

**Using CAPI for the Medicare Health Status Registry Reinterview Surveys: Design and Operation**  
**Judith T. Lessler, Battelle; Frank Mierzwa and Tim Smith, Research Triangle Institute**  
**Judith T. Lessler, Battelle, 100 Capitola Dr., Durham, NC 27713**

**Goals of Reinterview Study**

The Research Triangle Institute under contract from the Health Care Financing Administration (HCFA) is currently conducting a field experiment directed at identifying methods for carrying out a Medicare Beneficiary Health Status Registry Survey. If fully implemented, the Medicare Beneficiary Health Status Registry would be a longitudinal database containing survey data on characteristics of Medicare beneficiaries and claims data on use of Medicare services. The field experiment is investigating the cost and quality of the responses obtained for different versions of the questionnaire and different ways of soliciting response. A total sample of some 2400 Medicare beneficiaries are included in the study. Data are being collected using a combination of mail, telephone, and face-to-face interviewing. A subsample of 600 beneficiaries are participating in a reinterview study.

The reinterview study has several goals. One is to examine the validity of using shorter scales to measure characteristics that are measured with longer psychometric scales in other settings. A second goal is to examine the reliability of the responses to the original survey; and the third goal is to examine respondent's understanding of and their ability to accurately report certain medical conditions and procedures. Thus, the reinterview study aims to investigate both reliability and validity of items included in the survey.

**Issues Related to the Design of the Reinterview Study**

In a review of past research, Forsman and Schreiner (1991) noted that the reinterview method had been used to meet three different objectives: (1) to obtain the correct response, (2) to assess reliability, and (3) to assess interviewer performance. Among the factors that they noted as important to the design of the study were the design of the reinterview questions and the associated method of conducting the reinterview and

identifying discrepancies. To measure reliability, it is necessary to use the same questions possibly reworded to reflect changes in the time period since the original interview.

Two approaches have been used to measure examine validity or bias. Under one approach, researchers trying to obtain a more accurate answer reword questions or use a set of more focused questions to replace the original question. As noted above, this approach is being used in the field experiment for a subset of the questions in the original interview. A second approach for measuring bias is to repeat the original question and reconcile differences between the two responses through the use of either direct or indirect probing.

Forsman and Schreiner (1991) describe a three step process that uses the direct probing approach. The survey is re administered to the respondent; the reinterview response are compared to the original responses, and interviewer directly queries the respondent about the discrepant responses and asks the respondent to provide a "correct" response. Bergman, Kristiansson, and Safstrom (1991) report on a modified technique that is used in the Swedish Labor Force survey in which the interviewer asks a series of questions about discrepant items without alerting the respondent to the fact of the discrepancy. A survey expert then determines the correct responses.

Each method of reconciliation has certain advantages. The direct probing method allows the respondent to provide the correct answer, and the respondent is probably in the person best qualified to do this. The disadvantage is that respondents may be embarrassed by the direct approach and unwilling to cooperate with the interviewer in determining the correct response. The primary advantage of the indirect method is that it minimizes the embarrassment of the respondent thereby decreasing the risk that he or she may cover up an error. The direct probing method was adopted for the Medicare Registry field experiment.

Biemer and Forsman (1990) investigated problems with reinterview surveys. In comparing

the results of reinterview surveys that used reconciliation with those that merely reasked questions without subsequent reconciliation, they noted that research has shown that fewer discrepancies are found in a reinterview survey slated for reconciliation. Although the interviewers are not supposed to change the reinterview responses during the reconciliation, it appears that this does happen. It is easy to see how this might occur. In the original survey the respondent may report that she is 52 years old. She may also do so in the reinterview, but the interviewer may write down 25 in error. This is a discrepancy and should be included in any statistics that measure reliability because simple transcription errors do contribute to response variance. However, it is also easy to see how the interviewer would be likely to correct this transcription error rather than move it forward into the reconciliation process. The decision by RTI and HCFA to use CAPI for the Medicare Registry reinterview was partially based on the need to control this type of error and to get valid measure of simple response variance.

### **Design of the Reinterview Study.**

#### **Sample Design.**

Since the reinterview study entails making personal visits to the beneficiaries residence, the reinterview subsample was clustered in three geographic areas near RTI. A sample of 600 beneficiaries was selected--200 in each geographic area. Half the sample are new aged enrollees--mostly 65 years old; half are between 76 and 80 years old. The enrollee sample is stratified by race, sex, and, for the older cohort, by health status. The sample beneficiaries were randomly assigned to receive either a short, medium, or long questionnaire and either an introductory mailing or no introductory mailing. In addition, some small scale experiments are built into the questionnaires.

#### **Content and data collection procedures**

Questionnaires were mailed to all 2,510 sample members in May, 1993. A second questionnaire was sent to all nonrespondents in early July, approximately six weeks after the first mailing. A third request is scheduled to be sent to remaining nonrespondents in early August. Attempts to interview all remaining nonrespondents will made

by telephone 4 weeks after the third request by mail. Tracing and refusal conversion efforts are being made by telephone as necessary. Within the reinterview sample, field interviewers will attempt to interview all nonrespondents who do not return a questionnaire or complete an interview by telephone.

Data collected from the Time 1 questionnaires provide the basis for the reinterview. Once a questionnaire is received (or an interview is conducted with an interviewer), the data are edited and keyed. Special consistency codes are used in the editing process so that the original quality of the respondents' answers remains evident. For example, if a respondent circles two codes on an item that requires only one response, the editor uses a consistency code to indicate the type of error made. No effort is made by the editor to resolve the error. This editing process permits the field interviewer to ask specific questions during the reinterview that examine the cause of the response error.

Time 1 data are telecommunicated to a laptop PC used by a field interviewer. In addition to the telecommunicated data, the field interviewer receives an Assignment Control Form and copies of any pages from the original questionnaire containing "bad data" codes. These additional materials are sent by mail. Once the field interviewer receives the data and mailed materials, he or she attempts to contact the Time 1 respondent by telephone to schedule an appointment for the face-to-face reinterview.

#### **Time 2 Reinterview**

- (1) The 6-item Short Blessed Mental Status Questionnaire
- (2) Repetition of the Time 1 questions
- (3) A set of observational items for the interviewer to answer
- (4) A set of questions that assess the respondent's ability to hear
- (5) The Rand-38, a 38-item mental health inventory (self-administered paper and pencil form)
- (6) Probing of Consistent Items: For some questions, when the Time 1

and Time 2 answers are consistent, additional items are asked to verify the respondent's understanding the question and to collect more detailed information on the procedure or condition.

- (7) Probing of Inconsistent Items: It there is either a discrepancy between the Time 1 and Time 2 answers or "bad data" from either time, questions are asked to determine the correct response and the reason for the discrepancy.

The Short Blessed Mental Status Questionnaire is used as a screener to identify respondents who exhibit significant cognitive impairment. The quality of data provided by respondents who fail this examination is questionable. The scores from scale may allow identification of possible impairment found in the self-administered data.

The reinterview includes all of the questions found in the original questionnaire. Topics include general health status, receipt of medical care, life-time and current diagnoses of physical conditions, functional status, mental health status, cigarette and alcohol use, and demographic characteristics.

During the reinterview additional items pertaining to functional status are asked of some respondents. Each of the short, medium, or long version of the questionnaire contains some items to assess functional status. However, the full set is included only in the long version. To validate the use of the subset of these items used in the short and medium versions, all Time 2 reinterviews include the full set.

There are 19 items that can be probed when the respondent provides consistent answers at Time 1 and at Time 2. The probes are designed to obtain additional information about the condition or procedure that has been reported and verify the respondent's understanding of the terms used in the original question. For example, if a respondent reports having been diagnosed with diabetes at both interviews, the interviewer would ask whether or not the respondent was told to "do anything for your diabetes". If the answer is "yes", the interviewer

then asks for a description of what the respondent was told to do.

The final module in the reinterview consists of probes of any answers for which either discrepant responses or "bad data" responses have been provided. Twelve types of inconsistency are probed, based on pre-coded responses to the questions, "Don't Know", refusal, blank, and other "bad data" responses. The probes ask why answers were not provided, which of the two answers is correct, whether changed circumstances between Time 1 and Time 2 have resulted in different "correct" answers, and for other types of clarification.

After the interviewer completes the reinterview, Time 2 data are telecommunicated to RTI. In addition to the telecommunicated data, the field interviewer sends the completed Assignment Control Form and any other materials related to the case.

#### **CAPI design**

The reinterview is conducted by Computer Assisted Personal Interview (CAPI) using CASES (Computer-Assisted Survey Executions System) software. In addition, an external call-out to FoxPro from CASES is being used for a small portion of the reinterview. Interviewers use a Compaq Contura 20 Mhz microcomputer with an internal modem, 40 Mb hard disk drive, 2 Mb of RAM, and a 1.44 Mb floppy disk drive. The reinterview application utilizes approximately 20 Mb of disk space. The following components are included in the application:

- Time 1 data files
- Time 2 reinterview questions
- Time 2 data files
- Code for Time 1/Time 2 data comparisons
- Questions for consistent and inconsistent probing

Within the application, a number of function keys have been programmed to assist the field interviewer. Keys have been programmed to allow the field interviewer to backup, change an answer, go forward (after changing a response), to attach notes/comments mentioned by the respondent, to

code the reason why Question-by-Question (QxQ) information was used (if the respondent asks for clarification of a question or the response codes offered), and to break-off an interview and save the data captured (if the respondent is unable to complete the interview in one setting).

In addition to the reinterview application, a field tracking and telecommunications system, developed using Clipper software, has been loaded onto the laptop. The system allows field interviewers to enter event codes daily for results of reinterview attempts and to transmit data and event code information to and from RTI. Interviewers received extensive instruction on the use of this system and the process of "calling-in" to RTI. The field interviewers call-in to RTI daily to transmit data from completed cases and updated event code information. During the call, Time 1 data from newly assigned cases is also downloaded onto the field interviewer's laptop machine.

The field tracking and telecommunications system is completely menu driven, allowing for ease of use by the interviewer. To update event code information for a case, the interviewers need only to select "Track" from the menu, toggle down to the Case ID to update, and the system prompts the interviewer to enter the new event code. To complete the transmission process, interviewers need only to attach the telephone cord to the laptop machine and a wall outlet, and select "Send" from the menu. The system then dials into RTI, sets up the updated laptop files for transmission and uploads them to RTI, and downloads newly assigned Time 1 data files to the laptop system.

The reinterview laptop system also contains an external menu shell developed to reduce the interviewer's burden. Once the laptop machine is turned on, it boots to a menu screen where the interviewer can either choose tracking/telecommunications or CAPI reinterview. If an interviewer selects CAPI reinterview, he or she is prompted to enter the case ID. The interviewer is then prompted to re-enter the case ID to confirm that the ID was properly entered properly, and the system selects the appropriate

interview using a proxy or beneficiary gender-specific short, medium, or long version of the questionnaire. The CAPI application references to first two-digits of the case ID and a flag set in the Time 1 data for the case to determine which version of the questionnaire to provide.

During the reinterview, the consistency checks between reinterview and original responses are conducted on approximately 160 variables. The system to implement these checks is quite memory intensive and can take as long as 90 seconds to create the set of probes. The generation of the module to probe consistent responses is based on consistent responses for a subset of 19 items. Based on consistent responses provided by the respondent to these items, as many as 25 probes can appear in this module.

If the respondent does not answer a question or provides inconsistent responses, a module for probing inconsistent responses is generated. When a missing data value is coded for the original or reinterview, or the non-missing responses are inconsistent, a probe appears on the interviewer's laptop screen. The screen display for each of these probes is split in half, with the top portion containing the original question asked and a listing of the two responses obtained. The bottom portion contains the probe to ask the respondent. The interviewer studies each screen, and fills information concerning the discrepant item and responses displayed in the top half into the probe displayed in the bottom half. The interviewer then types into the laptop system the verbatim response offered by the respondent. A "code screen" appears after each verbatim response is entered. The interviewer then enters the appropriate code(s) based on the respondent's answer.

### **Pretesting and Revision of Methodology**

The process of reconciliation used in the reinterview involves asking additional probe questions which either directly or indirectly determine the correct answer. Additionally, these probe questions also aid in determining the source of any discrepancies. The comparison of Time 1 and Time 2 data during the reinterview requires somewhat opposing strategies.

An initial set of inconsistency probes were developed through laboratory testing in December, 1991 to capture the respondent's reasons for answering a question differently, or being unable to provide a response during the Time 1 or Time 2 interview. Interviewers were to ask the probe and type the respondent's verbatim response into the laptop computer. For each type of probe, a "code set" was developed for the interviewer to use to code the verbatim response.

In December, 1992 a field pretest of the CAPI system was implemented with 8 volunteer respondents. A short, medium, or long form questionnaire was mailed to each of the respondents to complete. The questionnaires were returned to RTI, edited and keyed, and the data were loaded into the laptop computer. A reinterview followed using CAPI in an effort to test the system and the set of probes for discrepant responses. Based on reinterviews of the 8 pretest respondents, we concluded that some of the probes needed revision, and that the set of response codes did not sufficiently cover the reasons offered by the respondents for providing discrepant responses.

Using the information gained during the pretest, an improved set of probes and response codes was developed in January, 1993. The set consisted of 11 probes and 17 response codes.

## Consistency Codes

### 10. Understanding Question

(11) Difficulty understanding question

### 20. Knowledge, recall or sensitivity

(21) Insufficient knowledge or information to answer (at T1, at T2, or both)

(22) Respondent received additional information about own health that affected T2 response

(23) Question sensitivity affected response (at T1, at T2, or both)

### 30. Response categories

(31) More than one response alternative seemed correct

(32) No response alternative seemed correct

(33) More general difficulty understanding or selecting a response

### 40. Recording response

(41) Failed to follow skip instructions at T1

(42) Failed to follow directions for recording response

### 50. Assistance received

(51) Assistance by someone other than interviewer at T1

(52) QxQ assistance from interviewer at T1, at T2 or both (53) QxQ assistance at Reconciliation

### 60. Respondent specific reasons

(61) Stopped answering questions at T1 because of burden

(62) Fatigue or physical limitation

(63) Refusal at T1, at T2 or both

(64) Changed mind

### 70. Other

(71) Respondent believes there is an error in our records for T1 response, for T2 response or for both responses

Given the list of 17 response codes, we turned our attention to how best to present them on a CAPI system screen so that the field interviewer could accurately and effectively use them. We grouped the 17 reasons into seven general categories:

- (1) Understanding Question
- (2) Knowledge, recall or sensitivity
- (3) Response categories
- (4) Recording response
- (5) Assistance received
- (6) Respondent specific reasons
- (7) Other

We developed a two-column screen layout containing each main category heading, followed by the set of specific reasons which fit within the general category. The field interviewers received extensive instruction on how to administer the inconsistency probes, and how to accurately use the response code set to code the provided verbatim response.

## Preliminary Results from the Pilot Study

While analysis of reinterview data is not scheduled to begin until October, 1993, there are some

experiences that can be reported at this time. When the procedures for the reinterview were first developed, there were four general concerns regarding the ability to collect the reinterview data. Specifically, these concerns centered on the

- Interviewers' ability to use the laptop system successfully,
- The possibility to complete the reinterview within 3 weeks after the Time 1 data were received,
- Respondents' ability and willingness to participate in a second interview, and
- Respondents' ability to resolve discrepancies.

The interviewers received intensive training on the laptop system. Only a few of the interviewers had prior experience using laptop PCs, and none had ever used a field tracking program telecommunication system. During the first few weeks of field operations, a few interviewers experienced some problems using the reinterview software; these difficulties were resolved through contact with their supervisor. None of the interviewers has reported any difficulty with the telecommunication system.

As the field procedures were being developed, a question arose as to when the reinterview should take place. Conducting the reinterview too soon after the original questionnaire was completed could result in answers based on the respondents' memories of their previous answers rather than on their "true" score. Conducting the reinterview too long after the first interview could result in discrepancies due to the passage of time rather than on a problem with the question. A goal was set to conduct the reinterview between 2 and 3 weeks after the first questionnaire was received at RTI. Initially, the completion time did not meet this goal, largely due to staff turnover during the first week of field data collection. This has been resolved and interviewers are now completing cases in a more timely fashion.

Overall, respondents seem quite willing and able to participate in the reinterview. The time required to conduct the reinterview (about 1 hour) and the

respondents' ability to provide answers about discrepancies seem not to have hindered the interviewers' efforts. Interviewers have reported some cases where the respondent would not let them through the door, but the interviewers were able to conduct the reinterview on the doorstep. Very few people have refused to participate in the reinterview all together.

It is not possible, at this time, to relate the number or types of errors respondents are making. Nor are data available on the respondents' explanations for the discrepancies since the data are just now beginning to arrive at RTI. However, it does seem that the reinterview procedures that are being implemented are working well from both the interviewer and respondent perspectives. Once all of the data have been assembled, they promise to provide a rich resource for analysis.

## REFERENCES

- Bergman, L., Kristiansson, K., and Safstrom, M. (1991). Decentralized CATI Versus Paper-and Pencil Interviewing: Effects on the Results in the Swedish Labor Force Surveys. Stockholm: STATISTICS SWEDEN
- Biemer, P. and Forsman, G. (1990) On the Quality of Reinterview Data with Application to the Current Population Survey. Paper presented at the Sixth Annual Research Conference, U.S. bureau of the Census, Washington D.C.
- Forsman, G. and Schreiner, I (1991) Design and Analysis of Reinterview: An Overview. In Biemer, P., et. al (1991) Measurement Errors in Surveys, New York: John Wiley & Sons.