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-15.0**

Title: Data Entry and Data Verification (Hand Entry)

Source: The University of Arizona

Data Entry and Data Verification

1.0 Purpose and Applicability

The purpose of this procedure is to define the steps involved in data entry and data verification of physical forms. It applies to the data entry and data verification of all physical forms generated by the NHEXAS Arizona Project, but particularly to those processed at Keypunch.

2.0 Definitions

2.1 BACKUP: (v.) The process of creating a duplicate of a file, directory, or drive to protect against data loss during a hardware or software failure. (n.) The duplicate copy created during this process.

2.2 CODE, GLOBAL: A set of standard codes used in data within the Respiratory Sciences Center designating the status of a data field in three cases: datum refused, datum non-applicable, and datum missing. It can also be a standard approach to coding questions with the same response structures.

2.3 DATA: Classified under this word are the following: DATA, ELECTRONIC; DATA, ENTERED; DATA, PHYSICAL; DATA PROCESSING BATCH; DATA RECORD; DATA, VERIFIED.

   2.3.1 DATA, ELECTRONIC: Data stored on some type of magnetic or optical medium (for example: floppy disk, hard disk, Bernoulli, tape).

   2.3.2 DATA, ENTERED: Electronic data entered for the first time into a computer database. Entered data are the product of "data entry."

   2.3.3 DATA, PHYSICAL: A datum or data written on a physical data form.

   2.3.4 DATA PROCESSING BATCH (DP BATCH): A number or alphanumeric code assigned to either a collection of household packets, or to a collection of one or more site or lab sheets within a site/lab packet.

   2.3.5 DATA RECORD: In the context of this SOP, this is a row of electronic data in a database.

   2.3.6 DATA, VERIFIED: Electronic data re-entered into the same table and database into which it was originally entered, and programatically compared against the original entered data. Verified data are the product of "data verification."

2.4 DATABASE, EMPTY: An R:BASE database structure that contains no data records except for parameter rows; or a non-R:BASE database structure that contains no data records.

2.5   FIELD: An area on a data entry form (i.e., screen) where a datum from a physical form is entered (see FORM, DATA ENTRY below); or an area on a physical form from which a physical datum is entered.

2.6   FIRST STUDENT DATA INPUT ASSISTANT: A Student Data Input Assistant who codes a stack, but who cannot review the coding of the same stack; or, a Student Data Input Assistant who does data entry of a stack, but who cannot do data verification of the same stack.

2.7   FORM, DATA ENTRY = A computer screen representation of a physical data form. The data on the physical form is entered into the computer database via the data entry form.

2.8   FORM, PHYSICAL [DATA] = The paper or "hard copy" version of a data form. This is also referred to as a "physical data form."

2.9   KEYPUNCH: The primary area in which data entry and data verification of NHEXAS Arizona field data takes place. It is located in the Respiratory Sciences Center, Room 2329; Arizona Health Sciences Center (AHSC); 1501 N. Campbell Avenue; University of Arizona; Tucson, AZ 85724.

2.10   LAN: Abbreviation for Local Area Network. This is any physical network technology that operates at high speed over short distances.

2.11   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.

2.12   OWNERSHIP CARD: A large (6" x 4" or larger) index card placed atop a stack designating who is responsible for the stack, as well as the data processing status of the stack. The card contains a keypunch staff member's name followed by "coding," "reviewing," "data entry," or "data verification." For example, Pat Smith's ownership cards would read "Pat's Coding," "Pat's Reviewing," "Pat's Data Entry," and "Pat's Data Verification."

2.13   PACKET: A sturdy, envelope-like container that can be fully closed and is large enough to hold the physical data form(s) generated by a study household, laboratory, research site, or data processing batch. One type of packet is used for one type of physical data forms (eg., manila envelopes would be used for all lab forms processed at the HRP site). Packets are either color coded, labeled according to their contents, or both. What are referred to as "household packets" are relevant to this SOP (see PACKET, HOUSEHOLD below).

2.14   PACKET, HOUSEHOLD: A packet containing the physical data forms for a study household.

2.15   PARAMETER ROWS: The five data records in an R:BASE table that define the following for each variable contained therein: (1) the upper range limit, (2) the lower range limit, (3) the global code for refusal, (4) the global code for non-applicable, and (5) the global code for missing.

2.16 PC: Abbreviation for Personal Computer. This is a microcomputer based on the Intel 8088/8086 instruction set and connected to SCORAZ through the LAN. PCs are nodes on the LAN.

2.17 SECOND STUDENT DATA INPUT ASSISTANT: A Student Data Input Assistant who reviews the coding of a stack, but who did not do the original coding of the stack; or, a Student Data Input Assistant who does data verification of a stack, but who did not do the original data entry of the stack.

2.18 SCORAZ: The SparcStation at the AHSC site that houses NHEXAS Arizona databases having one or more of the following statuses: (1) empty databases ready for data entry, (2) databases containing entered data, and (3) databases containing verified data. The status of any given database depends on the status of the current data processing batch.

2.19 SPARCSTATION: A small computer developed by Sun Microsystems and based on a RISC (Reduced Instruction Set Computer) Processor. The Sun SparcStation is a fast, multi-tasking, multi-user platform that runs the SunOS version of the Unix operating system. The node name of the Sun SparcStation pertinent to Keypunch is Scoraz.

2.20 STACK: A pile of physical forms representing only *one* type of form obtained from the packets of *one* data processing batch. A stack is sorted in ascending, numerical order by key variable(s), is secured by a large rubber band, and is identified by an ownership card at all times when not being processed.

3.0 References

None

4.0 Discussion

First, an intersubjective approach to data entry is employed. This helps to control any bias in data entry that would be produced if one person were to both enter and verify the same stack.

Second, stacks are entered and verified on PCs connected to the Respiratory Sciences Center LAN in which SparcStation Scoraz is a node. A tape backup of all files on Scoraz is made daily. This is a security measure to rely upon in case of LAN system failure.

5.0 Responsibilities

5.1 The Project Data Coordinator is responsible for
(a) installing, maintaining, and updating menu-driven program(s) for data entry and verification on Scoraz;
(b) providing written instructions for use of the menu-driven data entry and verification program;
(c) installing, maintaining, and updating empty databases on Scoraz;
(d) retrieving verified, electronic data from Scoraz upon notification from the Data Input Operator Supervisor;

(e)   retrieving household packets from Keypunch; and

(f)   answering any questions about data entry and data verification of NHEXAS Arizona data.

5.2   The Data Input Operator Supervisor is responsible for

(a)   supervising Student Data Input Operators in data entry and data verification;

(b)   notifying the Project Data Coordinator regarding problem(s) with the menu-driven program for data entry and data verification;

(c)   ensuring that this SOP is followed; and

(d)   ensuring that the coding and entry log form is filled out properly after each stack is entered and verified.

The Data Input Operator Supervisor may delegate any of her or his responsibilities.

5.3   The Student Data Input Assistant is responsible to the Data Input Operator Supervisor, and for

(a)   data entry of stacks;

(b)   data verification of stacks; and

(c)   keeping custody of the stacks during data entry and data verification.

6.0   Materials and Reagents

6.1   Materials

6.1.1 Stack(s)

6.1.2 "Coding, Entry, and Verification Record" form (referred to as coding and entry log form in this SOP) [Figure 1]

6.1.3 Ownership cards labeled appropriately

6.1.4 PC connected to the Respiratory Sciences Center LAN in which SparcStation Scoraz is a node

6.1.5 Instructions for using menu-driven data entry and data verification program

6.1.6 "Keypunch Problem Sheet" [Figure 2]

6.1.7 "COMPLETED" stamp

6.2   Reagents

None

7.0   Procedure

7.1   The Data Input Operator Supervisor assigns each stack to a *first* Student Data Input Assistant for data entry. Ownership cards are attached to the stacks appropriately.

7.2   The Student Data Input Assistant does data entry of the stack(s) into a data entry form.

7.2.1 The Student Data Input Assistant writes any data entry problems on the "Keypunch Problem Sheet," and also attaches post-its if necessary to clarify the location of the problem(s).

7.2.2 Upon completion, the Student Data Input Assistant submits the en-

tered stack(s) to the Data Input Operator Supervisor.

7.3   The Data Input Operator Supervisor assigns each entered stack to a *second* Student Data Input Assistant for data verification. Ownership cards are attached to the stacks appropriately.

7.4   The Student Data Input Assistant does data verification of the stack(s) into a data entry form.

   7.4.1 If the Student Data Input Assistant enters a value into a field that is different from the one originally entered, then the data verification program beeps, provides an error message containing the original entered value, and prompts for re-entry of the value. The actual value is entered and data verification proceeds as usual.

   7.4.2 After a physical form is verified, it is stamped with the word "COMPLETED" in the top margin.

   7.4.3 The Student Data Input Assistant writes any data verification problems on the "Keypunch Problem Sheet," and also attaches post-its if necessary to clarify the location of the problem(s).

   7.4.4 Upon completion, the Student Data Input Assistant submits the entered stack(s) to the Data Input Operator Supervisor.

7.5   The "Keypunch Problem Sheet" is attached to the outside of the batch of household packets once they have been re-assembled (see SOP# UA-C-6.0), and before they return to the Project Data Coordinator.

7.6   Quality Control

   7.6.1 Tolerance Limits

      Each physical form in a stack is entered into a database. Then, a second person does a 100% re-entry of the stack. This quality control measure has been shown to produce a 99% accuracy rate in data entry.

   7.6.2 Detection Limits

      All errors in data entry are detectable via data verification.

   7.6.3 Corrective Actions

      Any data entry errors encountered during data verification will be corrected.

8.0   Records

8.1   Data to Be Recorded from This Procedure

   None

8.2   Record Forms (Attached)

   8.2.1 Figure 1: "Coding, Entry, and Verification Record" form
   8.2.2 Figure 2: "Keypunch Problem Sheet"

8.3   Location of Record Forms

8.3.1 The batch tracking forms are returned to the Project Data Coordinator after the batch is cleaned. They are then stored in the "Data Processing Batch Sheets" notebook in room 128c of the HRP site.

8.3.2 The "Coding, Entry, and Verification Record" form is retained at Keypunch and stored in the "Coding and Entry Log" notebook.

8.3.3 All coding protocols are stored in the "Coding: NHEXAS Phase" notebook.

Figure 1: "Coding, Entry, and Verification Record" form

CODING, ENTRY, AND VERIFICATION RECORD

FORM NAME: _____

Form ID: UA-C-6.0-2.0

DP BATCH	CODING			Method	DATA ENTRY		COMMENTS
	Init	Date			Init	Date	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	
	Cod: ____ Rev: ____	__/__/__ __/__/__		[ ]scan [ ]keyp	Ent: ____ Ver: ____	__/__/__ __/__/__	

Figure 2:   "Keypunch Problem Sheet"

KEYPUNCH PROBLEM SHEET
NHEXAS Arizona

Stage:_____   Date: ___/___/__
Form Name:_____   DP Batch: ___/___/__
Page: ___ of ___

HHID	KEY VAR: ___	KEY VAR: ___	KEY VAR: ___	QUESTION #	PROBLEM

—For—Office—Use—Only—
Problems resolved by: _____   Date: ___/___/___   Form ID: UA-D-15.0-1.0
Comments:_____