

THE LONGITUDINAL STUDY OF AGING MEDICARE RECORD MATCH

Michele M. Chyba, National Center for Health Statistics
6525 Belcrest Rd. Hyattsville, MD 20782

KEY WORDS: Aging, Medicare, Match

Introduction:

Abstracting information from administrative records is an appealing alternative data collection method for survey researchers. It is appealing because it enables researchers to obtain information not collected in the survey instrument and to verify information reported by the respondent. Matching survey data to administrative records is not, however, without dilemmas.

One dilemma confronting the researcher is how to define a match. Should exact matches and probable matches be treated equally? The intent of this paper is to describe the procedures used when matching data from the Longitudinal Study of Aging to Medicare claims records (Part A and Part B).

Survey Background:

The Longitudinal Study of Aging (LSOA) is a cooperative venture of the National Center for Health Statistics (NCHS) and the National Institute of Aging. Telephone (CATI) interviews are conducted by the Bureau of the Census at their facility in Hagerstown, Maryland.

The outcome of this collaborative effort is a public use data tape which researchers can use to measure changes such as changes in living arrangements and physical limitations experienced by a cohort of men and women who were 70 years old and over in 1984.

The baseline information for the LSOA comes from the 1984 National Health Interview Survey (NHIS), the 1984 Health Insurance Supplement, and the 1984 Supplement on Aging (SOA). Reinterviews have been conducted in 1986 and in 1988. The 1990 reinterview is currently underway. The persons reinterviewed in 1986 were a sample of the 7,541 persons who were 70 years and older at the time of the 1984 SOA. Everyone who was 70 years old and over in the 1984 SOA was included in the 1988 reinterview sample. The 1990 reinterview sample included all individuals interviewed in 1988. Persons known to be deceased or those who asked not to be reinterviewed were excluded from each reinterview sample.

Each reinterview collected information about living arrangements, physical limitations, functional

limitations, and use of medical care (nursing home stays, hospital stays and contacts with doctors).

In addition to the information collected during the reinterviews, data for this cohort have been obtained from the National Death Index maintained by NCHS and the Medicare Automated Data Retrieval System (MADRS) maintained by the Health Care Financing Administration (HCFA).

Survey Data and Medicare Claim Matching:

The LSOA did not obtain detailed information such as diagnosis, length of stay, or charge about each hospitalization that was reported. The survey instrument asked only for the number of hospitalizations in the past 12 months. Detailed hospitalization information was obtained, instead, through accessing administrative records, namely, the MADRS Medicare claims records. At the time of the interview, the respondent or sample person was informed that providing their social security number or HIC number was voluntary. Records for those respondents who gave permission were matched to the MADRS.

The LSOA survey data have been matched to the MADRS twice. Each match was conducted after the reinterview field work was completed in 1986 and in 1988. The match will be conducted again following the completion of the 1990 reinterview. Version 2 of the LSOA Public Use Data Tape included Medicare data for 1984 through 1987. Version 3 of the public use data tape will include Medicare data for 1984 through 1989.

Including Medicare data on the LSOA Public Use Data Tape required three steps: 1. submitting a tape of social security numbers and health insurance claim (HIC) numbers to HCFA to be matched to their Master Enrollment File, 2. submitting a tape of social security numbers and HIC numbers that matched the Master Enrollment File to HCFA for matching to the MADRS file, and 3. matching the MADRS records to the survey respondent.

The LSOA has incorporated final bill records from the MADRS files for the following types of Medicare claims: inpatient, outpatient, home health, and hospice.

Record matching is not as easy as it sounds; it is

not a quick one, two, three process and it is done. The entire process cannot always be accomplished by using a computer program; some of the matching is completed by looking at the data and matching the records by hand. Often, regardless of the criteria that define a match, there are discrepancies between the survey data and the record data that need to be resolved. Discrepancies can occur because of a keying error or the legibility of the recorded data.

The steps followed to complete the matching process are illustrated in Figure 1. The first box on the top left in Figure 1 represents the social security numbers (SSN), the railroad retirement board (RRB) numbers and the health insurance claim (HIC) numbers obtained during the interviews.

Before the survey data were matched to the Medicare files, the social security numbers and health insurance claim numbers obtained in the surveys were compared to the Master Enrollment File.

The Master Enrollment File includes a number for every person eligible for Medicare regardless of whether or not he or she has filed a claim. The file includes deceased persons as well.

To accomplish the match to the Master Enrollment File, a file called a "Finder File" was prepared at NCHS according to HCFA specifications. This file contained SSN, RRB, and HIC numbers. To protect the anonymity of the sample person, the file never contained any other identifying information. To insure confidentiality, the LSOA files were merged with similar information from two other surveys conducted by NCHS. HCFA staff did not know to which survey the numbers belonged.

As illustrated in the second box (top right) in the figure, the "Finder File" was then linked by HCFA to the Master Enrollment File. The file NCHS received in return included the name, address, date of birth, date of death, sex, and race for each survey number included on the file that matched a record on the eligibility file.

The third box (second row, right) represents the corrections and comparisons of the results made at NCHS. Based on the results of comparing the HCFA and LSOA files, HIC numbers, addresses and/or name spellings were changed.

The survey and Master Enrollment File records were compared based on the name, address, sex, race, date of birth, date of death, SSN, and HIC number. If the survey and Master Enrollment File records matched on nine digits, name, and sex and the survey number was a social security number, it was replaced with the HIC number from the Master Enrollment

File. Records that matched on all nine digits, but not the first and last name were printed and examined. This procedure identified differences in name spellings as well as instances where the names were completely different. The differences could have occurred for one of several reasons: 1) the respondent had changed his or her name, 2) there was a keying error, or 3) the survey respondent could have reported the SSN, RRB, or HIC number incorrectly. Addresses were more likely to change during the 1988 matching process than during the previous matching process; about 2,300 persons had not been contacted and reinterviewed since the 1984 SOA. The address on the HCFA file could have been a household address, a mailing address, or a billing address. If the address on the HCFA file was a financial institution, the survey address was not changed.

Records that matched on number but not person were not included in the file that was sent to HCFA to be matched to the MADRS file. Instead, the person's record was flagged and the person was asked to provide their number in the subsequent LSOA reinterview.

Once the revisions were made, a revised file of numbers was sent to HCFA to be linked to the MADRS. The MADRS file contains 100 percent of all Medicare claims data for both Part A and Part B.

Matching to MADRS:

Matching MADRS Medicare claims to survey data was done carefully to avoid incorrect matches. To avoid erroneous linking, the matching process between the two files began with two variables: the HIC number and sex. Instances where the records matched on number but not on sex were scrutinized further.

Comparisons between other variables were made in an attempt to resolve differences between the survey data and Medicare claims data. The variables reviewed were age, date of birth, date of discharge, and date of death if there was one. Duplicate records were also discovered.

The majority of the cases where the records matched on number but not on sex were those situations where a spouse filed a claim under the other spouses' HIC number. It was obvious that the person on the HCFA file was not the same as the LSOA sample person. Some of the records matched a spouse who was not in the LSOA because the person was not 70 years old or over at the time of the SOA interview in 1984.

Newly matched data will be added to version 3 of the LSOA Public Use Data Tape. The indicator on the Person file will again be included as to whether or not there was a Medicare match.

The outcome of the first LSOA record match to the Medicare claims records revealed that 71 percent of the records matched. Information from the 1988 LSOA is expected to increase the percent of matched records because people were asked to provide social security numbers and HIC numbers if they were unknown in the previous interview.

As a result of the MADRS match, two files have been added to the LSOA Public Use Data Tape: a hospital record file and an other medical care use file. Shown below are the data items included on each of the files.

The first file includes data from Part A Medicare hospital records. There is one record for each discharge.

Data from Medicare Hospital Records Part A

LSOA identification number
date of birth
sex
date of discharge
length of stay
diagnostic codes (up to 5)
 DRG code
 MDC code (Major Diagnostic Categories not on MADRS)
surgical codes (up to 3)
total charge (rounded to nearest \$100)
type of hospital (general, psychiatric, TB)

The second file includes data from Part B Medicare records.

This file includes an indicator variable for home health care, hospice care, and outpatient care use for each year. There is one record for each person.

Data from Medicare Records

Part B

LSOA identification number
date of birth
sex
home health care indicator (yes/no for each year)
hospice indicator (yes/no for each year)
outpatient Medicare Part B indicator (yes/no for each year)

Results:

The following tables compare survey data to findings from the first match to the Master Enrollment File for the years 1984 to 1987. The match has been repeated in order to update information obtained in the prior MADRS match. Information from this second match were not available in time for inclusion in this paper.

Table 1 pertains to persons who were reported as having Medicare coverage. Most people who were covered in 1984 were covered in 1986. Of the 5,151 persons in the 1986 LSOA, 96.7 percent (4,983) had reported coverage in 1984. Of those who reported coverage, 10.3 percent did not provide the interviewer with a social security number. 17.8 percent of the numbers provided by respondents did not match a number on the Master Enrollment File. Of those who indicated that they did not have Medicare coverage, 76.2 percent provided a social security number. Curiously, 44.0 percent of those with no coverage reported a Medicare number that matched the Master Enrollment File.

The second table summarizes the reported 1984 and 1986 reinterview Medicare coverage. 4.6 percent of those who were self respondents in both interviews reported a change in coverage between the 1984 and the 1986 reinterviews. If a proxy was the respondent for one of the interviews, the percent of change in reported Medicare coverage nearly doubled.

Table 3 shows the distribution of Medicare coverage by type of respondent. In 1984, 97.1 percent of those who responded for themselves reported Medicare coverage; 91.2 percent of those covered provided their number and 81.1 percent of those covered who provided their numbers matched to the Master Enrollment File. Three percent reported no Medicare coverage, yet 80.7 percent provided a number. 58.7 percent of those who reported no Medicare coverage but gave a number matched to the Master Enrollment File. As one might expect, a match was more likely to occur if the number was provided by a self respondent rather than a proxy respondent. Regardless of whether or not Medicare coverage was reported, proxy respondents were 1) less likely to provide social security numbers and 2) less likely to provide social security numbers that matched.

Table 4 shows the results of the match to Medicare records by reported Medicare coverage in 1984 by respondent status. 8.8 percent who reported for themselves did not provide a number, and 17.3 percent provided a number that did not match. This

table shows again that proxy respondents generally did not provide social security numbers and that when they did the number would not match. Proxy and self respondents (72.0 percent) who reported Medicare coverage are shown in the last table.

The last table shows that 73.9 percent of the people who responded for themselves, were covered by Medicare and provided a number that matched to a Medicare claim. 53 percent of the numbers provided by self respondents resulted in a match to a hospital claim record. Although the number of matches for proxy respondents is small when compared to the self respondents, hospital use was found in 66.8 percent of the cases.

Conclusions:

Linking files and matching records is a complicated task but it is worth the effort. Matching the LSOA survey data to claims in the MADRS has added

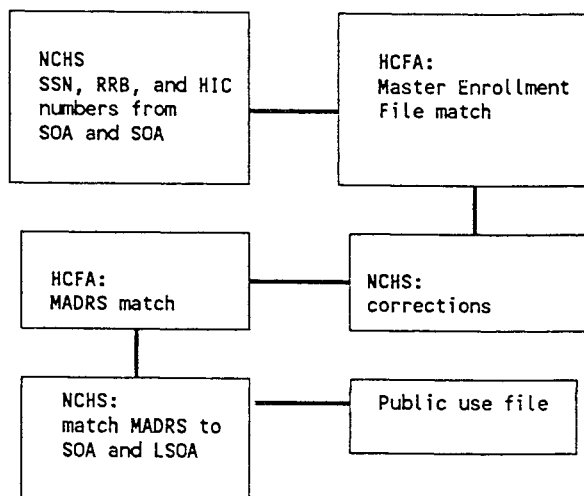
valuable data especially about hospitalizations to the LSOA data tape. Although proxy respondents reported social security numbers, the percent of matched records probably would be higher if the information necessary for the match was provided by a self respondent. A longitudinal survey such as the LSOA has an advantage in this situation in that if the information is unknown, the respondents can be asked for the information in a later interview.

References

MADRS File Documentation: Description of 100% - Part A/B - Skeleton. Dictionary OSDM. Health Care Financing Administration.

National Center for Health Statistics: 1988 Longitudinal Study of Aging Public Use File Version 2. Public Health Service. Hyattsville MD.

Figure 1. Longitudinal Study of Aging Medicare Match



SOURCE: NCHS, Longitudinal Study of Aging

Table 1. Comparison of reported Medicare coverage in 1984 with match to Master Enrollment file

Match Status 1984-1987	Total	Coverage in 1984	
		Yes	No
		Number in sample	
Total	5,151	4,983	168
		Percent distribution	
Total	100.0	100.0	100.0
No number given	10.8	10.3	23.8
Number given	89.2	89.7	76.2
No match	18.2	17.8	32.1
Match	71.0	71.9	44.0

SOURCE: NCHS, Longitudinal Study of Aging, 1986.

NOTE: No includes 8 people with coverage not reported.

Table 2. Medicare coverage and respondent as reported in 1984 and in 1986

Coverage in 1984 and 1986	Total	Respondent in 1984 and 1986		
		Self/ Self	Self/ Proxy	Proxy Proxy
		Number in sample		
	4,113	3,001	860	252
		Percent Distribution		
Total	100.0	100.0	100.0	100.0
No change	93.4	95.4	91.3	90.9
Yes/yes	93.1	94.5	89.9	88.1
No/no	1.2	1.0	1.4	2.8
Change	5.7	4.6	8.7	9.1
Yes/no	4.0	3.1	6.3	7.5
No/yes	1.7	1.5	2.4	1.6

SOURCE: NCHS, Longitudinal Study of Aging, 1986.

NOTE: Excludes people with respondent or coverage not recorded and people with no 1986 interview.

Table 3. Comparison of Medicare coverage reported in 1984 and match with 1984-1987 Master Enrollment file

Coverage reported on 1984 interview	Total	Respondent in 1984	
		Self	Proxy
		Number in sample	
	5,072	4,578	494
		Percent distribution	
With Medicare coverage	96.8	97.1	94.5
If covered, gave number	89.8	91.2	76.7
If number given, matched	80.1	81.1	69.8
Without Medicare coverage	3.2	2.9	5.5
If not covered, gave number	77.8	80.7	63.0
If number given, matched	57.1	58.7	47.1

SOURCE: NCHS, Longitudinal Study of Aging, 1986.

NOTE: Excludes people with either respondent or coverage not reported.

Table 4. Match status of people who were reported as covered by Medicare in 1984

Match status 1984-1987	Total	Respondent in 1984	
		Self	Proxy
		Number in sample	
Total	4,910	4,443	467
		Percent distribution	
Total	100.0	100.0	100.0
No number given	10.2	8.8	23.3
Number, no match	17.8	17.3	23.1
Match	72.0	73.9	53.5
No hospital use	33.1	34.7	17.8
Hospital use	38.9	39.2	35.8

SOURCE: NCHS, Longitudinal Study of Aging, 1986.

NOTE: Excludes people with either respondent or coverage not reported.

Table 5. People who were reported as covered by Medicare in 1984 and had match to Master Enrollment file

Match status 1984-1987	Total	Respondent in 1984	
		Self	Proxy
		Number in sample	
Total with match	3,535	3,285	250
		Percent distribution	
Total with match	100.0	100.0	100.0
No hospital use	46.0	47.0	33.2
No use	18.9	19.3	13.6
Other only	27.1	27.7	19.6
Hospital use	54.0	53.0	66.8
Hospital only	6.9	6.7	10.4
Both	47.1	46.4	56.4

SOURCE: NCHS, Longitudinal Study of Aging, 1986.

NOTE: Excludes people with either respondent or coverage not reported.