

Standard Operating Procedures

Subject: Error Correction

Date: 15Jun2006

Purpose: Define procedures for correcting errors and making other changes or additions to data files for the Data Management and Analysis Center (DMAC).

Scope: These standards apply to computerized data files managed by the Data Management and Analysis Center.

Responsibility: Making changes to a project's data files for the purpose of correcting errors, filling in missing values, or adding new fields is the responsibility of the programmer assigned to that project. It is also the programmer's responsibility to document and validate these changes.

Making a change to a value in a computer data file may include five steps: identification, authorization, implementation, documentation, and verification. The minimum requirements for each step will be detailed below:

Identification

The need for a change in the data may be identified in a number of ways. An error in data entry may be detected through the verification process, checking values in the data against the original instrument. Errors may also be identified using an error detection program (see error detection SOP). In cases where the instrument is vague or answers require interpretation, there may be decisions that will amend values that have previously been entered. Sometimes entire observations must be deleted in cases of duplicate or erroneous entry.

In every case it is necessary to identify the unique identifying values for the observation(s) to be changed, the old (incorrect) value, and the new (corrected) value.

Authorization

Any change to a project's data should be approved by a responsible project staff member. The authorization should be recorded in writing so that it may be filed in the project's documentation. This may take the form of an error request sheet, an email, a memo, or simply corrections indicated on a SAS printout.

Implementation

Changes to a data file must be made by creating a new version of the dataset. In this way there will exist versions of the data before and after the changes were made. Changes must be easily reversible.

Documentation

All changes to a data file must be permanently documented so that the record of all data values can be traced from the time of entry. Documentation should include the authorized request for the change (discussed above), a log of the program making the change, and a comparison of the data before and after the change showing that it was made as requested.

Verification

A record of each change may be sent to the project staff who authorized it for their final review and approval. Programming staff are responsible for checking that the correct changes have been made and no unexpected changes have been made, irregardless of whether or not the project staff choose to also verify these changes.