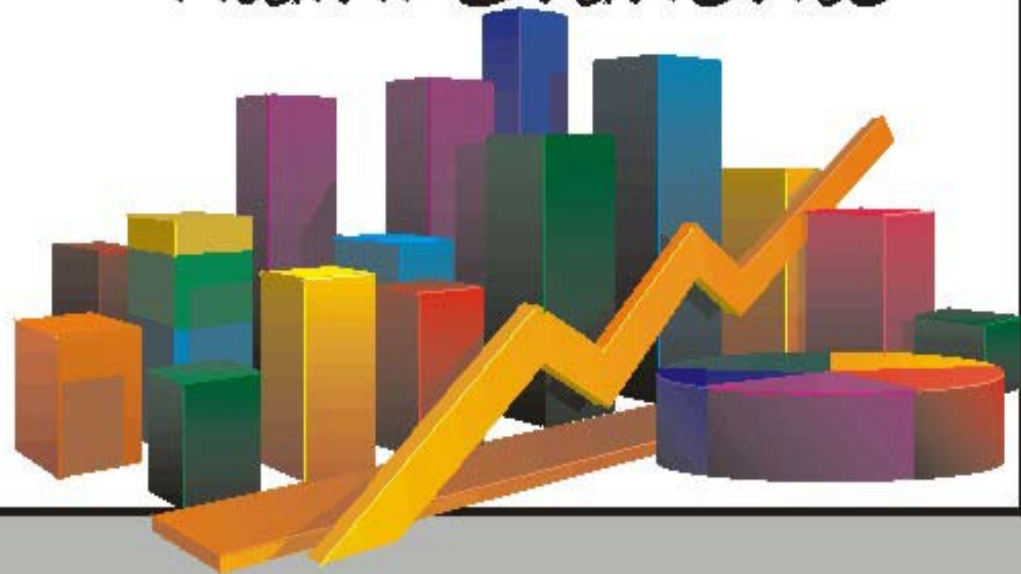


# *National Center for Health Statistics*



*National Center for  
Health Statistics*



# Source Point Data Editing in Health Surveys (Topic IV) Supporting Paper

*Kenneth W. Harris, National Center for Health Statistics (USA)*

*October 20-22, 2003*

The National Center for Health Statistics (NCHS) is the Federal agency responsible for the collection and dissemination of the nation's vital and health statistics. To carry out its mission, NCHS conducts a wide range of annual, periodic, and longitudinal sample surveys and administers the national vital statistics registration systems. These sample surveys and registration systems form four families of data systems: vital event registration systems (4), population based surveys (5), provider based surveys (5), and followup/followback surveys (1).



Much of what happens to the data covered by these data systems, from collection through publication, depends on the family to which they belong. At most steps along the way various activities and operations are implemented with the goal of making the data as accurate as possible. These activities and operations are generally categorized under the rubric, “data editing.”



In the 1990 Statistical Policy Working Paper 18, “Data Editing in Federal Statistical Agencies,” data editing is defined as:

Procedures designed and used for detecting erroneous and/or questionable survey data (survey response data or identification type data) with the goal of correcting (manually and/or via electronic means) as much of the erroneous data (not necessarily all of the questioned data) as possible, usually prior to data imputation and summary procedures.



Source point data editing refers to editing survey data by any means of access to either the interviewer (or other data collector), the respondent, or records within a limited time following the original interview or data collection. This time limit reflects the period within which the persons involved can reasonably be expected to remember details of the specific interview or, in the case of data collected from records, a time within which there is reasonable expectation that there has been no change to the records which would affect the data collected.



Most of the NCHS data systems implement the majority of their data editing procedures after data entry. In fact, NCHS's population based surveys are the only ones that, as a group, have instituted significant data editing procedures at or near the point of data collection. Just one provider based survey, the National Nursing Home Survey, and the lone Followup/Followback Survey, the Longitudinal Study of Aging II, also use source point data editing. These data systems use automated procedures during data collection which identify problems that offer the opportunity to recontact the data suppliers in a timely fashion. The four registration systems also use a limited number of automated procedures after data collection. A total of eighteen specific automated software checks were identified as being used by the NCHS data systems.



## Frequency of Selected Source Point Data Editing Among 11 NCHS Data Systems

*Automated Skips (7)	Re-entering and deleting members from a roster (8)	*Allowing other comments and notes (6)
*Backing up and changing data during the interview (7)	Context sensitive help (7)	Out of range values (11)
*Automatic adjusting of skip patterns for changed entries (7)	Forcing correction before proceeding if a data entry is invalid (11)	Inconsistency (11)
*Ability to access and change data file after interview is over but during SPDE timeframe (6)	Warning of unusual data entries (9)	Illegal blanks (4)
Re-entering and correcting information about existing members in a roster (11)	Forced confirmation before storing data (2)	Invalid dates (11)
Re-entering and adding members to a roster (9)	*Allowing the interviewer to enter comments on each question (4)	Invalid characters (e.g. alpha characters in a field that requires a numeric answer) (11)

**\*Not Applicable for Registration Systems**

# Data Collection Followup

**Are respondents ever recontacted after data collection is over? (10)**

**\*Does your survey recontact respondents to check for data falsified by the interviewer? (5)**



# Conclusion

Source point data editing is a valuable tool for improving the completeness and quality of data collected in NCHS health surveys. New procedures are regularly being tested and evaluated. Because of certain data collection constraints with regard to some provider based surveys, it is likely that NCHS will never have 100 percent coverage. Nevertheless, the number of surveys employing source point data editing is expected to increase over time.

# **Kenneth W. Harris**

## **Contact Information**

**Phone: 301-458-4262**

**Email: [KWH1@CDC.GOV](mailto:KWH1@CDC.GOV)**

**Address: CDC-NCHS**

**3311 Toledo RD**

**HYAT IV, RM 3210**

**Hyattsville, MD 20783, USA**

