

National Human Exposure Assessment Survey (NHEXAS)

Arizona Study

Quality Systems and Implementation Plan for Human Exposure Assessment

The University of Arizona
Tucson, Arizona 85721

Cooperative Agreement CR 821560

Standard Operating Procedure

SOP-UA-D-24.0

Title: Cleaning: Arizona Lab Data

Source: The University of Arizona

U.S. Environmental Protection Agency
Office of Research and Development
Human Exposure & Atmospheric Sciences Division
Human Exposure Research Branch

Notice: The U.S. Environmental Protection Agency (EPA), through its Office of Research and Development (ORD), partially funded and collaborated in the research described here. This protocol is part of the Quality Systems Implementation Plan (QSIP) that was reviewed by the EPA and approved for use in this demonstration/scoping study. Mention of trade names or commercial products does not constitute endorsement or recommendation by EPA for use.

Cleaning: Arizona Lab Data

1.0 Purpose and Applicability

The purpose of this procedure is to define the particular steps involved in cleaning the electronic data generated from data entry of the Arizona Lab Data forms. It applies to electronic data corresponding to the Arizona Lab Data Forms that were scanned and verified by the Data Staff for NHEXAS Arizona, the Border Study, or other Health and the Environment projects.

2.0 Definitions

- 2.1 **BORDER STUDY:** An alias for "Total Human Exposure Arizona: A Comparison of the Border Communities and the State" conducted in Arizona by the University of Arizona/Battelle/Illinois Institute of Technology consortium.
- 2.2 **DATA CLEANING:** The process of locating and correcting data processing and field technician errors. They can be individual level errors in the electronic and physical data, or they can be system level errors in the data collection, packaging, coding, entry, and cleaning procedures themselves. This process is also referred to as "data validation."
- 2.3 **DATA, ELECTRONIC:** Data stored on some type of magnetic or optical medium (for example: floppy disk, hard disk).
- 2.4 **DATA, ENTERED:** Electronic data scanned into a data file using Teleform scanning software. Entered data are the product of "data entry."
- 2.5 **DATA, VERIFIED:** Electronic data that has cleared through the Teleform Verification process. In the Verification process Teleform reviews all of the entered data and displays any possible errors. These potential errors are reviewed by a Data Technician. Once all of the errors are fixed the data is saved to an ASCII file.
- 2.6 **DATA CLEANING BATCH:** A collection of electronic data, along with their corresponding physical forms. Data cleaning batches are formed after one or more data processing batches (see DATA PROCESSING BATCH below) are scanned. The data cleaning batches are then cleaned (see DATA CLEANING) and appended to the master database (as described in UA-D-44.X). Each data cleaning batch is assigned a numeric descriptor of the form MMDDYY, where MM is the month the batch was created, DD is the day the batch was created, and YY is the year the batch was created. If more than one batch is created on the same day, each batch after the first is assigned a descriptor of the form MMDDYY_N, where N denotes the batch as being the Nth batch created that day.
- 2.7 **HEALTH AND ENVIRONMENT PROJECTS (or H&E):** An umbrella title for all projects funded to M.D. Lebowitz and/or M.K. O'Rourke (or their designees) which examine purported or real relationships among environmental factors and any aspect of human health.
- 2.8 **HRP SITE:** The **H**ealth **R**elated **P**rofessions building, located at 1435 North Fremont Avenue, Tucson, AZ 85719. This is an annex of the Arizona Prevention Center and the primary site of NHEXAS Arizona, the Border Study, or other Health and the Environment projects.

- 2.9 MASS DATA MESSAGE PROCESS (or MDM) = The data processing program used by NHEXAS Arizona, the Border Study, and other Health and Environment projects.
- 2.10 NHEXAS Arizona: Acronym for National Human EXposure Assessment Survey, a research project conducted in Arizona by the University of Arizona/Battelle/Illinois Institute of Technology consortium.

3.0 References

SOP UA-D-44.X (Operation Manual of the Mass Data Massage Program) (This Volume)

4.0 Discussion

A copy of the Arizona Lab Forms are given in Appendices A, B, C, D, E, F, G, H, and I. These copies have each variable labeled for easy reference. Questionnaires and Forms are reviewed for completeness and accuracy prior to being transferred to the data section of the project. These checks are designed to generate flawless forms and questionnaires prior to data entry. Once entered and verified the data is subject to logic checks through the application of Dictionaries and other projects. Dictionaries for the lab forms have not been generated. Once generated these dictionaries will define the files structures for the Arizona Lab Forms databases, as well as the logic and range checks that are performed on each field. For further reference, see SOP# UA-D-4.X.

Sometimes despite all best efforts, errors are found in data. These errors must be corrected to produce a valid database. At this time, all anticipated problems have been addressed. "Cleaning" protocols are developed concurrently with data entry to address the unanticipated problems. Specific corrections will be documented as outlined in SOP # UA-D-38.X.

5.0 Responsibilities

The Project Data Coordinator is responsible for writing the data cleaning instructions for the Arizona Lab Forms when data problems are found during the data entry process.

6.0 Materials and Reagents

- 6.1 Local Area Network
- 6.2 Purple Pen
- 6.3 Copy of Data Dictionary for specific form.

7.0 Procedure

7.1 Steps Followed

I. Data Technicians

A. Entering data

1. Scan and verify the form as described in UA-D-34.X.

B. Cleaning Data

1. Log into UNIX environment.
2. Type "data" to run the *MDM*.
3. Select form type:
 - a. The user will enter "2" for Lab Data
 - b. The *MDM* will then prompt the user for a specific form. The user will then enter:
 - i. "1" for 25mm and 37mm pre-weighing
 - aa. If this option is selected, the user will be prompted to select either the header or the detail groups. The user will enter "1" for the header, and "2" for the detail group.
 - ii. "2" for 25mm and 37mm post-weighing
 - aa. If this option is selected, the user will be prompted to select either the header or the detail groups. The user will enter "1" for the header, and "2" for the detail group.
 - iii. "3" for Sentinel filter pre-weighing
 - aa. If this option is selected, the user will be prompted to select either the header or the detail groups. The user will enter "1" for the header, and "2" for the detail group.
 - iv. "4" for Sentinel filter post-weighing
 - aa. If this option is selected, the user will be prompted to select either the header or the detail groups. The user will enter "1" for the header, and "2" for the detail group.
 - v. "5" for XRF analysis
 - vi. "6" for Soil characterization
 - vii. "7" Vacuum Dust characterization
 - viii. "8" for 24-Hour Food Diary Check
 - aa. If this option is selected, the user will be prompted to select either the header, the detail groups, the second page, or the supplement. The user will enter "1" for the header, "2" for the detail group, "3" for the second page, or "4" for the supplement.
 - ix. "9" for Vacuum filter pre-weighing
 - aa. If this option is selected, the user will be prompted to select

either the header or the detail groups. The user will enter "1" for the header, and "2" for the detail group.

x. "0" to return to the previous menu

4. The *MDM* shall prompt the user for an action. The user then selects an operation from a given list. Operations include:
 - a. Create a new batch
 - b. Clean existing batch
 - c. QA an existing batch
 - d. [Examine Descriptive Master Database]
5. If the user selects the Clean existing batch option:
 - a. The *MDM* will list all possible batches to clean.
 - b. If there are no batches, the program will inform the user, and prompt the user to press the <enter> key, which will return the user to the previous menu.
 - c. If there are batches to be cleaned, the user will be prompted for a specific batch.
 - d. Once a batch is selected, it is preprocessed (checking for errors) for cleaning, and the user is given information regarding the status of the preprocessing.
 - e. If no errors are found within in the data, the *MDM* will send mail to the Data Coordinator informing him of the newly cleaned batch.
 - f. If errors are found in the data, the user is informed as to the number of errors found within the batch.
 - i. For each error found in the batch, the user will be shown the key variables to locate the record containing the error, which is displayed between dashed lines.
 - ii. The user is then prompted for an action, the action being either the changing the erroneous value, the skipping of the error, a manual change to a different variable in the current record, a manual change to any variable in any record, or the user may quit.
 - g. If the user wishes to change the erroneous value:
 - i. Then "C" must be entered.
 - ii. The variable name is then displayed, and the *MDM* prompts for a new value.
 - iii. Once a new value is given, the user is given the option to accept the value given, to accept a value formatted by the *MDM*, or to abort the operation.
 - iv. If the operation is not aborted, the user is prompted for a reason for the change. Once a reason is given, the update is completed.
 - h. If the user wishes to skip the error:
 - i. He/she must enter "S".
 - ii. This will cause the *MDM* to skip the current error, but not the current

record.

- i. If the user wishes to change the value of a different variable in the current record:
 - i. He/she must enter "M".
 - ii. The user will then be prompted for the number of changes to be made.
 - iii. Once this number is entered, the *MDM* will then prompt for the variable to be changed.
 - iv. Once a valid variable is specified, the user is prompted for the new value.
 - v. Once a new value is given, the user is given the following options:
 - aa. Accept the value given
 - bb. Accept a value formatted by the *MDM*
 - cc. Abort the operation.
 - vi. If the operation is not aborted, the user is prompted for a reason for the change. Once a reason is given, the update is completed.
- j. If the user wishes to change the value of a variable in a different record:
 - i. He/she must enter "R".
 - ii. The user will then be prompted for the number of changes to be made.
 - iii. Once this number is entered, the *MDM* will then prompt for the key variable values of the record to be modified.
 - iv. Once valid values are given, the *MDM* prompts the user for the variable to be changed.
 - v. Once a valid variable is specified, the user is prompted for the new value.
 - vi. Once a new value is given, the user is given the following options:
 - aa. Accept the value given
 - bb. Accept a value formatted by the *MDM*,
 - cc. Abort the operation.
 - vii. If the operation is not aborted, the user is prompted for a reason for the change.
 - viii. Once a reason is given, the update is completed.
- k. If the user wishes to quit, he/she must hit the "Q" key. The user is then returned to the menu defined in I.B.4.

8.0 Records

- 8.1 All records are automatically generated by the *MDM*.
- 8.2 Records of all the forms in a cleaning batch are printed out when the batch is created. This list

is then attached to the cleaning batch.

- 8.3 Records of the changes made to the data are located in the following directory tarred with their associated batch: /rsc53/NHEXAZdata/master/labaz/<Specific Form>.
- 8.4 Records of the cleaning batches which have been appended to the master data base are located in the following directory: /rsc53/NHEXAZdata/master/labaz/<Specific Form>. The list is kept in the file "read.me".
- 8.5 All changes to the hard copy of the form must be dated and completed in purple or red ink.

Inclusions:

- Appendix A: Food Diary Check Form (2 pages)
- Appendix B: Sentinel Filter Pre-Weighing Form (1 page)
- Appendix C: Sentinel Filter Post-Weighing Form (1 page)
- Appendix D: Teflo Filter Pre-Weighing Form (1 page)
- Appendix E: Teflo Filter Post-Weighing Form (1 page)
- Appendix F: Vacuum Filter Pre-Weighing Form (1 page)
- Appendix G: Vacuum Dust Characterization (1 page)
- Appendix H: Soil Characterization (1 page)
- Appendix I: XRF Analysis (2 pages)

PAGE 2

24-Hr Food Diary Check

HHID: <u>HHID</u> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>	F.S. <input type="checkbox"/>	Start Date: <u>Evntdate</u> <input type="text"/> / <input type="text"/> / <input type="text"/>	Sample ID#: <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>	Status <input type="checkbox"/>
<div style="display: flex; justify-content: space-between;"> HHIDFS lig-ID lig-stat </div>				
1. Liquid Sample ID#: <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> Status <input type="checkbox"/>				
2. Total Volume: <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> ml <u>Tot-Vol</u>				
QC: <input checked="" type="checkbox"/> []				
3. Alcohol <u>Alcsmell</u> <u>Alcdiary</u> A. Smell: Y <input type="checkbox"/> N <input type="checkbox"/> N/A (def.) <input type="checkbox"/> B. Diary: Y <input type="checkbox"/> N <input type="checkbox"/> N/A (def.) <input type="checkbox"/>				
COMMENTS: _____				
Comments: _____				

Formstat

Office Use Only

Form Status:	<input type="radio"/> 1.Cmp <input type="radio"/> 2.N Cmp <input type="radio"/> 3.P Cmp <input type="radio"/> 4.Re-col <input type="radio"/> 5.Ref <input type="radio"/> 7.Dest <input type="radio"/> 8.N/A <input type="radio"/> 9.Miss	Tech. ID <u>QCBV</u> MO <u>QDATE</u> DAY <u>QDATE</u> YR <u>QDATE</u>	Tech. ID <u>DEBY</u> MO <u>DEDATE</u> DAY <u>DEDATE</u> YR <u>DEDATE</u>	
	QC: <u>QCBV</u>	DE: <u>DEBY</u>	DP Batch: <u>DPBATCH</u>	QXV: <u>LFOD1</u>
	QA: <u>QABV</u>	DP Batch: <u>DPBATCH</u>		
Chain of custody initiated (sig.): _____				
Consigned to packet on []: ____/____/____ Box UA G4-2.0				



SENTINEL FILTER PRE-WEIGHING BACK-UP FORM

Formstat

SENTINEL FILTER POST-WEIGHING BACK-UP FORM

Item#

Form Type: 123

FORM UA-L-9.0-4.0

Tech. ID

Technician: Tech ID

Init.

Eventdate

Equilibration Start Date: 7/7/

Eventtime

Equilibration Start Time: :

Postdate

1. Post-Weighing Date:

Posttime

2. Post-Weighing Time: :

Temp

3. Temperature:

Humidity

4. Relative Humidity: %

FILE NAME:

Scale

6. Scale:

AE 163 (UA ID# A195535)

AE 166 (UA ID# A201335)

#V-1200 (ID# 50090339)

N/A (def.)

COMMENTS:

COMMENT CODES

Standard Weight

Sample ID#

Weight

Status

QC

1. Samples

2.

3.

4.

5.

6.

7.

8.

9.

10.

Std-ID1

Std-ID2

Formstat

Office Use Only

Form Status:

1. Cmp

2. N Cmp

3. P Cmp

4. Re-cool

5. Ref

7. Dest

8. N/A

9. Miss

Tech. ID

MO

DAY

YR

QC: QCBY

QC DATE: / /

QA: QABY

QA DATE: / /

Tech. ID

MO

DAY

YR

DE: DEBY

DE DATE: / /

DP Batch: DPBATCH

QXV: L S P O 1

Chain of custody initiated (sig.):

Consigned to packet on: [] / /

Box UA G4-2.0

12986

Appendix D: Teflo Filter Pre-Weighing Form

25mm and 37mm TEFLO FILTER PRE-WEIGHING BACK-UP FORM

Form Type: 120 FORM UA-L-9.0-1.0	Technician: Tech ID Init. Tech ID	Equilibration Start Date: Event date / / Equilibration Start Time: Event time : :
1. Pre-Weighing Date: Pre date		Post-Weighing Technician: Post tech Init. Tech. ID
2. Pre-Weighing Time: Pre time		All Weights Are Recorded In Grams (g)
3. Temperature: Temp		
4. Relative Humidity: Humidity %		
5. FILE NAME: _____		
6. Scale: Scale <input type="radio"/> AE 163 (UA ID# A195535) <input type="radio"/> AE 166 (UA ID# A201335) <input type="radio"/> #V-1200 (ID# 50090339) <input type="radio"/> N/A (def.)		
COMMENTS: _____ _____ _____ _____ _____ _____ _____ _____ _____ _____		Sample ID# Pre weigh Weight Status QC
		2.
		3.
		4.
		5.
		6.
		7.
		8.
		9.
		10.
COMMENT CODES: Comment 1 Comment 2 Comment 3		Standard Weight Std-ID1
Std-ID2 		

Formstat

Office Use Only

Form Status: <input type="radio"/> 1. Cmp <input type="radio"/> 2. N Cmp <input type="radio"/> 3. P Cmp <input type="radio"/> 4. Re-col <input type="radio"/> 5. Ref <input type="radio"/> 7. Dest <input type="radio"/> 8. N/A <input type="radio"/> 9. Miss	QC: QCBY QA: QABY Init. 	Tech. ID MO DAY YR QCDATE / / QADATE / / 	Tech. ID MO DAY YR DE: DEBY Init. DP Batch: DPBATCH QXV: LPPR1
---	---	--	--

Chain of custody initiated (sig.): _____

Consigned to packet on: [] ____/____/____

Box UA G4-2.0

28614



25mm and 37mm TEFLO FILTER POST-WEIGHING BACK-UP FORM

Chain of custody initiated (sig.): _____

Consigned to packet on: [] / / Box UA G4-2.0

3324



SOIL CHARACTERIZATION

8670

Appendix I: XRF Analysis

Itemnum		TechID		XRF ANALYSIS												
Form Type <div style="border: 1px solid black; padding: 2px; display: inline-block;">1 2 4</div>		Tech: <div style="border: 1px solid black; width: 20px; height: 20px; display: inline-block;"></div> <small>Init.</small>		Analysis Date: <div style="border: 1px solid black; width: 20px; height: 20px; display: inline-block;"></div> / <div style="border: 1px solid black; width: 20px; height: 20px; display: inline-block;"></div> / <div style="border: 1px solid black; width: 20px; height: 20px; display: inline-block;"></div>				HHID: <div style="border: 1px solid black; width: 20px; height: 20px; display: inline-block;"></div> F.S. <div style="border: 1px solid black; width: 20px; height: 20px; display: inline-block;"></div>								
FORM: UA-L-10.0-1.0				Analysis Time: <div style="border: 1px solid black; width: 20px; height: 20px; display: inline-block;"></div> : <div style="border: 1px solid black; width: 20px; height: 20px; display: inline-block;"></div> <small>Evntdate</small>				HHID <div style="border: 1px solid black; width: 20px; height: 20px; display: inline-block;"></div> HHIDFS <div style="border: 1px solid black; width: 20px; height: 20px; display: inline-block;"></div>								
				Header: QC: <input checked="" type="checkbox"/> []												
1. Sample ID #: <div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> Status: <input type="checkbox"/>												3. Application: <input type="radio"/> 1. Thin Film <input type="radio"/> 2. Soils <input type="radio"/> 88. N/A (def.)				
2. Source Time: Fe-55: <div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> Sec. Fe-time Cd-109: <div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> Sec. Cd-time Am-241: <div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> Sec. Am-time												4. XRF Cup #: _____				
Element	READING												STANDARD DEVIATION			
Pb	<div style="border: 1px solid black; width: 100px; height: 20px; display: inline-block;"></div> Pb-read												<div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> Pb-std			
As	<div style="border: 1px solid black; width: 100px; height: 20px; display: inline-block;"></div> As-read												<div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> As-std			
Cd	<div style="border: 1px solid black; width: 100px; height: 20px; display: inline-block;"></div> Cd-read												<div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> Cd-std			
Ni	<div style="border: 1px solid black; width: 100px; height: 20px; display: inline-block;"></div> Ni-read												<div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> Ni-std			
Cr	<div style="border: 1px solid black; width: 100px; height: 20px; display: inline-block;"></div> Cr-read												<div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> Cr-std			
Ba	<div style="border: 1px solid black; width: 100px; height: 20px; display: inline-block;"></div> Ba-read												<div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> Ba-std			
Mn	<div style="border: 1px solid black; width: 100px; height: 20px; display: inline-block;"></div> Mn-read												<div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> Mn-std			
Se	<div style="border: 1px solid black; width: 100px; height: 20px; display: inline-block;"></div> Se-read												<div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> Se-std			
V	<div style="border: 1px solid black; width: 100px; height: 20px; display: inline-block;"></div> V-read												<div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> V-std			
Cu	<div style="border: 1px solid black; width: 100px; height: 20px; display: inline-block;"></div> Cu-read												<div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> Cu-std			
Zn	<div style="border: 1px solid black; width: 100px; height: 20px; display: inline-block;"></div> Zn-read												<div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> Zn-std			
K	<div style="border: 1px solid black; width: 100px; height: 20px; display: inline-block;"></div> K-read												<div style="border: 1px solid black; width: 40px; height: 20px; display: inline-block;"></div> K-std			



Element	READING		STANDARD DEVIATION	
Ca	<input type="text"/>	Ca-read	<input type="text"/>	Ca-std
Co	<input type="text"/>	Co-read	<input type="text"/>	Co-std
Fe	<input type="text"/>	Fe-read	<input type="text"/>	Fe-std
Mo	<input type="text"/>	Mo-read	<input type="text"/>	Mo-std
Tl	<input type="text"/>	Tl-read	<input type="text"/>	Tl-std
Ag	<input type="text"/>	Ag-read	<input type="text"/>	Ag-std
Sr	<input type="text"/>	Sr-read	<input type="text"/>	Sr-std
U	<input type="text"/>	U-read	<input type="text"/>	U-std
Th	<input type="text"/>	Th-read	<input type="text"/>	Th-std
Sn	<input type="text"/>	Sn-read	<input type="text"/>	Sn-std
W	<input type="text"/>	W-read	<input type="text"/>	W-std
Ti	<input type="text"/>	Ti-read	<input type="text"/>	Ti-std
Rb	<input type="text"/>	Rb-read	<input type="text"/>	Rb-std
Ir	<input type="text"/>	Ir-read	<input type="text"/>	Ir-std
Hg	<input type="text"/>	Hg-read	<input type="text"/>	Hd-std

Formstat

Office Use Only

Form Status: <input type="radio"/> 1. Cmp <input type="radio"/> 2. N Cmp <input type="radio"/> 3. P Cmp <input type="radio"/> 4. Re-coil <input type="radio"/> 5. Ref <input type="radio"/> 7. Dest <input type="radio"/> 8. N/A <input type="radio"/> 9. Miss	Tech. ID		MO	DAY	YR	Tech. ID		MO	DAY	YR
	QC: QCBY		QC DATE			DE: DEBY		DE DATE		
	QA: DABY		QA DATE			DP Batch:		QXV: LXRF		
	Init.		Init.			Init.		Init.		

Chain of custody initiated (sig.):

Consigned to packet on: [] / /

Box UA G4-2.0

Data Use	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	G	H	I	J
Only:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

41358

