judgment of the reviewer.

Where are Data Validation Criteria Defined?

- ✓ Contract Lab Program Statements of Work
- ✓ Quality Assurance Plans
- ✓ Sampling and Analysis Plans
- ✓ Analytical Methods
- ✓ Laboratory Standard Operating Procedures

What Constitutes a "Data Package"?

- ✓ Sample Results
- ✓ Quality Control Data
- ✓ Calibration Data
- ✓ Standards Data
- ✓ Raw Data (instrument output)

Organic Data Validation Evaluation Factors

- ✓ Holding Times
- ✓ Matrix Spikes/LCS/ PE Samples
- ✓ Target Compound Identification (GC/MS)
- ✓ QL (CLP = CRQL)
- ✓ Tentatively Identified Compounds (GC/MS)
- ✓ System Performance
- ✓ Tune Performance (GC/MS)
- ✓ Initial Calibrations
- ✓ Continuing Calibrations
- ✓ Blanks (Lab/Field)
- ✓ Surrogates
- ✓ Internal Standards (GC/MS)
- ✓ Overall Assessment

Inorganic Data Validation Evaluation Factors

- ✓ Holding Times
- ✓ Initial And Continuing Calibration Verification
- ✓ CRDL/QL Check Stds
- ✓ Calibration Blanks
- ✓ Laboratory And Field Blanks
- ✓ Laboratory Control Sample
- ✓ Laboratory And Field Duplicates
- ✓ Matrix Spike
- ✓ Post Digestion Spike / Analyte Addition Test
- ✓ Serial Dilution
- ✓ Internal Standards
- ✓ Quantitation
- ✓ Overall Assessment

Region 9 Tiered Structure For Data Validation

- ✓ Tier 1 validation level is a summary forms review.
- ✓ Tier 2 is a focused review, usually of selected analytes or sample locations.
- ✓ Tier 3 is considered the “traditional” full validation.