

# Labeling Externally Validated Data: The Black Box Opened !

Anand Mudambi  
US Environmental Protection Agency

*NEMC  
August 2009*



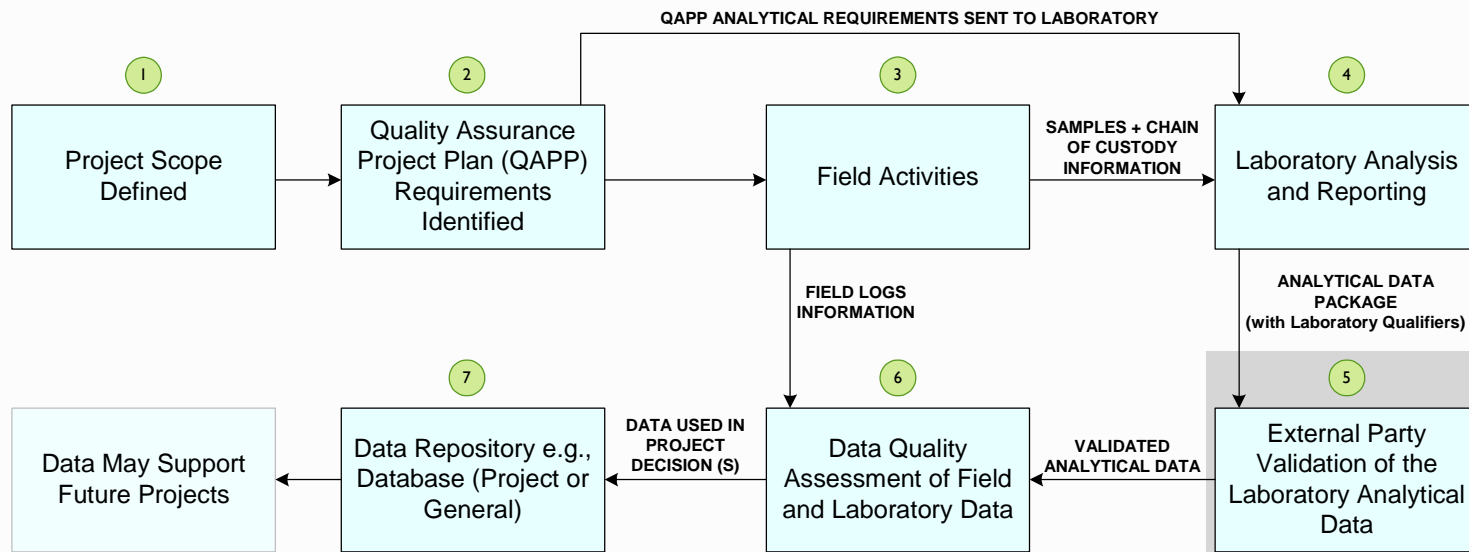


OR

- Telling me what you did when you validated my data AND not taking all day about it!!!



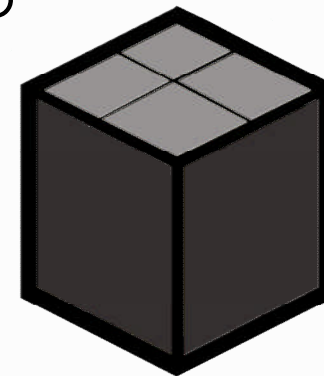
# Typical Project Data Generation and Review Process





## The Analytical Data Validation “Black Box”

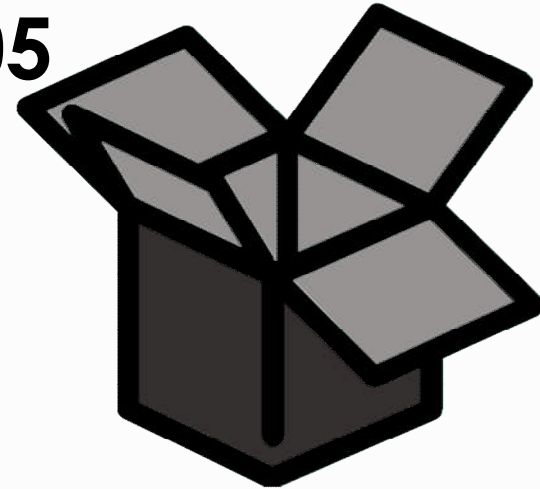
- Each year Federal Programs spend considerable resources to validate laboratory analytical data.
- Validation guidelines vary from program to program.
- No consistent mechanism to indicate what was actually checked during the validation process.





## Opening up the Black Box !

- **Guidance for Labeling  
Externally Validated  
Laboratory Analytical  
Data for Superfund Use  
EPA-540-R-08-005  
January 2009**





## Guidance Document Stats

- Title: *Guidance for Labeling Externally Validated Laboratory Analytical Data for Superfund Use*
- Location on the web:  
<http://www.epa.gov/superfund/policy/pdfs/EPA-540-R-08-005.pdf>
- Length: 21 pages
  - Main document: 7 pages
  - 5 Appendices: 14 pages



## Guidance Goals

- Improve communication within Superfund about scope and content of lab analytical data verification and validation.
- Encourage appropriate use of data for
  - Task at hand
  - Future decisions
- Note: Guidance can be used by other programs and agencies





## Approach Taken

- Facilitate communication through the use of “labels” that summarize verification and validation checks.
- Checks are grouped into stages.
- Each stage builds on the checks from previous stage.
- Labels also describe nature of review process (manual and/or electronic)
- Based on the general guidelines used for data validation



## Stages of Validation Checks

- Completeness
- Compliance
  - Sample-related QC
  - Instrument-related QC
- Recalculation
- Instrument output review

Note: Examples of checks in each stage are given in Appendix A of the Guidance Document.



## Completeness Checks

- To make sure that the requested data deliverables are provided.
- To determine that data requested are actually present in the deliverables.
- To ensure consistency within the deliverable (e.g., between hardcopy and electronic deliverables).





## Compliance Checks

- To compare analytical Quality Control (QC) results with the acceptance criteria, requirements or guidelines present in the regional data validation documents, analytical methods or contract.
  - *Sample-Related QC* (e.g., blank contamination, surrogate recoveries)
  - *Instrument-Related QC* (e.g., instrument calibrations, tunes)





## Recalculation Checks

- The laboratory reported values (e.g., sample results, instrument calibration results) are verified by recalculation using instrument output data reported by the laboratory.
- Confirms that correct formulae and values were used in calculation of results.





## Instrument Output Checks

- Actual instrument outputs should be checked to ensure that the laboratory reported analytes have been correctly identified and quantitated (e.g., are mass spectra properly identified? Are peak integrations correct?).





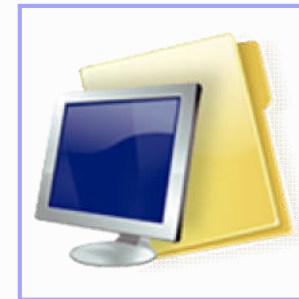
## Validation Checks, Stage, and Labels

Checks	Validation Stage	Labels
Completeness	Stage 1	S1V
Completeness & Sample QC	Stage 2a	S2AV
Completeness, Sample QC & Instrument QC	Stage 2b	S2BV
Completeness, Sample QC, Instrument QC & Recalculations	Stage 3	S3V
Completeness, Sample QC, Instrument QC, Recalculations & Instrument Outputs	Stage 4	S4V



## Process Labels

- Electronic review only E
- Manual validation M
- Electronic and manual EM





## Example Labels For Validated Data

- Stage 2b validation by electronic tools only:
  - **S2BVE**
- Stage 3 validation by both electronic and manual processes:
  - **S3VEM**

Note: Labels combining all validation stages and processes are given in Appendix B of the Guidance



## Desired Outcome

- Third party reviewers/validators associate the validated data with its validation stage as data is shared with decision makers.
- Data users quickly recognize the nature of review performed on data prior to use.
- Future use of data is facilitated by labels that travel with data (e.g., kept associated with data in repositories)



## Benefits of Providing Validation Labels

- Opens up the “Validation Black Box” by telling data recipients and users in a short and succinct manner:
  - Checks used to validate the data
  - Process used to validate the data





## Benefits of Providing Validation Labels

- Allows for automation of some stages of the data validation process (e.g., use of electronic tools for compliance checks and recalculations)
- Helps integrate manual and electronic processes used for data validation





## Next Steps

- Provide guidance on implementation of this guidance
  - Short paper on benefits and suggested contract language
- Ensure data repositories have fields for the labels which are linked to the validated analytical data





## Working on Implementation

- EPA Regions 4 and 10
- EPA's Electronic Data Exchange and Evaluation System (EXES) and US Army Corps of Engineers' Automated Data Review (ADR) software



## FINAL STEP!!!

- UP TO YOU
- Repeat after me: “I will never use validated data that does not have a label attached to it!!!”





## Contact Information

- Anand Mudambi  
Phone: 202-564-2817  
Email: [mudambi.anand@epa.gov](mailto:mudambi.anand@epa.gov)

