Vital and Health Statistics

A Method to Redefine Stays on the 1985 National Nursing Home Survey

Series 2: Data Evaluation and Methods Research No. 115

This report describes a method for standardizing definitions of episodes of nursing home care in the 1985 National Nursing Home Survey. The method shows how the information on nursing home admissions and discharges collected on the Current and Discharged Resident Questionnaires can be used to redefine the endpoints of nursing home stays. The report also explains how errors caused by missing and inconsistent nursing home admission and discharge data were resolved.

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Centers for Disease Control
National Center for Health Statistics

Hyattsville, Maryland
March 1992
DHHS Publication No. (PHS) 92-1389

Suggested Citation

Jonas BS, Madans JH, Rothwell ST, Bush MA, Feldman JJ. A method to redefine stays on the 1985 National Nursing Home Survey. Vital Health Stat 2(115). 1992.

Library of Congress Cataloging-in-Publication Data

A Method to redefine stays on the 1985 National Nursing Home Survey.
p. cm. – (Vital and health statistics. Series 2, Data evaluation and methods research ; no. 115) (DHHS publication ; no. (PHS) 92–1389)
"March 1992."
Includes bibliographical references.
ISBN 0–8406–0455–6
1. National Nursing Home Survey (U.S.)—Evaluation. 2. Nursing homes—United States—Length of stay—Statistics. I. National Nursing Home Survey (U.S.) II. National Center for Health Statistics (U.S.) III. Series. IV. Series: DHHS publication ; no. (PHS) 92–1389. V. Series. DHHS publication ; no. (PHS) 92-1389.
[DNLM: 1. Length of Stay—United States—statistics. 2. Long Term Care—United States—statistics. 3. Nursing Homes—utilization—United States. 4. Research Design. W2 A N148vb no. 115]
RA409.U45 no. 115
[RA997]
362.1'0723 s—dc20
[362.1'6'0973]
DNLM/DLC
for Library of Congress 91–36707
CIP

National Center for Health Statistics

Manning Feinleib, M.D., Dr.P.H., *Director*

Jacob J. Feldman, Ph.D., *Associate Director for Analysis and Epidemiology*

Gail F. Fisher, Ph.D., *Associate Director for Planning and Extramural Programs*

Peter L. Hurley, *Associate Director for Vital and Health Statistics Systems*

Robert A. Israel, *Associate Director for International Statistics*

Stephen E. Nieberding, *Associate Director for Management*

Charles J. Rothwell, *Associate Director for Data Processing and Services*

Monroe G. Sirken, Ph.D., *Associate Director for Research and Methodology*

David L. Larson, *Assistant Director, Atlanta*

Division of Analysis

Jennifer H. Madans, Ph.D., *Acting Director*

Jennifer H. Madans, Ph.D., *Deputy Director*

Diane Makuc, Dr.P.H., *Chief, Analytical Coordination Branch*

Sandra T. Rothwell, *Chief, Longitudinal Analysis Branch*

Jennifer H. Madans, Ph.D., *Acting Chief, Demographic Analysis Staff*

Contents

List of text tables

List of figures

Symbols	
- - -	Data not available
. . .	Category not applicable
-	Quantity zero
0.0	Quantity more than zero but less than 0.05
Z	Quantity more than zero but less than 500 where numbers are rounded to thousands
*	Figure does not meet standard of reliability or precision

A Method to Redefine Stays on the 1985 National Nursing Home Survey

by Bruce S. Jonas, Ph.D., Jennifer H. Madans, Ph.D., Sandra T. Rothwell, M.P.H., Mary Ann Bush, M.S., and Jacob J. Feldman, Ph.D., Division of Analysis, National Center for Health Statisitcs

Introduction

The purpose of this report is to describe a method for standardizing definitions of episodes of nursing home care in the 1985 National Nursing Home Survey (NNHS). This standardization is necessary because not all nursing homes use the same definition of a nursing home stay and, thus, reported lengths of stay are not always comparable. Length of stay in nursing homes is important for policy decisions such as allocation of resources and planning for long-term care insurance. The construction of length of stay estimates is complicated in surveys, such as the NNHS, which use facility based definitions of a nursing home stay as the sampling frame for current and discharged residents. It is necessary to consider a resident's entire pattern of nursing home usage, including multiple nursing home stays and intervening hospital utilization, in calculating length of stay. Studies of length of stay have been published using the 1977 NNHS (1,2). However, in that survey, information was not collected on nursing home stays other than the sampled stay. More recent surveys do include information on additional nursing home stays and hospital utilization. Adjustments to length of stay estimates have been made by Spence and Wiener (3) using data from the 1985 NNHS, and Short et al. (4) using the 1987 National Medical Expenditure Survey. The methods used differ from the method presented in this report.

Background

The 1985 National Nursing Home Survey (NNHS) collected extensive information about nursing homes and their residents. The sample frame consisted of all nursing homes listed in the 1982 National Master Facilities Inventory (NMFI), nursing homes identified in the 1982 Complement Survey of the NMFI, facilities that opened between 1982 and June 1984, and hospital-based nursing homes identified by the Health Care Financing Administration. The sample frame contained 20,479 nursing homes. The sample was selected using a stratified two-stage probability design. The first stage was the selection of the individual facilities. A sample of 1,220 nursing and related care homes was selected, of which 1,079 nursing homes (88.4 percent) participated in the survey. The second stage was the selection of current residents and discharges within homes. This stage was carried out by the interviewers at the time of their visits to the facilities. The sample frame for current residents consisted of all people on the register of the facility on the evening prior to the day of the survey. A sample of five or less current residents per facility was selected. The sample frame for discharges consisted of all discharges (whether the resident was alive or dead) that occurred during the 365 days prior to the survey date. A sample of six or fewer discharges per facility was selected (5).

Data were collected on 5,243 individuals (97 percent response rate) who were current residents in these institutions at the time of contact (Current Resident Sample) and on 6,023 discharges (95 percent response rate) that occurred during the 365 days preceding the date of contact (Discharged Resident Sample). Detailed information was collected from nursing home facility records and appropriate staff regarding dependence in activities of daily living; functional impairments; diagnoses; the receipt of services; cognitive and emotional status; sources of payments; and hospital use and nursing home stays prior to, during, and subsequent to the sampled stay. The data collection instruments were the Current Resident Questionnaire (CRQ) and the Discharged Resident Questionnaire (DRQ).

One of the objectives of the NNHS was the collection of information on patterns of nursing home use. However, the sample design of the NNHS is not compatible with this objective. First, information was collected on discharge events rather than on discharged residents. Since an individual could have been admitted and discharged several times during the course of a 365-day period and since each discharge is listed separately, individuals can appear on the sampling list multiple times. It is also possible for a current resident to be included in the discharge sampling frame if he or she was discharged during the 365 days prior to the survey and then readmitted to the sample facility. Thus, an individual might be selected into the sample(s) more than once.

The second problem with the sample design concerns how endpoints of a nursing home stay are defined. Stay endpoints are identified by the facility and there is variation across facilities as to how stays are defined, particularly in how temporary transfers to hospitals are treated. Some nursing homes consider any transfer to the hospital to be a formal discharge and the resident's return to be a formal readmission as illustrated by facility A in figure 1. In other facilities, as illustrated by facility B, no formal discharge is made and the nursing home stay includes the hospital stay. These administrative rules concerning official discharge practices determine the content of the discharged resident sampling list and the admission date associated with the listed stays in both the current resident and discharge samples. Thus, while the two examples in the figure represent equivalent episodes of nursing home use, the manner in which they would be listed and subsequently sampled would lead to length of stay calculations that would not lead to this conclusion. The discharge sampling list would contain one listing for facility B but two for facility A. Selection of either of the two stays in facility A would yield data on only part of the nursing home stay, and the length of stay calculated from the recorded admission and discharge dates would underestimate the true length of stay. In order to calculate length of stay in a consistent fashion, it is necessary to develop a uniform set of rules for the definition of a nursing home stay in terms of the appropriate admission and discharge dates.

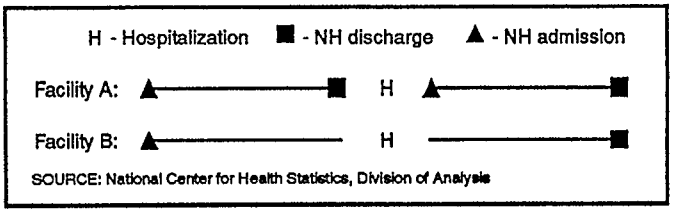

Figure 1. Illustration of variation in the definition of nursing home stays

Additional questions were added to both the CRQ and DRQ to obtain information on nursing home utilization that would allow the analyst to control the effects of facility variation in recordkeeping. The purpose of this report is to describe how the information on nursing home admissions and discharges collected on the Current and Discharged Resident Questionnaires can be used to redefine the endpoints of nursing home stays. As has been found in many other surveys, which collect retrospective information on dates of occurrence, the data on additional nursing home use contained considerable missing and inconsistent entries. Thus, the data first had to undergo extensive editing. These editing procedures are described below. This report describes how these editing and redefinition procedures affect the data themselves. The results presented are unweighted and are not intended to be used as national estimates.

Procedures for redefinition of admission and discharge dates of the sampled stay

The first step in the redefinition process was the creation of a person based file where each record contained the data for a single resident. Multiple records of the same individual were sorted chronologically and only the earliest record was retained in the file. No data are lost as a result of this procedure since each record should contain all information on a resident's nursing home and hospital use. This resulted in files which contain 5,200 current resident cases and 5,981 discharged resident cases.

The procedures for redefinition of admission and discharge dates were developed with the goal of providing a uniform approach to identifying the endpoints of each nursing home stay. A nursing home stay is defined as the time spent in the nursing home inclusive of time spent in an acute care hospital if the resident returned to the nursing home after the hospital stay without spending any time in the community or in another nursing home. The questions added to the CRQ and DRQ to help make this determination are listed in figure 2. Information was obtained on the dates of other stays the resident had in that facility. Space was available to code up to eight pairs of additional admission and discharge dates. For current residents these stay pairs occurred prior to the sampled admission date. However, for the discharged resident sample, the stay pairs can be any combination of stays prior to the sampled admission and subsequent to the sampled discharge. Information was also obtained on whether the discharge was to a short-stay or general hospital. When the stay just prior to the sampled stay ended with a discharge to a hospital, additional information was collected in the CRQ on the number of nights the resident spent in the hospital during that stay. For this one interval only, it is possible to determine whether the interval between the two nursing home stays was spent exclusively in a hospital. For the other stays, information on the duration of the intervening hospital stay was not collected.

For up to 8 stays:

- On what other dates was resident admitted to and discharged from this facility?
- Was this discharge to a short-stay or general hospital?

For first stay only:

- Number of nights resident spent in hospital?

Figure 2. Questions added to the 1985 National Nursing Home Survey to obtain nursing home and hospital utilization history

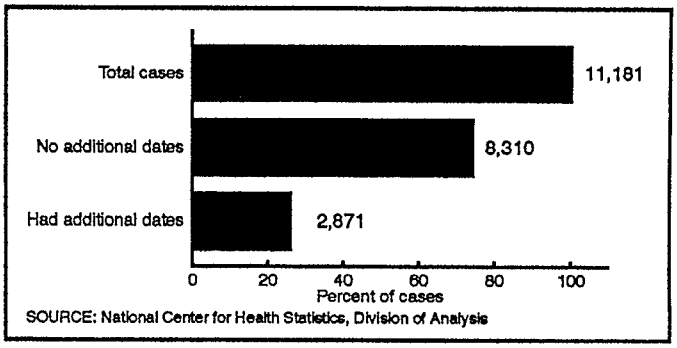

Figure 3. Number and percent of cases in the combined CRQ/DRQ sample with and without information on additional nursing home stays

The number of cases for which information on additional stays in the nursing home was reported is given in figure 3. No additional stays are listed for 74 percent of the cases (8,310 residents). Thus, redefinition of admission and discharge dates to reflect the true parameters of the nursing home stay might be required for 2,871 cases or 26 percent of the combined CRQ and DRQ samples. Specifically, the admission date for the sampled stay was to be redefined if 1) the resident had previously been in that facility, 2) the resident had been discharged to a hospital at the end of the previous stay, and 3) the entire interval between discharge and the admission date of the sampled stay had been spent in the hospital. Similarly the discharge date for the sampled stay was to be redefined if 1) the resident had been discharged to a hospital at the end of the sampled stay, 2) the resident returned to the nursing facility, and 3) the entire interval between nursing home discharge and subsequent readmission had been spent in the hospital. Because information on place of discharge was collected for all reported stays in the nursing home, but information on the number of days spent in the hospital was only obtained for the most recent hospitalization immediately prior to the sampled stay, the following rule was developed to determine which intervals should be considered to have included only a hospital stay. Intervals of 21 days or less were treated as exclusive hospital stays while intervals of more than 21 days were treated as having included some time spent in the community. The goal was to preclude redefinition if it appeared the resident spent any time in the community.

The 21-day cutoff rule was developed by analyzing the hospitalization that occurred prior to the sampled stay where number of nights spent in the hospital was obtained. Ninety-nine percent of the intervals that were 21 days or less were spent exclusively in short-stay or general hospitals. Conversely, only 38 percent of the intervals that were greater than 21 days were spent exclusively in short-stay or general hospitals (table A).

Table A. Number and percent of intervals spent entirely in hospital by interval length

Interval length	Nature of interval	
	Hospital stay	Mixed
	Number	
21 days or less	472	1
More than 21 days	44	72
	Percent	
21 days or less	99.8	0.2
More than 21 days	37.9	62.1

If the following three conditions are met: 1) the resident had a previous stay in the home, 2) the stay ended with a discharge to a hospital, and 3) the interval between the discharge and subsequent readmission is 21 days or less, the new admission date is set to the admission date for the previous stay. The process continues with the investigation of each preceding stay until a true discharge is encountered. Similarly, if the following three conditions are met: 1) the resident had a subsequent stay in the home, 2) the sampled stay ended with a discharge to a hospital, and 3) the interval between the discharge and subsequent readmission is 21 days or less, the new discharge date is set to the discharge date for the subsequent stay. The process continues with the investigation of each subsequent stay until a true break is encountered.

A schematic representation of the redefinition process is presented in figure 4. In the first scenario, the admission date is not changed because the previous stay did not end with a discharge to a hospital. In the second and third scenarios, the admission date is set to the admission date of the previous stay because conditions for redefinition are met. In these examples, the admission date is not set to an even earlier admission date because conditions for redefinition are not met in the second interval preceding the sampled stay. In the fourth scenario the admission date is set to an earlier date that includes two previous stays because conditions for redefinition are met in both intervals. The sampled discharge dates were redefined in an analogous manner in the forward direction.

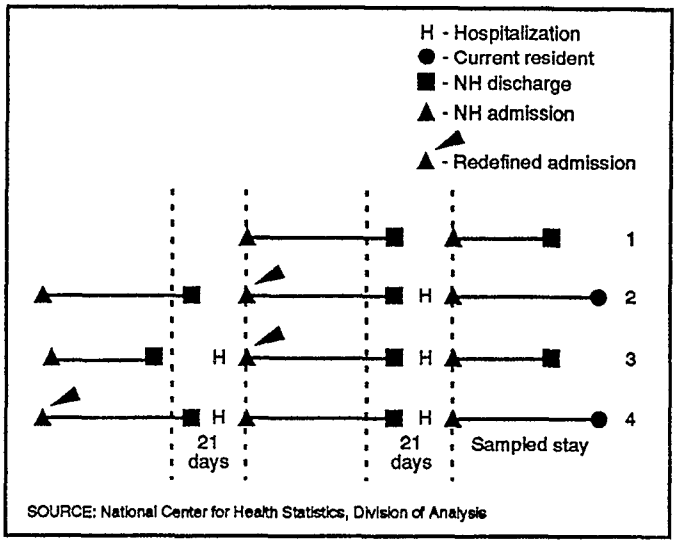

Figure 4. Schematic representation of the sampled admission/discharge date redefintion procedure

Procedures for editing date inconsistencies

The information on the dates of additional stays contained errors and omissions, requiring extensive editing to resolve the inconsistencies. Of the 2,871 cases which contained information on other stays in the nursing home, 1,084 (37.8 percent) had date inconsistencies that required editing. For purposes of this discussion the entire set of eight additional admission and discharge dates is referred to as the "date array" and any one admission and discharge date pair is referred to as a "stay pair."

The report of any date information was considered documentation of a nursing home stay. The sampled admission date, sampled discharge date (DRQ only), and interview date were taken to be correct and the date array was evaluated in relation to these dates. Information from the entire date array was used to reconstruct stay pair information. Stay pairs were deleted only as a last resort and admission or discharge dates were not moved across stay pairs. Dates were instead corrected or, if necessary, imputed whenever possible. When two or more solutions were available for reconstructing a consistent date array, we chose the one that changed the least number of dates in the existing date array. An effort was made to be conservative in the reconstruction of date arrays with respect to the 21-day cutoff that would determine eligibility for date redefinition. Finally, the same mechanism was used to resolve all cases with the same type of inconsistency.

Date arrays were reviewed, errors detected, and edit rules were applied to resolve errors. A consistent date array is defined as one with the earliest admission and discharge pair in the first position, with subsequent stay pairs following in chronological order, with missing date codes (i.e., "989898") filling the remainder of the array, and with date intervals that are mutually exclusive, i.e., no two stay pairs have overlapping or embedded date information. In addition, for current residents, all dates listed must fall before the sampled admission date. For discharged residents, the dates can fall before the sampled admission date or between the sampled discharge date and the interview date.

All records were reviewed using computer edit programs and manual inspection to determine whether they met the consistency requirements. Corrections made to inconsistent date arrays can be classified into three groups: 1) re-sorting of dates, 2) changing dates, and 3) imputing missing dates.

Re-sorting dates

The simplest errors were those where the stay pairs were out of chronological order. The entire order could be reversed or individual stay pairs could be out of chronological order. Occasionally admission and discharge dates within a stay pair were transposed. These types of errors were corrected first and involved reordering the stays.

Changing inconsistent dates

After the date array was sorted into proper sequence, the records with remaining problems were reviewed to determine whether, in the judgment of the authors, they appeared to have transcription or keying errors. Errors of this type were corrected. Errors that did not appear to be the result of transcription or keying errors required further attention. For example, stay pairs embedded within other pairs were deleted and the longer stay was maintained on the file. Overlapping stay pairs were combined into one longer stay covering the entire time period. Stay pairs occurring after the sampled admission date for current residents, which could not be corrected through one of the above rules, were deleted. Finally, discharge dates for the most recent stay for discharged residents, which indicated that the subject was still a resident in the facility at the time of the field interview, were checked against question 11d, "Is _____ still a resident [of the sample facility]?" and question 11e, "Was _____ discharged alive?" When the answers to these questions indicated that the resident had been discharged, the discharge date was corrected either by imputing a discharge date (see section on imputation below) or by inserting the date of death depending on the information provided in question 11d and question 11e.

Imputing missing dates

Date imputation rules were developed for cases in which one or more dates in the date array contained missing information. Two types of missing date information occurred. In the first type, the missing dates are bracketed on either side by valid dates. In this case the rule employed is to impute the missing date to the midpoint of the interval defined by the valid dates. For example, in the following date array, (010184 989898)

(010185 020185) (021585 030785) (031185 031985) (989898 989898) (989898 989898) (989898 989898) (989898 989898), the first pair's discharge day, month, and year were imputed to be the midpoint between the first and second admission dates (070184).

In the second type, missing dates could be at one or both ends of the valid subset of the date array. In this case the rule employed was to delete the stay pair. For example, consider a record with the following date array: (989898 121584) (010185 020185) (021585 030785) (031185 031985) (032185 032585) (032785 033185) (041085 043085) (989898 989898), where the sample admission date is 091085, the sample discharge date is 111585 and the interview date is 121785. Since this first missing admission is at the beginning of the series and before the sampled admission date, there is no possibility of using the information in the rest of the series to determine the range within which the date might fall. In this case no midpoint imputation is possible. The decision to delete such stays, thus not considering them in the redefinition, is the most conservative alternative since it results in the shortest length of stay. The figures and tables that follow in the report are based on this decision.

An alternative method for dealing with these missing dates would be to set the missing date to the date preceding or following the valid discharge or admission date in the stay pair. In order to preserve the first stay pair in the example above, the missing admission date would be set to "121484," the day preceding the existing discharge date. The choice of a single day represents a conservative limit on the length of the newly preserved stay, and this imputation method would also likely underestimate the true length of stay. If this second approach were employed, the results of the redefinition process would be slightly different. Fifty-six additional cases would have their admission dates redefined, 40 current residents and 16 discharged residents, increasing the total number of redefined admission dates specified in table C, page 8, from 1,434 to 1,490. The percent of total cases undergoing redefinition of admission and/or discharge dates would increase slightly from 17.6 percent to 18.1 percent resulting in only minor changes in figure 5, page 8. Missing discharge dates at the end of the series would be handled in an analogous manner. However, there were no cases of this type where the imputed date would have become the redefined discharge date. Thus, there would be no difference in the results of the redefinition for discharge dates between the two methods of handling missing end dates. Of the 16 additional discharged resident cases where the alternative imputation method resulted in redefinition of the admission date, one had a redefined discharge date using another redefinition rule. Therefore, the total number of discharged resident cases with any redefinition of dates increases from 1,319 to 1,334 in table E, page 10. The median length of stay for these cases would change minimally.

It should be noted that cases often had more than one of the types of errors described above. These cases were handled using the most appropriate combination of rules. The complexity of these cases was such that writing computer algorithms to effect the rule logic was impractical and most of the cases were scrutinized individually and corrected by hand.

The types of inconsistencies found among the 1,084 cases are shown in table B. In 129 cases (11.9 percent) the only correction required was the reordering of the date array. In 773 cases (71.3 percent), changes in one or more dates were required in addition to any possible reordering. Finally, in 182 cases (16.8 percent), the corrections included imputation of one or more dates and may also have included reordering or date correction.

Table B. Number and percent distribution of cases by type of date corrections

Correction	*Number*	*Percent distribution*
Total	1,084	100.0
Re-sorting only	129	11.9
Date correction	773	71.3
Date imputation	182	16.8

Results of date redefinition

Figure 5 shows the results of the date redefinition procedure. As noted, 74 percent of cases did not contain information on additional stays in the nursing home and, therefore, were not candidates for redefinition. Eighteen percent of the total cases, but over two thirds of those eligible for redefinition, had their sampled admission or discharge dates redefined. The remaining 8 percent had additional dates but the dates for the sampled stay were not redefined. The admission date was redefined in 648 (68 percent) of the CRQ cases eligible for redefinition. Of the 1,920 DRQ cases eligible for redefinition, only the admission date was redefined in 572 (30 percent) cases, only the discharge date was redefined in 533 (28 percent) cases, and both endpoints were redefined in 214 (11 percent) cases. In total, 1,434 cases had their admission date redefined and 747 had their discharge date redefined (tables C and D).

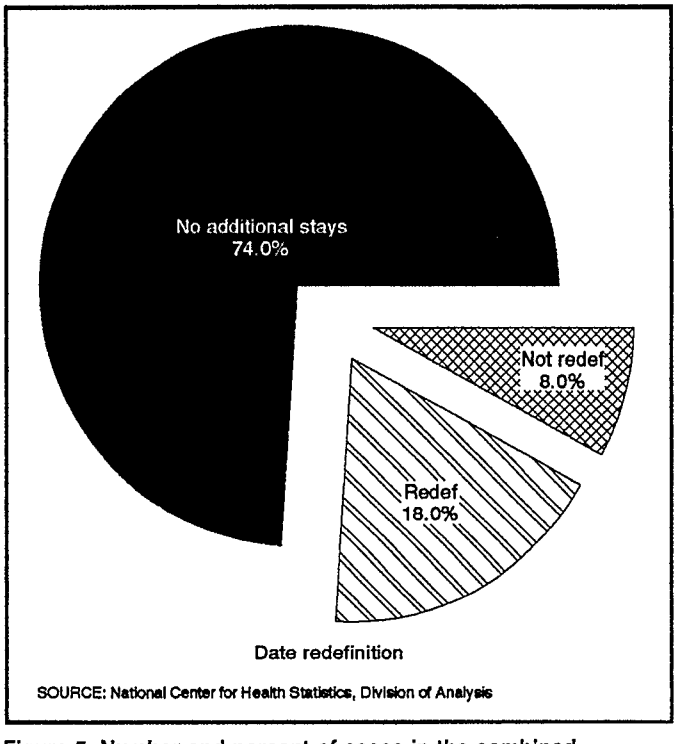

Figure 5. Number and percent of cases in the combined CRQ/DRQ sample by whether dates were redefined

Table C. Number and percent distribution of eligible cases by redefinition status of admission dates

Redefinition status	*Number*	*Percent distribution*
Total eligible	2,871	100.0
Admission date redefined	1,434	49.9
Admission date not redefined	1,437	50.1

Table D. Number and percent distribution of eligible cases by redefinition status of discharge dates

Redefintion status	*Number*	*Percent distribution*
Total eligible	1,920	100.0
Discharge date redefined	747	38.9
Discharge date not redefined	1,173	61.1

Figure 6 shows the magnitude of the changes made to the admission dates for those current residents and discharged residents in which the sampled admission date was redefined. In 56 percent of the cases, the admission date was redefined as having occurred at least 1 year earlier than the recorded admission date. In nearly 20 percent of the cases the redefined date is more than 4 years earlier than the recorded date. Thus, redefinition tends to make a large difference in the date for a majority of the cases where the admission date was redefined.

The impact of redefinition on the discharge date for those discharged residents undergoing a change in discharge date is shown in figure 7. In order to be selected into the discharge sample, the discharge had to have occurred within the 365 days prior to the field interview. As a result, discharge dates could only be brought forward by a maximum of 365 days. The difference between the redefined and recorded discharge date is more than 3 months in 56 percent of the cases. In about one third of the cases this difference is 6 months or more. Thus, even when the range of time difference is restricted to less than 1 year, redefinition makes marked differences in the sampled discharge date.

The impact of redefining admission and discharge dates is illustrated by comparing the distribution of length of stay based on recorded admission and discharge dates

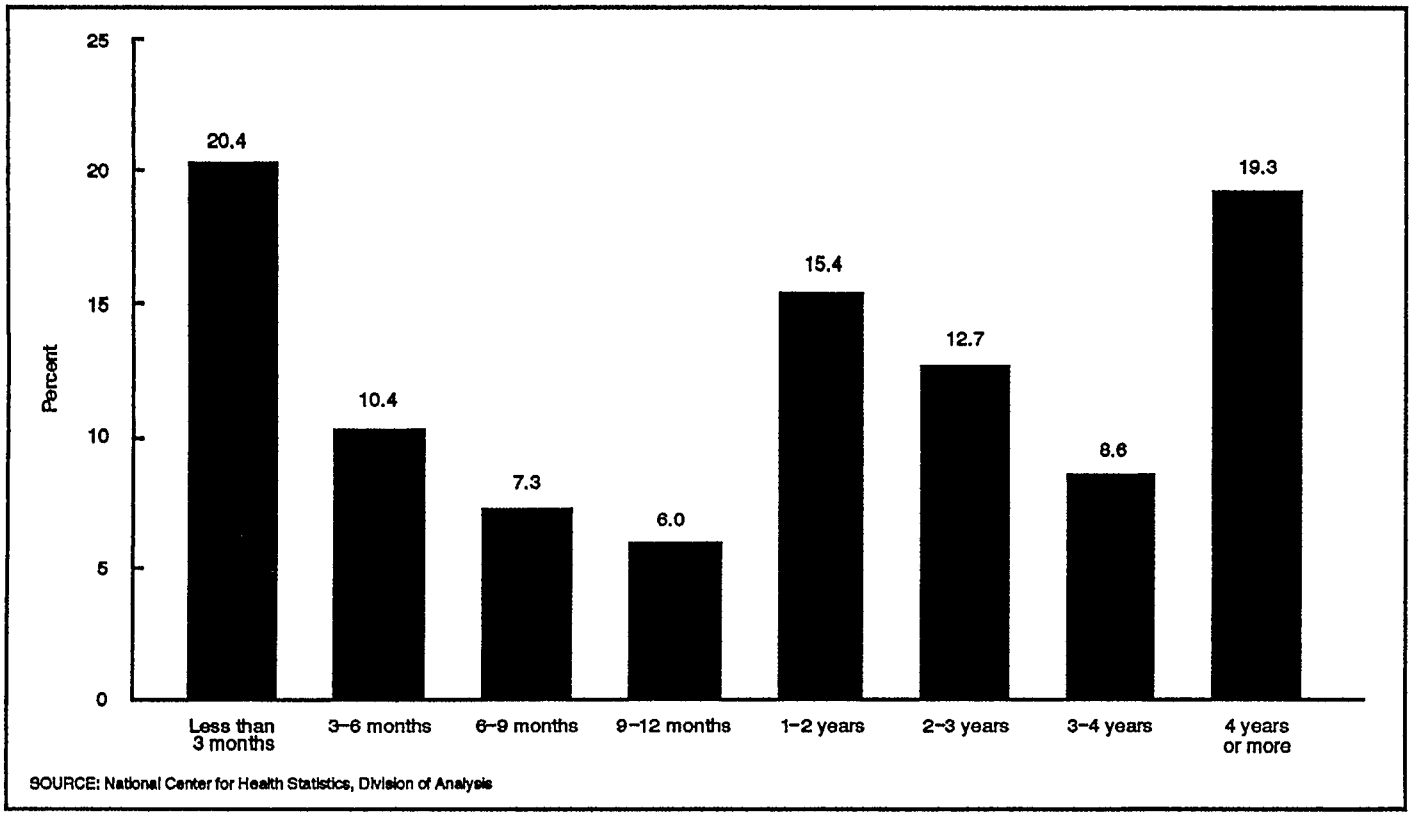

Figure 6. Percent of cases in the combined CRQ/DRQ sample by difference between sampled and redefined admission dates (N=1,434)

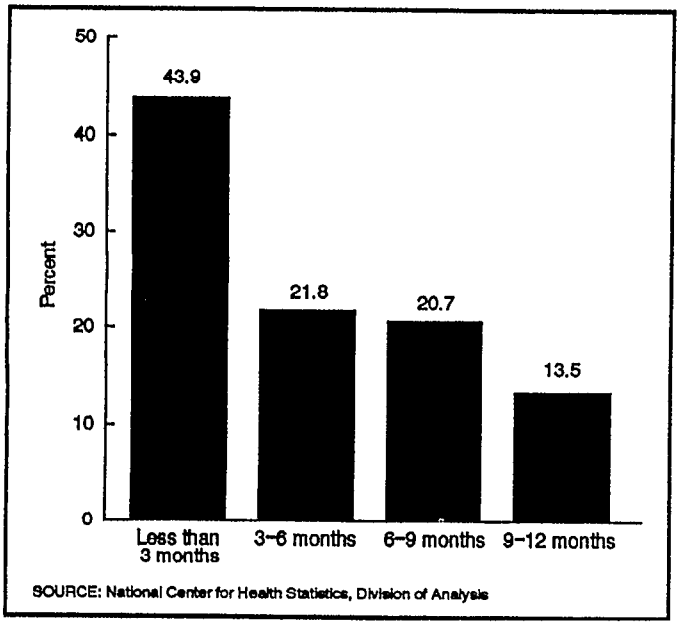

Figure 7. Percent of cases in the DRQ sample by difference between sampled and redefined discharge dates (N=747)

to length of stay based on the redefined dates (table E). Using the original 5,981 discharged residents, the median length of stay increases 58 percent from 135 days when based on the recorded length of stay to 213 days when based on redefined length of stay. Focusing on only those 1,319 discharged cases where dates were redefined, the median length of stay increases from 134 days when based on recorded endpoints to 562 days when based on redefined endpoints, an increase of 419 percent. In addition, about 55 percent of stays are less than 6 months long when the length of stay is based on recorded dates for both the total group and those cases undergoing redefinition. When redefined dates are used, 47 percent of the total sample of discharges have a length of stay less than 6 months and only 22 percent of the subset of redefined stays have stays under 6 months. Similarly, 34 percent of the 5,981 sample stays have a length of 1 year or more when using recorded dates, but 41 percent of discharge cases have stays 1 year or more in length when redefined dates are used. Of the 1,319 redefined cases, 30 percent of the discharges have recorded lengths of stay of 1 year or more whereas almost 62 percent of the redefined discharges have recalculated lengths of 1 year or more.

Furthermore, 37 percent (494) of the 1,319 discharged residents with redefined endpoints, had their discharge date set to the interview date since the redefinition process determined that they were still in the nursing home at the time of the interview. Subsequent waves of data collection in the National Nursing Home Survey Followup will extend the discharge date estimates for these residents as well as for current residents and therefore the statistics presented above must be considered conservative bounds on the differences between sampled and redefined length of stay.

Table E. Number and percent distribution of Discharged Resident Questionnaire (DRQ) cases by length of stay (LOS) for sampled and redefined admission and discharge dates

Length of stay	Sampled LOS		Redefined LOS	
	Number	Percent distribution	Number	Percent distribution
All DRQ cases				
Total	5,981	100.0	5,981	100.0
Less than 3 months	2,662	44.4	2,249	37.6
3–6 months	621	10.4	590	9.9
6–9 months	406	6.8	412	6.9
9–12 months	269	4.5	294	4.9
1–2 years	643	10.7	736	12.3
2–3 years	392	6.6	441	7.3
3–4 years	225	3.8	306	5.1
More than 4 years	763	12.8	953	16.0
Median (days)	135	...	213	...
DRQ cases undergoing redefinition				
Total	1,319	100.0	1,319	100.0
Less than 3 months	571	43.3	158	12.0
3–6 months	171	12.9	140	10.6
6–9 months	107	8.1	113	8.5
9–12 months	73	5.5	98	7.4
1–2 years	154	11.7	247	18.7
2–3 years	94	7.1	143	10.9
3–4 years	45	3.5	126	9.6
More than 4 years	104	7.9	294	22.3
Median (days)	134	...	562	...

Effects of data editing on date redefinition

Because the procedures used to edit the date arrays could affect redefinition and length of stay computations, the edit rules were designed to err in the direction of not redefining dates and therefore limiting length of stay. Table F shows the percent of edited and unedited cases undergoing redefinition of the sampled admission dates for the combined current resident and discharged resident sample. Thirty-one percent of the edited cases were redefined while 61 percent of the unedited cases were redefined. Thus, for the cases where inconsistent information was found and corrected, the percent of those cases redefined is half of that for the cases that were correct to start. In addition, if redefinition occurred, the change in dates was more limited. As shown in figure 8, the admission dates for edited cases were more likely to change by less than 3 months, 3–6 months, and 6–9 months than are the unedited cases.

Table F. Number and percent distribution of cases eligible for redefinition by redefinition status of admission date and edit status of date arrays

Redefinition status of admission date	*Edit status of date array*		
	Cases	*Edited*	*Not edited*
		Number	
Total	2,871	1,084	1,787
Redefined	1,434	341	1,093
Not redefined	1,437	743	694
		Percent distribution	
Total	100.0	100.0	100.0
Redefined	49.9	31.5	61.2
Not redefined	50.1	68.5	38.8

Similar results were found for discharge dates. Table G shows the percent of edited and unedited cases undergoing redefinition of sampled discharge dates. Eleven percent of the edited cases were redefined while 47 percent of the unedited cases were redefined. Again, if redefinition occurred, the change in discharge dates was more limited for edited cases. As indicated in figure 9, the discharge dates for edited cases were more likely to change by less than 3 months (78 percent versus 42 percent).

Combining the results found for admission and discharge dates, table H shows that the length of stay was changed in 30 percent of the edited cases while the length of stay was affected in almost 80 percent of the unedited

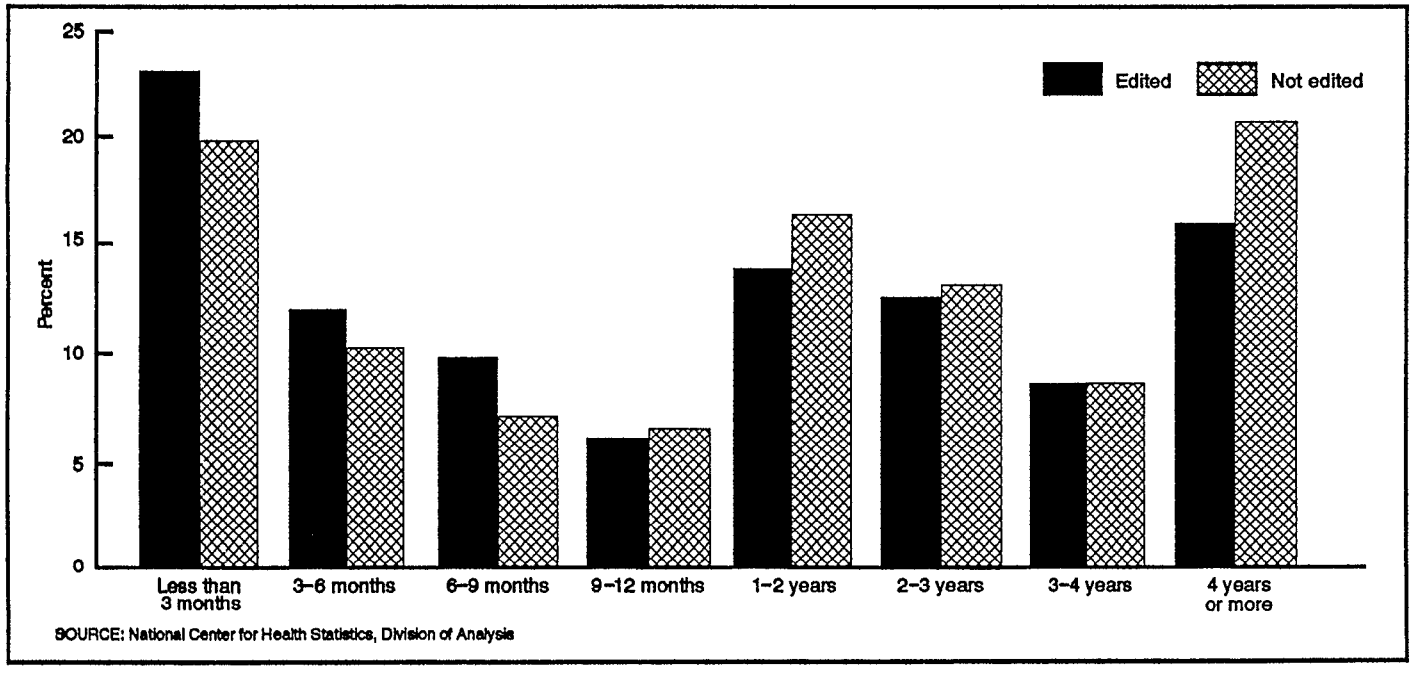

Figure 8. Difference between sampled and redefined admission dates by whether cases were edited, CRQ and DRQ (N=1,434)

Table G. Number and percent distribution of discharged resident questionnaire cases eligible for redefinition by redefinition status of discharge date and edit status of date arrays

Redefinition status of discharge date	Edit status of date array		
	Cases	Edited	Not edited
		Number	
Total	1,920	423	1,497
Redefined	747	45	702
Not redefined	1,173	378	795
		Percent distribution	
Total	100.0	100.0	100.0
Redefined	38.9	10.6	46.9
Not redefined	61.1	89.4	53.1

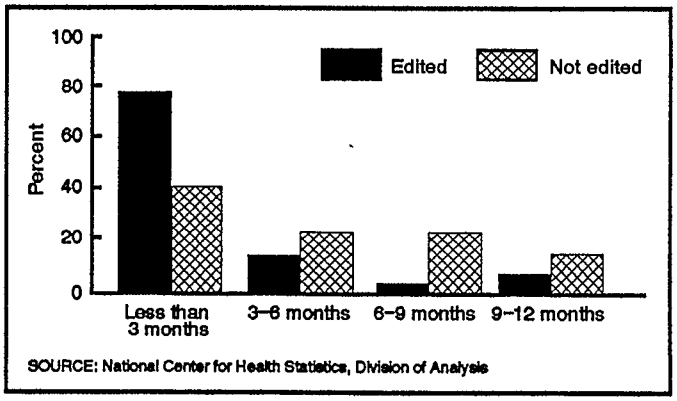

Figure 9. Difference between sampled and redefined discharge dates by whether cases were edited, DRQ only (N=747)

cases which contained additional dates. Figure 10 shows the difference between sampled and redefined length of stay for edited versus unedited cases. Edited cases have a larger percent in the less than 3-month interval (35 percent versus 30 percent), while the unedited cases have

Table H. Number and percent distribution of discharged resident questionnaire cases eligible for redefinition by redefinition status of length of stay (LOS) and edit status of date arrays

Redefinition status of LOS	Edit status of date array		
	Cases	Edited	Not edited
		Number	
Total	1,920	423	1,497
Redefined	1,319	129	1,190
Not redefined	601	294	307
		Percent distribution	
Total	100.0	100.0	100.0
Redefined	68.7	30.5	79.5
Not redefined	31.3	69.5	20.5

larger percents in the next three intervals: 3–6 months, 6–9 months, and 9–12 months. The remaining intervals of 1 year or more tend to be about even between edited and unedited cases. However, the percents for edited cases are based on relatively small numbers. The net impact of the date edits is that both the proportion of edited cases undergoing redefinition as well as the magnitude of the increase in length of stay for those edited cases that were redefined is less than that for cases that were correct as recorded.

A specific example of a decision to err conservatively with respect to admission and discharge date redefinition is the midpoint rule for imputed dates. An examination of date array patterns revealed that in over 93 percent of the cases where a discharge date was missing, the interval between the two successive admission dates was greater than 42 days. Thus, imputation using the midpoint date in these cases kept the separation interval greater than 21 days on either end. The imputed date served primarily to mark the place in the date array and there was no redefinition over the interval.

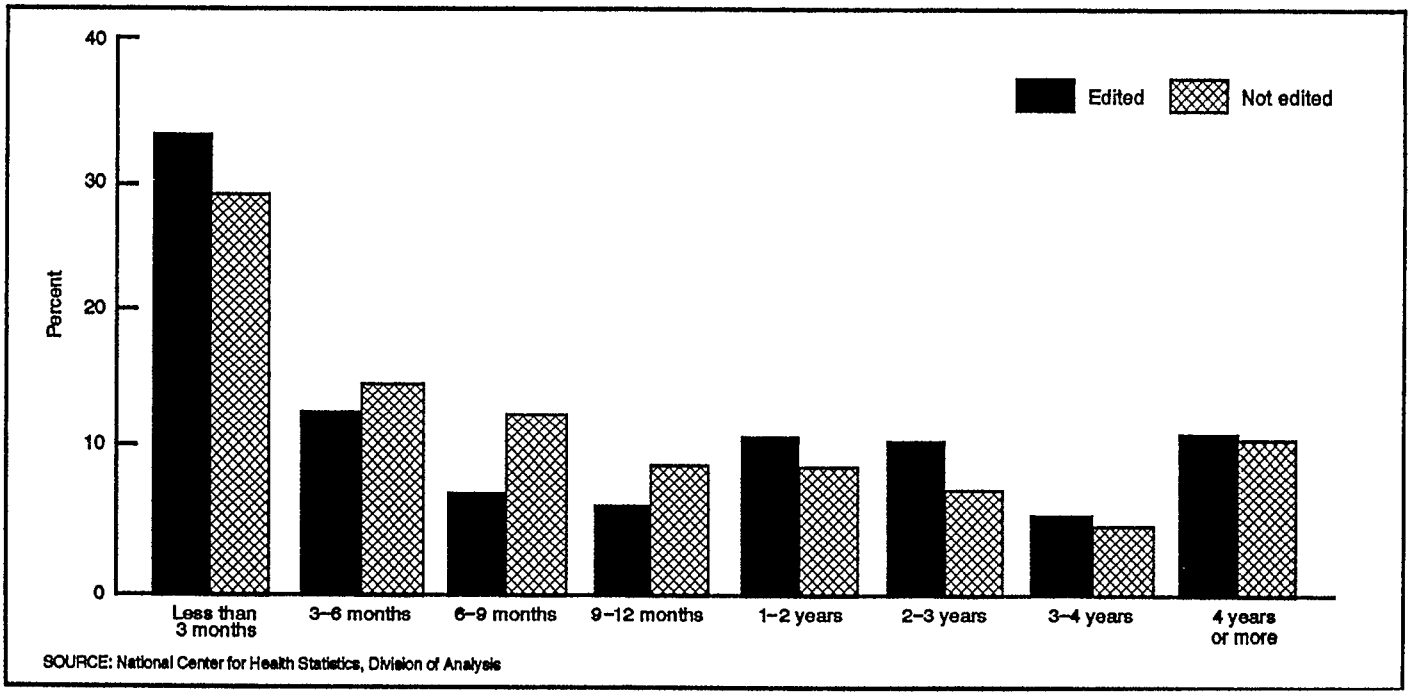

Figure 10. Difference between sampled and redefined length of stay by whether cases were edited, DRQ only (N=747)

Summary

A method has been described whereby episodes of nursing home utilization can be reconstructed using new information collected in the National Nursing Home Survey (NNHS). Additional questions made it possible to combine nursing home stays that were broken only by visits to short-stay hospitals and to improve compatibility of stays reported by facilities that do and do not consider such hospital visits as reasons for formal discharge. Additional stays in the same nursing home were found in one fourth of the cases in the NNHS. However, data on additional stays had high levels of missing and incomplete data. Forty percent of those cases required further editing. As noted, many cases were so complicated that the date fields needed to be evaluated and corrected manually. Editing was performed in such a manner that, when data on additional stays were incomplete, the editing tended to limit further redefinition of recorded admission and discharge dates. When data on additional stays were present, admission and discharge dates were redefined in 70 percent of the cases overall: 56 percent for the edited cases and 77 percent for the unedited cases. Redefinition tended to increase markedly the length of stay for the affected cases and this translates into a change in length of stay estimates for the entire sample. For the entire discharged resident file, the median length of stay based on the recorded endpoints was 135 days as compared to 213 days for redefined endpoints. In addition, the 1985 National Nursing Home Survey Followup will provide more information on this cohort. For many of the current residents as well as for discharged subjects who were still residents of the sample nursing home on the 1985 NNHS interview date, subsequent data collection waves will extend the length of stay and further increase the estimate of median length of stay.

References

1. Meiners MR, Trapnell GR. Long-term care insurance premium estimates for prototype policies. Medical Care 22: 901–911. 1984.
2. Liu K, Manton KG. The length of stay pattern of nursing home admissions. Medical Care 21: 1211–1222. 1983.
3. Spence DA, Wiener JM. Nursing home length of stay patterns: Results from the 1985 National Nursing Home Survey. The Gerontologist 30: 16–20. 1990.
4. Short PF, Cunningham P, Mueller C. Standardizing nursing-home admission dates for short-term hospital stays. Medical Care 29: 97–103. 1991.
5. Hing E, Sekscenski E, Strahan G. The National Nursing Home Survey: 1985 summary for the United States. National Center for Health Statistics. Vital Health Stat 13(97). 1989.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1992— 312-082/60002

New Electronic Data Product Releases

CDC
CENTERS FOR DISEASE CONTROL

National Center for Health Statistics

National Hospital Discharge Survey on Diskettes—1989

Diskettes containing data on hospitalization in the United States during 1989 have recently been produced by the National Center for Health Statistics (NCHS) and are available for sale from the National Technical Information Service. The diskettes, along with documentation, provide automated access to data on hospital utilization by age and sex of patient, geographic region of the United States, diagnosed conditions, and surgical and nonsurgical procedures performed. The data are from NCHS' National Hospital Discharge Survey (NHDS), an annual survey designed to obtain information from medical records of sampled inpatients.

The data diskettes include software for accessing the NHDS database fields, computing rates for selected populations, and creating smaller files for use with other software packages. The database files were developed by downloading text files similar to the detailed tables in Vital and Health Statistics, Series 13, "Detailed Diagnoses and Procedures for Patients Discharged From Short-Stay Hospitals, United States."

The technical characteristics and the software and hardware requirements for the diskettes are:

Technical Characteristics:

- First-listed Diagnoses
- Procedures
- Days of Care
- All-listed Diagnoses (available upon request)

Hardware and Software Requirements:

A. Hardware

- 640K RAM
- IBM compatible computer

B. Software

- PC/MS DOS release 2.0 or higher
- Multi-year data access system software (included)
- dBase III files

(Diskettes for data year 1989 can be purchased by completing the order form on the **back**. To order NHDS diskettes for data years 1985–88, contact the National Technical Information Service at the address below.)

Mail Order To:

NTIS
National Technical Information Service
U.S. DEPARTMENT OF COMMERCE
Springfield, VA 22161
(703) 487-4650-TELEX 89-9405

Purchaser:

Telephone_____

Name_____

Organization_____

Address_____

City, State, Zip_____

Attention:_____

Method of Payment

☐ Charge my NTIS deposit account no._____

☐ Purchase order no._____

☐ Check enclosed for $_____

☐ Charge my: American Express
Visa MasterCard Card expiration date_____
Account no._____
Signature_____
(Required to validate order)

☐ Send me an application for an NTIS Deposit account (NTIS-PR-33)

Ship to: (Enter if different from address at left)

Name_____

Organization_____

Address_____

City, State, Zip_____

Diskettes **1989 Data Files**	QUANTITY	UNIT PRICE	TOTAL PRICE
3½" PB 91-506881		$80.00	
3½" (HD) PB 91-506873		$80.00	
5¼" PB 91-506865		$80.00	
5¼" (HD) PB 91-506857		$80.00	
1989 All-listed Files	QUANTITY	UNIT PRICE	TOTAL PRICE
3½" (HD) PB 91-507111		$50.00	
5¼" PB 91-507087		$50.00	
5¼" (HD) PB 91-507079		$50.00	
Documentation PB 91-204362		$15.00	

TOTAL $ _____

Vital and Health Statistics series descriptions

SERIES 1. **Programs and Collection Procedures**—Reports describing the general programs of the National Center for Health Statistics and its offices and divisions and the data collection methods used. They also include definitions and other material necessary for understanding the data.

SERIES 2. **Data Evaluation and Methods Research**—Studies of new statistical methodology including experimental tests of new survey methods, studies of vital statistics collection methods, new analytical techniques, objective evaluations of reliability of collected data, and contributions to statistical theory. Studies also include comparison of U.S. methodology with those of other countries.

SERIES 3. **Analytical and Epidemiological Studies**—Reports presenting analytical or interpretive studies based on vital and health statistics, carrying the analysis further than the expository types of reports in the other series.

SERIES 4. **Documents and Committee Reports**—Final reports of major committees concerned with vital and health statistics and documents such as recommended model vital registration laws and revised birth and death certificates.

SERIES 5. **Comparative International Vital and Health Statistics Reports**—Analytical and descriptive reports comparing U.S. vital and health statistics with those of other countries.

SERIES 6. **Cognition and Survey Measurement**—Reports from the National Laboratory for Collaborative Research in Cognition and Survey Measurement using methods of cognitive science to design, evaluate, and test survey instruments.

SERIES 10. **Data From the National Health Interview Survey**—Statistics on illness, accidental injuries, disability, use of hospital, medical, dental, and other services, and other health-related topics, all based on data collected in the continuing national household interview survey.

SERIES 11. **Data From the National Health Examination Survey and the National Health and Nutrition Examination Survey**—Data from direct examination, testing, and measurement of national samples of the civilian noninstitutionalized population provide the basis for (1) estimates of the medically defined prevalence of specific diseases in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics and (2) analysis of relationships among the various measurements without reference to an explicit finite universe of persons.

SERIES 12. **Data From the Institutionalized Population Surveys**—Discontinued in 1975. Reports from these surveys are included in Series 13.

SERIES 13. **Data on Health Resources Utilization**—Statistics on the utilization of health manpower and facilities providing long-term care, ambulatory care, hospital care, and family planning services.

SERIES 14. **Data on Health Resources: Manpower and Facilities**—Statistics on the numbers, geographic distribution, and characteristics of health resources including physicians, dentists, nurses, other health occupations, hospitals, nursing homes, and outpatient facilities.

SERIES 15. **Data From Special Surveys**—Statistics on health and health-related topics collected in special surveys that are not a part of the continuing data systems of the National Center for Health Statistics.

SERIES 16. **Compilations of Advance Data From Vital and Health Statistics**—These reports provide early release of data from the National Center for Health Statistics' health and demographic surveys. Many of these releases are followed by detailed reports in the Vital and Heath Statistics Series.

SERIES 20. **Data on Mortality**—Various statistics on mortality other than as included in regular annual or monthly reports. Special analyses by cause of death, age, and other demographic variables; geographic and time series analyses; and statistics on characteristics of death not available from the vital records based on sample surveys of those records.

SERIES 21. **Data on Natality, Marriage, and Divorce**—Various statistics on natality, marriage, and divorce other than as included in regular annual or monthly reports. Special analyses by demographic variables; geographic and time series analyses; studies of fertility; and statistics on characteristics of births not available from the vital records based on sample surveys of those records.

SERIES 22. **Data From the National Mortality and Natality Surveys**—Discontinued in 1975. Reports from these sample surveys based on vital records are included in Series 20 and 21, respectively.

SERIES 23. **Data From the National Survey of Family Growth**—Statistics on fertility, family formation and dissolution, family planning, and related maternal and infant health topics derived from a periodic survey of a nationwide probability sample of women 15–44 years of age.

SERIES 24. **Compilations of Data on Natality, Mortality, Marriage, Divorce, and Induced Terminations of Pregnancy**—Advance reports of births, deaths, marriages, and divorces are based on final data from the National Vital Statistics System and are published annually as supplements to the Monthly Vital Statistics Report (MVSR). These reports are followed by the publication of detailed data in Vital Statistics of the United States annual volumes. Other reports including induced terminations of pregnancy issued periodically as supplements to the MVSR provide selected findings based on data from the National Vital Statistics System and may be followed by detailed reports in Vital and Health Statistics Series.

For answers to questions about this report or for a list of titles of reports published in these series, contact:

Scientific and Technical Information Branch
National Center for Health Statistics
Centers for Disease Control
Public Health Service
6525 Belcrest Road, Room 1064
Hyattsville, Md. 20782

301-436-8500

DEPARTMENT OF
HEALTH & HUMAN SERVICES

Public Health Service
Centers for Disease Control
National Center for Health Statistics
6525 Belcrest Road
Hyattsville, Maryland 20782

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

BULK RATE
POSTAGE & FEES PAID
PHS/NCHS
PERMIT NO. G-281

DHHS Publication No. (PHS) 92-1389, Series 2, No. 115