Series 2
No. 123

Vital and Health Statistics

From the CENTERS FOR DISEASE CONTROL AND PREVENTION / National Center for Health Statistics

Comparison of Two Surveys of Hospitalization: The National Hospital Discharge Survey and the NHANES I Epidemiologic Followup Study

September 1997

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

CENTERS FOR DISEASE CONTROL AND PREVENTION

Suggested citation

Russell LB, Milan E, and Jagannathan R. Comparison of two surveys of hospitalization: the National Hospital Discharge Survey and the NHANES I Epidemiologic Followup Study. Vital Health Stat 2(123). 1997.

Library of Congress Catalog Card Number

Russell, Louise B.
Comparison of two surveys of hospitalization : the NHDS and the NHEFS.
p. cm. — (Vital and health statistics. Series 2, Data evaluation and methods research ; no. 123) (DHHS publication ; no. (PHS) 97-1397)
Authors: Louise B. Russell, Edwin Milan, Radha Jagannathan.
''August 1997.''
Includes bibliographical references.
ISBN 0-8406-0530-7
1. Hospital utilization—United States—Statistics. 2. Medical care surveys—United States. I. Milan, Edwin. II. Jagannathan, Radha. III. National Hospital Discharge Survey (U.S.) IV. National Health and Nutrition Examination Survey (U.S.). Epidemiologic Followup Study. V. National Center for Health Statistics (U.S.) VI. Title. VII. Series. VIII. Series: DHHS publication ; no (PHS) 97-1397.
[DNLM: 1. Hospitalization—United States—statistics. 2. Length of Stay—United States—statistics. 3. Data Collection—United States. W2 A N148vb no. 123 1997]
RA409.5.U45 no. 123 RA981.A2
362.1'0723 s—dc21
[362.1'1'0973]
DNLM/DLC 97-36313
for Library of Congress CIP

For sale by the U.S. Government Printing Office
Superintendent of Documents
Mail Stop: SSOP
Washington, DC 20402-9328
Printed on acid-free paper.

Vital and Health Statistics

Comparison of Two Surveys of Hospitalization: The National Hospital Discharge Survey and the NHANES I Epidemiologic Followup Study

Series 2:
Data Evaluation and Methods Research
No. 123

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

Hyattsville, Maryland
September 1997
DHHS Publication No. (PHS) 97–1397

Contents

Detailed Tables

Figures

Abstract

Objectives

This report compares hospitalization data from the NHANES I Epidemiologic Followup Study (NHEFS) with data from the National Hospital Discharge Survey (NHDS), the benchmark for hospitalization in the United States, for men and women 35 years and older for the period 1971–87. The comparison is intended to help analysts evaluate the validity and generality of analyses based on the NHEFS.

Methods

Hospital stays per 1,000 population and average lengths of stay are compared year by year for each age-sex group and for the entire period. Regression analyses test for differences between the two surveys by age and sex, and for differences in trends over time and the effect of the Medicare program's prospective hospital payment system.

Results

Hospital stays per 1,000 population were lower in NHEFS than in NHDS in all age-sex groups at the beginning of the period, but the differences had almost disappeared by 1987. Lengths of stay, although somewhat longer in NHEFS, matched NHDS more closely. Differentials by age and sex were similar in the two surveys for both hospital stays per 1,000 population and length of hospital stay.

With its extensive information on baseline risk factors, the NHEFS offers a unique opportunity to study determinants of hospitalization in a representative sample of U.S. adults. The evaluation presented here suggests two points for researchers who want to use the NHEFS. First, including age as a control should largely correct for differences in age distribution between NHEFS and NHDS. Second, a time trend should also be included to capture the effects of several factors that caused the count of stays to be low in the early years of NHEFS followup.

Keywords: *hospitalization • health surveys • longitudinal studies*

Comparison of Two Surveys of Hospitalizations

by Louise B. Russell, Ph.D.; Edwin Milan, M.A.; and Radha Jagannathan, M.S., Rutgers University

Introduction

The National Hospital Discharge Survey (NHDS) has sampled hospital stays from non-Federal, short-stay hospitals annually since 1965. Designed to estimate annual hospital stays, hospital bed days, and average lengths of stay by age and sex, and to provide information on diagnoses and procedures, NHDS is the primary source of data on national hospitalization patterns (1). However, it has no information on patients beyond standard demographic characteristics collected on hospital admission forms (1,2). Because of the limited information on patient characteristics and because data are collected only for persons who use hospital care rather than for all persons at risk of hospitalization, the NHDS does not permit analyses of the determinants of hospital use.

The Epidemiologic Followup Study (NHEFS) of participants in the first National Health and Nutrition Examination Survey (NHANES I) has traced a sample of 14,407 adults representative of the noninstitutionalized U.S. population since they were first examined for NHANES I in 1971–75. Baseline data on each respondent include an array of potential risk factors measured by physician examinations, laboratory tests, and interviews. Followup surveys collected information on health-related outcomes, including overnight stays in health facilities. Thus, for the first time, it is possible to study the relationship between baseline risk factors and subsequent hospital use in a representative sample of adults.

Researchers have begun to use the hospitalization data for a variety of analyses. LaCroix and colleagues examined risk factors for pneumonia hospitalization and death (3). Reichman et al. studied the relationship between serum vitamin A and subsequent development of prostate cancer using hospitalization (or death) to measure incidence (4). Sichieri and colleagues studied risk factors for hospitalization with gallbladder disease among black people (5) and gallstone disease in all women (6). Miller and colleagues, including two authors of this report (Russell and Milan), related biomedical risk factors to subsequent hospital use, regardless of diagnosis, in adults 45 years and older (7). Thus, NHEFS has opened a field of study offering an array of possibilities that should be of increasing interest as the followup period lengthens.

This report compares patterns of hospitalization by age and sex in NHDS

We are grateful to Jeffrey L. Carson, Ellen J. Idler, Jane E. Miller, and William C. Taylor for helpful comments on this report. Walter Gonzalez prepared the tables and figures in computer format.
This project was supported by grant HS 07002 from the Agency for Health Care Policy and Research. Baseline data for the Health and Nutrition Examination Survey I (1971–75), the Revised Health Care Facility Stay File for the 1982–84 followup, and the Health Care Facility Stay File for the 1986 followup were provided by the Inter-university Consortium for Political and Social Research. The Health Care Facility Stay File for the 1987 followup was provided by the National Center for Health Statistics. All data were originally collected by the National Center for Health Statistics. Neither the original collector of the data nor the Consortium bear any responsibility for the analyses or interpretation presented in this publication.

with those in NHEFS for the period 1971–87. It is good research practice to compare a new data set with established data sets to help evaluate the validity and generality of results based on the new data set. Because this report examines all hospital stays, not just those for a particular diagnosis, it should be useful to other researchers for what it shows about the comparability of the two data sets.

Comparisons are made of annual stays per 1,000 population and average length of stay calculated from NHEFS with the same measures from NHDS for men and women in three age groups: 35–44 years, 45–64 years, and 65 years and older. The comparison covers the period from 1971 (the first year of NHANES I) through 1987 (the most recent year of NHEFS followup for which public use tapes were available at the time this analysis was conducted). The period begins 6 years after the start of NHDS and ends in the last year before a new methodology for data collection was introduced. This timing avoids the start-up years when the survey was being refined (1)and the changeover to the new methodology, which may affect trends (2).

Four questions are addressed in this report:

- How do hospital stays per 1,000 population and average lengths of stay compare in the two surveys?
- How does the year-to-year variability in these measures compare?
- Do the two surveys show similar trends over time?
- Do the surveys show similar differences among age and sex groups?

Background

NHDS is based on a stratified sample of approximately 500 non-Federal short-stay hospitals (2). Each year, data on several hundred thousand completed hospital stays are collected from the medical records of these hospitals. Because NHDS draws on medical records, dates of admission and discharge are almost always known, as are diagnosis and procedure codes. The hospital stays are weighted for sampling probability and nonresponse and adjusted to match known totals of hospital beds in each stratum and hospital stays in each sample hospital to arrive at national estimates of hospital stays and hospital days (1,2). Number of hospital stays and number of days in the hospital per 1,000 population are calculated using the civilian population, including institutionalized persons, as of July 1 of the survey year. Most of the NHDS data used for this report were taken from various issues of the *Statistical Abstract of the United States* (8). The National Center for Health Statistics supplied missing data.

NHEFS is based on a stratified sample of people—the 14,407 noninstitutionalized adults aged 25–74 at baseline (1971–75) who received a physical examination in NHANES I. These individuals were traced through followup interviews scheduled in 1982–84, 1986, 1987, and 1992 (data for 1992 were not publicly available at the time of this analysis). Almost 95 percent of the original cohort was successfully traced through 1987 (9).

During the followup surveys, information on overnight stays in health facilities was collected (9,10). Only respondents for whom an interview was completed were included in this process. Respondents (or proxies for deceased persons) were asked about overnight stays during the period covered by the interview—dates, reasons, and name and location of each facility. In special cases, where no proxy could be interviewed for a decedent, the death certificate was used to obtain information on hospital stays. With respondents' permission, facilities were contacted and asked to supply information from their records about the identified stays. The public use files classify each facility from which information was requested as a hospital, a nursing home, or ''unknown/out of the country.''

These methods of data collection distinguish three groups of hospital stays in NHEFS (7). In the first and largest group (54.4 percent of stays in the sample), stays were matched with abstracts, that is, they were identified by respondents and information was supplied by the hospitals. The second group (22.1 percent) consists of stays identified by the hospital, but not reported by the respondent. The third group (23.5 percent) is comprised of stays reported by respondents, but for which there were no hospital records. In some cases, the hospital could find no record of the stay; in others, the respondent refused permission to request information or the hospital did not respond to the request.

Hospital stays in the first and second groups, ''abstract-matched stays,'' are comparable to data from NHDS because the information comes from hospital records and is usually complete. Stays in the third group, ''report-only stays,'' present a problem. Exact dates of admission and discharge are not known, so length of stay is unknown. Even the year is often unknown because the respondent either supplied no dates or reported a range of years. This problem is discussed in the methods section.

Because both NHDS and NHEFS include stays in non-Federal short-stay hospitals, which accounted for 92 percent of all hospital stays nationally over the period (11), estimates from the two surveys should look similar. But the surveys differ in ways that preclude identical estimates. While some mark NHDS as the benchmark, others simply mean that each survey is best suited to different uses.

First, NHDS is based on a much larger sample of stays than NHEFS—several hundred thousand each year compared with 38,202 stays for all 14,407 NHEFS participants between baseline and 1987. Thus, estimates from NHDS will be more precise and less variable from year to year. Second, NHDS has complete information on virtually all stays, while more than one-fifth of NHEFS stays have missing information because respondents' reports could not be matched with abstracts.

Third, NHDS surveys a narrower range of hospitals than NHEFS does. NHDS includes only non-Federal short-stay hospitals, while NHEFS

includes all overnight stays regardless of facility type or ownership. Because NHEFS includes Federal as well as non-Federal hospitals, and also includes chronic disease, rehabilitation, and mental hospitals, it has the potential to provide a more complete picture of hospitalization. Finally, because it samples hospital stays, NHDS captures the experience of individuals admitted from institutions as well as those admitted from the community. NHEFS included only persons who were not institutionalized at baseline and who may have been healthier than average because they were able to travel to the examination site. Coverage of all hospitals will increase the number of hospital stays, while a healthier population will decrease them.

The term ''stays'' is used in this report to describe the data from both surveys, although there is a difference in definition between the surveys. NHDS collects data on stays completed during the year (discharges) and assigns them to the year of discharge. NHEFS stays were assigned to the year of admission for the analyses reported here.

Methods

Data

The sample used for this report was initially selected for research on the determinants of mortality and hospitalization. It consists of adults who were 25–74 years old at baseline; had valid information on age, race, sex, systolic blood pressure, and serum cholesterol; and were successfully traced through 1987. Of the original 14,407 people who participated in NHANES I, 13,324 were selected. Those dropped included 581 who were not traced and 502 who lacked data on blood pressure or cholesterol. For this report, another 391 persons were dropped: 370 traced persons who were not successfully interviewed after baseline and 21 who died before age 35 years, the lower age limit for our comparisons. (All others under 35 at baseline were included in the year they reached 35.) Thus, the NHEFS estimates in this report are based on 12,933 adults, 89.8 percent of the original cohort.

To create a record for each respondent of all hospital stays during the surveyed period, three NCHS-created files were used. The Revised Health Care Facility Stay File was used for the 1982–84 followup (the original file was revised to make it comparable with the files for later followups), and the Health Care Facility Stay Files were used for the 1986 and 1987 followups. Hospital stays were assigned to the year of admission to avoid dropping those for which the admission date was known, but the discharge date unknown. Each stay was assigned to an age group on the basis of the person's age at admission. As in NHDS, transfers were counted as separate stays. Of the 115 abstract-matched stays that were dropped, 6 were completely contained within another stay and were not coded as transfers. The remaining 109 were recorded as having zero length of stay and were dropped on the assumption that these individuals had died before arrival (such cases are also out of scope for NHDS). However, subsequent information showed that the stays had been less than one day and should not have been considered out of scope.

Many report-only stays could not be dated precisely because the respondent gave a range of years or no date, and 853 report-only stays were dropped because of problems with the date of admission: 696 had no date, 53 had dates preceding the respondent's baseline interview, and 104 had a range of dates entirely or mostly preceding the baseline interview. The remainder could only be dated in terms of whether they occurred between baseline and first followup in 1982–84 or between first followup and 1987. When the respondent reported a range that included 1982–84, the stay was assigned to the first period if most of the years in the range fell before 1983 and to the second if most fell after 1983.

After these adjustments, a total of 31,353 stays (82.1 percent of the 38,202 recorded for the original cohort) remained in the file—24,187 (77.1 percent) matched by abstracts and 7,166 (22.9 percent) reported only by the respondent. Between baseline and 1987, 9,241 respondents had at least one hospital stay. The remainder had no stays (3,507) or their only stays were coded ''facility unspecified or out of the country'' (42 respondents), ''dead on arrival'' (18 respondents), or were reported only by the respondent and either lacked dates or occurred during a period mostly or entirely preceding the baseline (125 respondents).

Average length of stay was calculated as the number of nights spent in the hospital divided by the number of stays. One respondent was apparently in the hospital at the time of the 1982–84 interview, and the date of discharge was not (as was usually done) entered at the next followup. That stay was retained for calculating stays per 1,000, but dropped from the calculations of average length of stay.

Hospital stays per 1,000 person-year equivalents and average length of stay were calculated separately for men and women in the three age groups available for the NHDS that best matched those constructed from the NHEFS: 35–44 years, 45–64 years, and 65 years and older. To account for the actual time each respondent was at risk for a hospital stay, the number of days the respondent was in the sample each year was calculated, adjusting for date of baseline or followup interview, or for death. Then the days for all persons in the appropriate age-sex group who were in the sample for at least part of the year were summed and divided by 365 (366 in leap years) to get person-year equivalents for that year. Thus NHDS and NHEFS are both reported as stays per 1,000 person-years. For simplicity, the rates are described as per 1,000 persons or population in the rest of this report.

NHANES I consisted of two nationally representative samples. The NHEFS data were weighted, using weights proposed by the National Center for Health Statistics, so that together, the two samples are representative of the national population (12). Although they are appropriate for the baseline period, these weights are less appropriate for later years when the national population's age-sex distribution was

different. They are not ideal weights for comparison with NHDS, but they are the best available for analyses of NHEFS. Table 1 compares the age distributions within the three age groups for the U.S. civilian population and the NHEFS.

As the table indicates, the aging of the NHEFS cohort prevents the creation of an exact match with NHDS for the oldest (65 years and older) and youngest (35–44 years) age groups. At the beginning of the comparison period, no one in NHEFS (with the exception of one person whose age was initially understated) was older than 74 years. It was not until 1987 that the NHEFS sample had aged enough to include persons 75 and older in proportion to their share in the national population. Also, the youngest members of the NHEFS sample turned 35 in the years 1981–85 (10 years after their baseline interviews). By 1987, the proportion of those in the youngest age group who were 35–39 was well below the national proportion.

Statistical Analyses

Hospital stays reported only by respondents, and not matched with medical abstracts, are presented in the next section, but were not included in the analyses used to compare the NHDS and the NHEFS. While including report-only stays gives more complete counts of the stays experienced by the NHEFS sample, trends cannot be examined and lengths of stay cannot be calculated for these stays.

To compare the NHDS and NHEFS, means and standard deviations were computed for each of the two 17-year series for each age-sex group. The correlations between the two series were also calculated. Then, again separately for each age-sex group, hospital stays per 1,000 persons and average length of stay from each series were regressed against time, to test for trend, and a categorical variable for the years 1983–87, to capture any effect of Medicare's prospective payment system. Finally, the age-sex groups were combined for each series, and stays per 1,000 and length of stay were regressed on time and categorical variables for age and sex, to test whether differences among groups were similar in NHDS and NHEFS.

Results: Comparison of Surveys

The results of the analyses address the four questions raised in the introduction.

Levels

Rate of Stays

When only abstract-matched stays are counted, hospital stays per 1,000 population in NHEFS are substantially below the levels in NHDS for all ages and both sexes (table 2, first two columns). The means calculated from NHEFS are about 60 percent as high as those in NHDS, except for men 45–64, for whom the ratio is 74 percent.

Report-only stays make up much of the difference between the NHEFS and NHDS estimates (table 3). Comparison of the two periods baseline to 1982 and 1983–87 show that with report-only stays included, stays per 1,000 persons in NHEFS equaled NHDS levels for men 45–64, but were somewhat lower for other groups. In the later period, abstract-matched stays were generally a higher percent of all stays, which may mean that more stays were successfully matched with hospital records.

Length of stay

The average length of stay is higher in NHEFS than in NHDS in all age-sex groups (table 4, first two columns). The differences are smaller for the older age groups. Within age groups, the differences are smaller for women than for men.

Annual Variability

Rate of Stays

Figures 1 and 2 show that for all age-sex groups, there is more year-to-year variability in NHEFS than in NHDS and that there is a tendency for NHEFS to approach NHDS over time. Standard deviations (table 2, column 3) support the picture of higher variability, except for the youngest groups where the secular decline in NHDS, not reflected in NHEFS, makes variability less in the NHEFS. Correlation coefficients (table 2, last column) show that the NHEFS and NHDS are moderately well correlated for men and women 65 years and older (0.69 and 0.77, respectively) and less well correlated for the other four groups.

Length of Stay

Although the NHEFS estimates of average length of stay are higher and their year-to-year variability is greater, they follow the trends in NHDS (figures 3 and 4). For men and women 65 years and older and for women 45–64, the high variability is limited to the early years and may reflect sampling variability, which diminished as more people (and thus more stays) were brought into the sample. For the other three groups (men 35–44 years old and 45–64 years old, and women 35–44 years old), variability remained high throughout the period. Standard deviations (table 4, third column) also show the higher variability of the NHEFS series. For the youngest age groups, there is effectively no correlation between NHDS and NHEFS. For the older groups, the correlations range from 0.45 for middle-aged women to 0.84 for women 65 and older, with both age groups of men between these extremes.

Trends

Rate of Stays

Figures 1 and 2 show that the NHEFS data approach NHDS data over time and thus necessarily have a different time trend. The differences in trend remain even when the baseline years, 1971–75, are omitted.

Regression results (table 5) support the visual impression. For the middle-aged and elderly groups, NHEFS shows much larger annual increases over the period than NHDS does. For men 35–44 years old, the annual

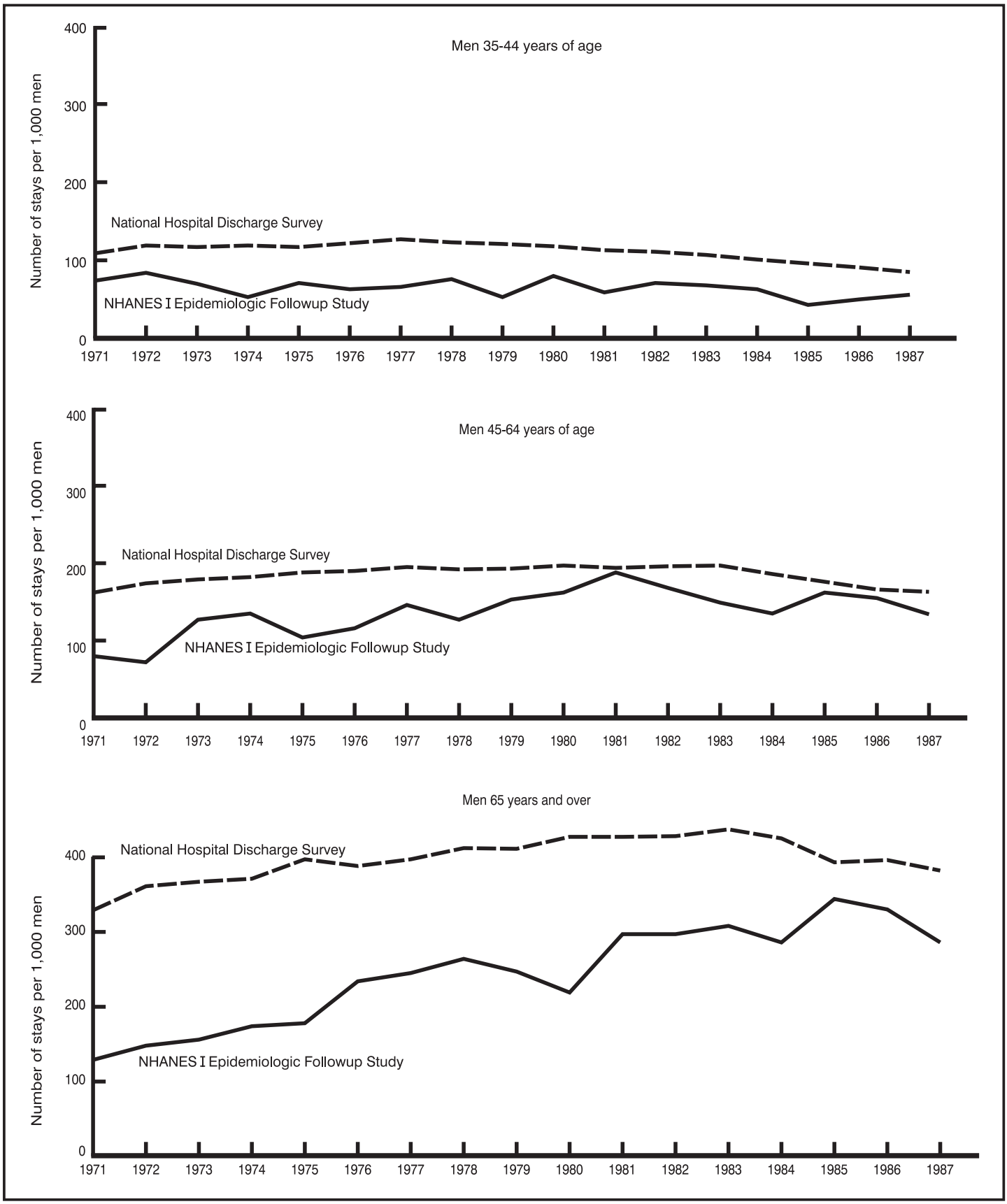

Figure 1. Number of hospital stays per 1,000 men 35 years and over by age group and year: National Hospital Discharge Survey and NHANES I Epidemiologic Followup Study, 1971–87

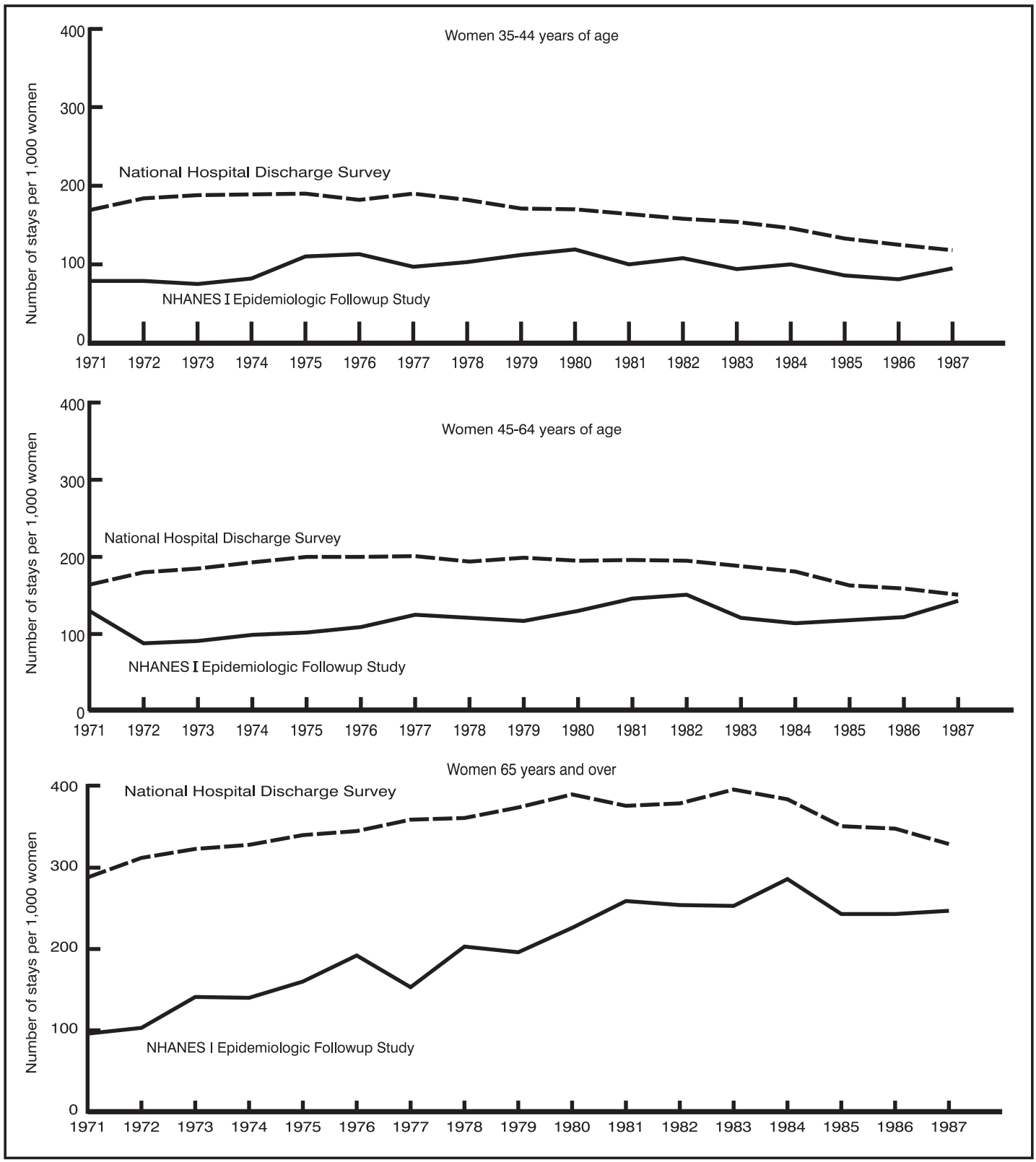

Figure 2. Number of hospital stays per 1,000 women 35 years and over by age group and year: National Hospital Discharge Survey and NHANES I Epidemiologic Followup Study, 1971–87

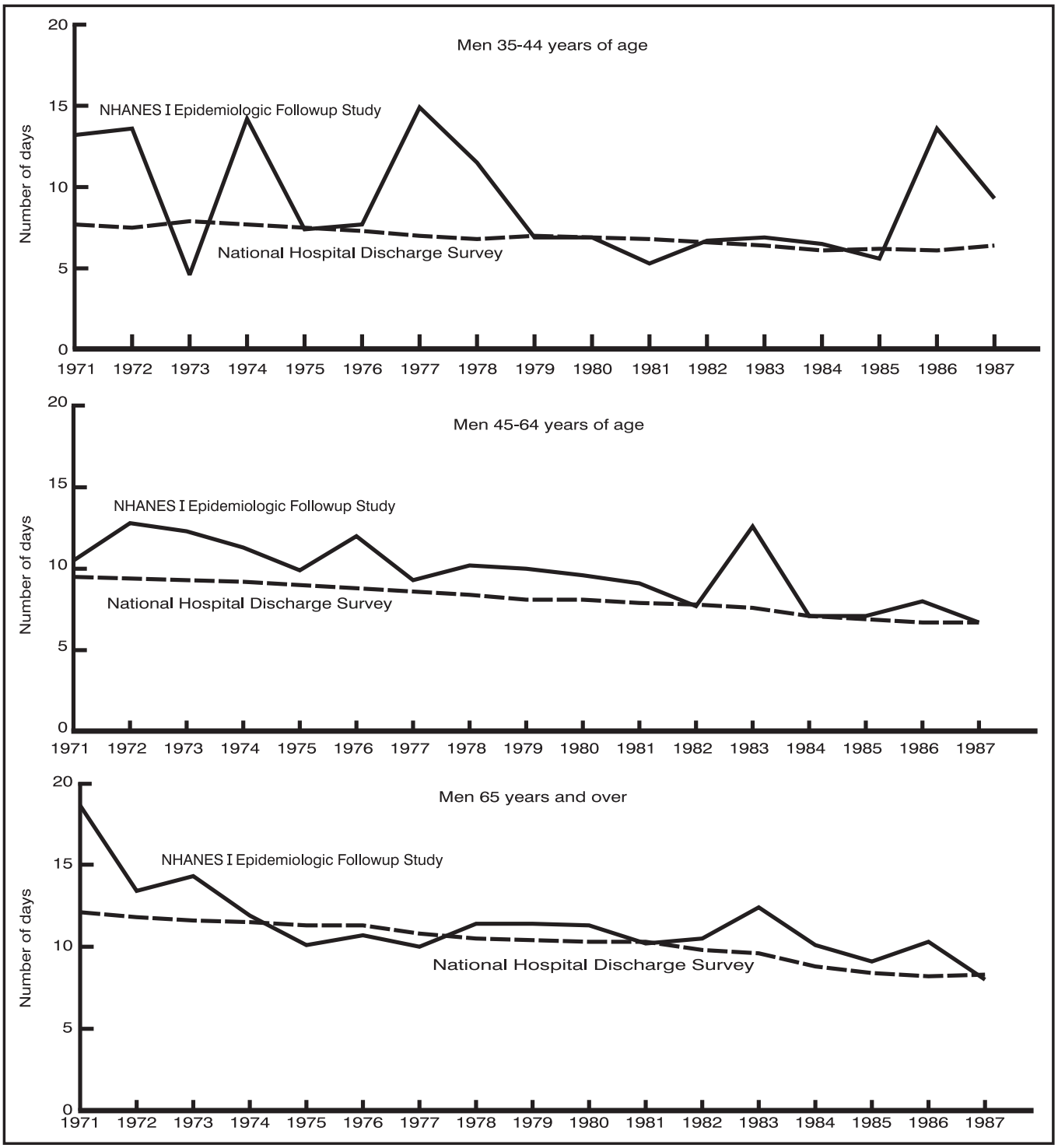

Figure 3. Average length of hospital stay for men 35 years and over by age group and year: National Hospital Discharge Survey and NHANES I Epidemiologic Followup Study, 1971–87

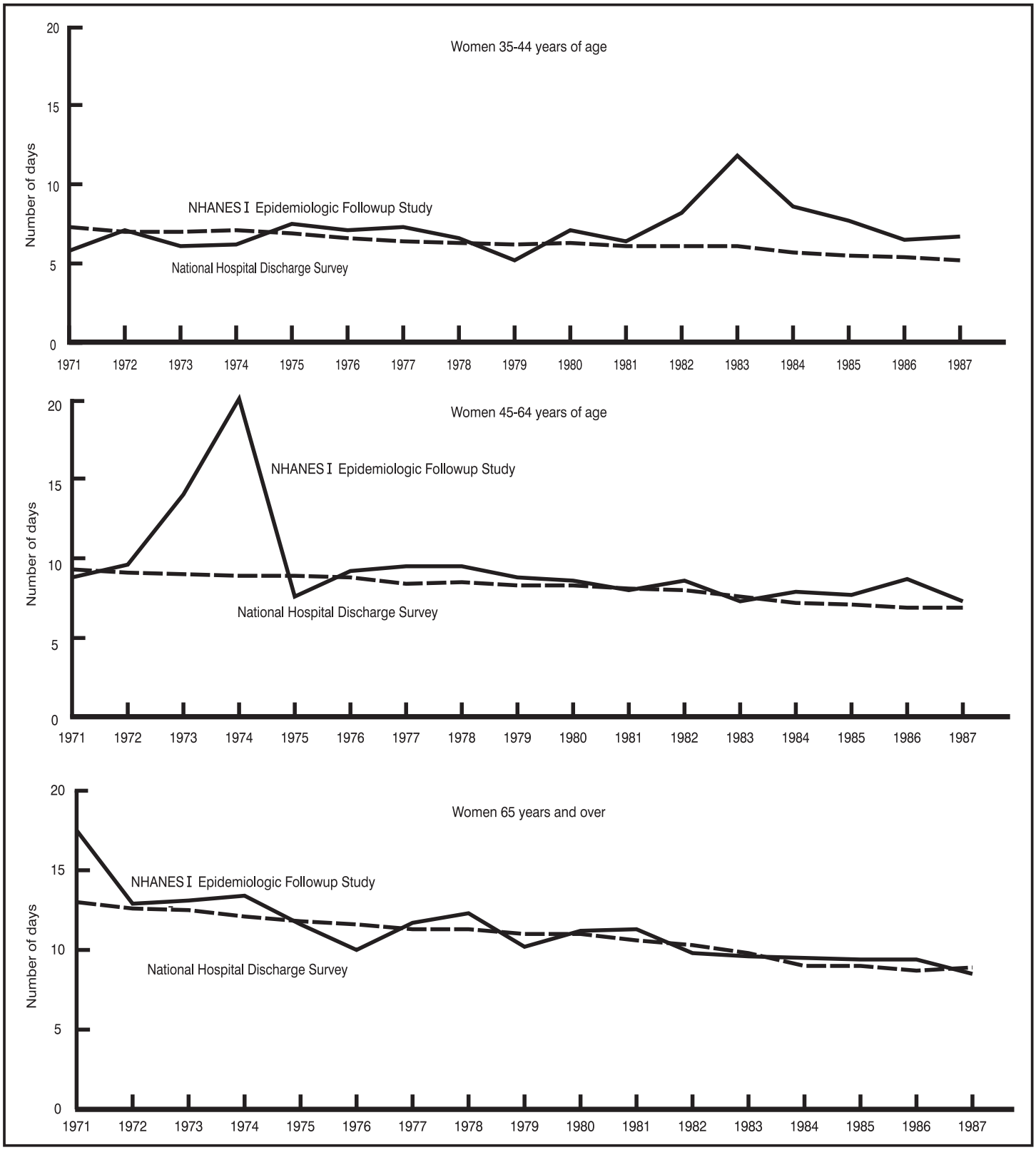

Figure 4. Average length of hospital stay for women 35 years and over by age group and year: National Hospital Discharge Survey and NHANES I Epidemiologic Followup Study, 1971–87

changes are insignificant for both surveys. For women in this age group, the trend for NHDS is negative, while that for NHEFS is positive.

There is better agreement on the impact of 1983–87. All these coefficients are negative, showing a decline in hospital stays per 1,000 in those years, and most are significant (P< 0.05).

Length of Stay

Figures 3 and 4 show that NHEFS generally follows the trend of NHDS, but at a higher level and with more variability. In accordance with this finding, regression results (table 6) show higher intercepts for the NHEFS and lower R-squareds. Although annual changes show that lengths of stay declined over the period in both NHEFS and NHDS, the average change is larger and less likely to be significant in NHEFS. NHDS also consistently shows a further decline in 1983–87, which is statistically significant for the older groups, but NHEFS does not.

Age-Sex Differentials

Rate of Stays

Age differences are as expected, and when age is included in the regressions, both NHDS and NHEFS show no statistically significant difference between the sexes in number of hospital stays per 1,000 population (table 7). Compared with people 45–64 years old, younger people have significantly fewer hospital stays per 1,000, and the estimated difference is almost identical—46 stays per 1,000 in NHDS and 47 stays per 1,000 in NHEFS. People 65 years and older have significantly more hospital stays, but the estimated difference is only half as large in NHEFS as in NHDS (94 compared with 190 stays per 1,000), reflecting the younger age distribution in this group in NHEFS over most of the period. The gap was smaller when the baseline years, 1971–75, were omitted. Regressions without time produced identical results.

Length of stay

Both surveys show broadly similar differences by age (table 7). Compared with middle-aged people, hospital stays are shorter for people 35–44 years old and longer for people 65 years and older. Both NDHS and NHEFS show that hospital stays average 1.5 days less for the youngest group. For the elderly population, hospital stays are 2.4 days longer than for the middle-aged population in NHDS and 1.7 days longer in NHEFS. The younger age distribution in NHEFS probably explains at least part of the difference. NHDS shows no difference between the sexes in length of stay when age is included in the regression analysis, but NHEFS shows a statistically insignificant difference of almost one day. Again, omitting time did not change the results.

Discussion

This report compares patterns of hospitalization in two national surveys. The National Hospital Discharge Survey (NHDS) samples stays in non-Federal, short-stay hospitals to estimate annual stays, bed days, and average length of stay, by age and sex. The NHANES I Epidemiologic Followup Study (NHEFS) has traced a nationally representative sample of adults 25–74 years old since 1971–75 to link baseline data on risk factors to subsequent health-related outcomes. These outcomes included overnight stays in all types of health facilities. The two surveys are conducted differently and for different purposes so they will not yield identical results for hospitalization. Nonetheless, it is helpful for researchers to understand where there are differences and what those differences suggest for methods of analysis.

By age-sex group, stays per 1,000 persons in the NHEFS were lower than stays per 1,000 in the NHDS. The gap was large in the beginning, but got steadily smaller over the period (figures 1 and 2). Lengths of stay, although somewhat higher in NHEFS, matched NHDS more closely. Differentials by age and sex were similar in the two surveys for both number of hospital stays per 1,000 and length of stay.

The analysis was based on hospital stays in the NHEFS that were matched with hospital abstracts, 77.1 percent of all stays in the sample. Although abstract-matched stays per 1,000 population were only about 60 percent of NHDS levels for the period, they were lowest in the early years and rose rapidly over the period until, by 1987, they were close to NHDS in all age-sex groups. When hospital stays reported by respondents but not confirmed by hospitals were included, and comparisons were made for two subperiods, NHEFS approached NHDS levels more closely. The ‘‘report-only’’ stays cannot be dated precisely enough to examine trends in more detail. If they could be, NHEFS levels might equal or exceed those of NHDS in the more recent years in all groups.

Several factors contribute to the differences between NHEFS and NHDS. Of primary importance is the different nature of the two samples. The NHDS samples discharges from non-Federal short-stay hospitals every year, thus capturing the hospital experience of the civilian population. The NHEFS is a longitudinal study of adults who were 25–74 years old when they were selected (1971–75). At baseline, the cohort was representative of the population 25–74 years old, but with the passage of time it no longer includes people at the younger ages and has started to include people 75 years and older. In addition, members of the original cohort have been lost to followup and no new members have been added to replace them or to represent in-migrants in the U.S. population. Thus the NHEFS cohort is not representative of adults 25–74 years old in any period other than baseline.

For persons aged 65 years and older, the different age distributions of NHEFS and NHDS help explain the more rapid rise in hospital stays per 1,000 persons in NHEFS over the period. At baseline, NHEFS included only people 65–74 years old, who have lower rates of hospital use than older people. Persons 75 years and older have become represented only as these

respondents aged. Until 1987, the proportion of the NHEFS elderly population 75 years and older did not match the U.S. elderly population. In subsequent followup years, the age distribution of NHEFS elderly persons should more closely match that of the nation's elderly population.

Loss to followup of some of the original NHEFS participants also affects the comparison. The analysis was based on 12,933 of the 14,407 adults in the NHEFS cohort. Most of the omitted respondents were lost to followup altogether or were traced but not interviewed after baseline. Those lost during the early years may have been disproportionately individuals who died, especially because the National Death Index was not yet available to facilitate tracing these individuals. Because hospital use is typically high just before death, loss of these individuals may have caused hospital stays per 1,000 people to be biased downward in the early years of followup. Loss to followup is likely to increase with the length of the followup period.

In addition, the weights applied to make the NHEFS sample representative of the national population may influence the comparability of the two surveys over time. These weights are appropriate for the years 1971–75, when the NHANES I was conducted. Applied to the full NHEFS cohort of 14,407 adults, they produce nationally representative data for those years. Although the age-sex distribution of the population has changed significantly since 1971–75, appropriate weights, adjusted for loss to followup, are not available for more recent years. Even if they were available, they could not compensate for the lack of sample persons at younger ages.

Differences in the health of the populations surveyed could also contribute to differences in number of hospital stays per 1,000 persons and lengths of stay. NHDS includes all persons with hospital stays while NHEFS participants, none of whom were in institutions at baseline and all of whom were capable of traveling to the examination site, were probably healthier on average than the general population. It would be reasonable to expect this advantage to diminish over time as the NHEFS sample regresses toward the mean. Such regression could be another explanation for the steeper rise in stays per 1,000 in the NHEFS compared with the NHDS.

Because comparison with the NHDS was not the primary purpose of the research for which the NHEFS sample used in this report was drawn, it did not determine the criteria for sample selection. The NHEFS sample selected omitted 502 individuals who lacked valid data on blood pressure or cholesterol. It is not known whether the hospitalization experience of these individuals differed from the experience of those retained for the sample. In addition, 109 hospital stays were dropped because they were listed as having zero lengths of stay, which was misinterpreted to mean that the individual was dead on arrival. Omission of these stays makes a small contribution to the lower levels of stays in the NHEFS.

Finally, hospital stays may have been undercounted in NHEFS in the earlier years because the time between baseline and first followup (approximately 10 years) was much longer than the periods between subsequent followups. Hospitalization data were collected by first asking respondents to recall their experience since the last interview. Because the first followup was not initially planned, participants were not encouraged to keep records and may have forgotten some hospital stays. In addition, the long delay may have made it impossible to locate hospitals or records for a larger proportion of stays than in later followups.

Much of the difference in hospitalization rates between NHDS and NHEFS is attributable to the ''report-only'' stays, which account for more than one-fifth of all stays recorded in the NHEFS. Some of these stays may be included in those recorded as having been reported only by the hospital. They would not have been matched with respondents' reports if, for example, the respondent's memory was faulty about the hospital in which the stay occurred. However, others are legitimate stays that could not be matched with hospital records. This is particularly clear in cases where the hospital had closed or the respondent refused permission to contact the hospital.

Thus, ''report-only'' hospital stays are an important source of information. Research on the determinants of hospitalization should include these stays whenever possible. They can readily be used in analyses concerned with whether a stay occurred over the followup period or with the number of stays that occurred. However, when the timing of stays is important for the analysis, they cannot be used because most of them lack dates—even the year of admission.

Because timing is important, most research using NHEFS has been, and much will continue to be, based on abstract-matched stays. The analyses for this report used only abstract-matched stays and suggest that they approximate total stays reasonably well. Although abstract-matched stays appear to undercount the total substantially in the early years of the period, they approach NHDS levels by 1987.

It is equally important that patterns by age and sex in NHEFS data correspond reasonably well with those in NHDS for the period as a whole. For the oldest groups (men and women 65 years and older), which have the largest numbers of hospital stays, the correlations between NHEFS and NHDS are high. This finding suggests that the two surveys are measuring the same phenomenon and inferences from one can apply to the population of the other. Differences by age and sex are similar in NHEFS and NHDS for both number of hospital stays per 1,000 population and average length of stay. The correspondence between the two surveys in age-sex differences suggests that differences by risk factors may also be similar. Thus, analyses of risk factors for hospitalization based on NHEFS are likely to be valid for the population as a whole.

Still, the analysis shows that differences between the two surveys may be important for some research topics. For example, NHDS shows a drop in both number of hospital stays per 1,000 and lengths of stay during the period after implementation of

Medicare's prospective payment system for hospitals (1983–87), but NHEFS does not. The factors promoting an increase in number of stays appear to have overwhelmed the effect of this period in NHEFS. Also, the two surveys are not as highly correlated for younger age groups as for the elderly population. The poor correspondence is due, in part, to the smaller numbers of hospital stays in these groups and the associated higher sampling variability. NHDS is based on a much larger sample of hospital stays than NHEFS, and its data collection methodology allows more completeness and consistency from year to year.

This analysis suggests two specific analytical points for researchers who plan to use hospitalization data from the NHEFS. First, analyses should include a time trend. The trend will probably capture the effect of several factors that caused the count of stays to be low in the early years of followup, and thus cannot be interpreted as a true secular trend. Second, including age as a control variable, which is standard practice in most studies, is essential for these data and should largely correct for differences in age distribution between NHEFS and NHDS.

The NHEFS offers a resource not available in NHDS or in any other survey: an array of risk factors measured at baseline for use in analyzing subsequent hospitalization in a representative sample of U.S. adults. The primary reason for collecting the hospitalization data and for attempting a complete count was to provide a means of validating self-reports of diagnoses. But the result is of much wider usefulness. Because of the size and nature of the sample and the data on baseline risk factors and other health outcomes for the same individuals, the NHEFS data on hospitalization offer an unusual opportunity for studying the determinants of hospital use. This report is intended to help researchers better understand the NHEFS hospitalization data.

References

1. Simmons WR, Schnack GA. Development of the design of the NCHS Hospital Discharge Survey. National Center for Health Statistics. Vital Health Stat 2(39). 1977.
2. Graves EJ. National Hospital Discharge Survey: annual summary, 1990. National Center for Health Statistics. Vital Health Stat 13(112). 1992.
3. LaCroix AZ, Lipson S, Miles TP, White L. Prospective study of pneumonia hospitalizations and mortality of U.S. older people: The role of chronic conditions, health behaviors, and nutritional status. Pub Health Rep 104(4): 350–60. 1989.
4. Reichman ME, Hayes RB, Ziegler RG, et al. Serum vitamin A and subsequent development of prostate cancer in the first National Health and Nutrition Examination Survey Epidemiologic Follow-up Study. Cancer Res 50(8): 2311–15. 1990.
5. Sichieri R, Everhart JE, Roth HP. Low incidence of hospitalization with gallbladder disease among blacks in the United States. Am J Epidemiology 131(5): 826–35. 1990.
6. Sichieri R, Everhart JE, Roth HP. Prospective study of hospitalization with gallstone disease among women: Role of dietary factors, fasting period, and dieting. Am J Public Health 81(7): 880–84. 1991.
7. Miller JE, Russell LB, Davis DM, et al. Biomedical risk factors for hospital admission in older adults. Med Care. In press.
8. U.S. Bureau of the Census. Statistical abstract of the United States. Washington. 1974–90.
9. Cox CS, Rothwell ST, Madans JH, et al. Plan and operation of the NHANES I Epidemiologic Followup Study, 1987. Vital Health Stat 1(27). 1992.
10. Cohen BB, Barbano HE, Cox CS, et al. Plan and operation of the NHANES I Epidemiologic Followup Study: 1982–84. Vital Health Stat 9(22). 1987.
11. American Hospital Association. Hospital statistics, 1989–90 edition. Chicago, Illinois: American Hospital Association, 1989: Table 1.
12. National Center for Health Statistics. Estimation of sample weights for all sample persons in the NHEFS. Memorandum dated November 3, 1989.

Table 1. Number and percent distribution of men and women 35 years and older, by year and age group, according to sex and population: United States and NHANES I Epidemiologic Followup Study, selected years, 1971–87

Year and age group	Men		Women	
	U.S. civilian population	NHEFS[1]	U.S. civilian population	NHEFS[1]
1971		Number in thousands		
35–44 years	10,889	591	11,725	692
		Percent distribution		
35–39 years	47.8	43.9	48.3	44.4
40–44 years	52.2	59.0	51.7	55.6
		Number in thousands		
45–64 years	20,226	1,007	22,191	1,229
		Percent distribution		
45–54 years	55.8	55.0	54.9	57.9
55–64 years	44.2	45.0	45.1	42.1
		Number in thousands		
65 years and over	8,559	339	12,002	467
		Percent distribution		
65–74 years	64.7	100.0	59.5	99.1
75 years and over	35.3	–	40.5	0.9
1975		Number in thousands		
35–44 years	10,897	8,677	11,629	9,671
		Percent distribution		
35–39 years	50.4	49.4	51.0	50.8
40–44 years	49.6	50.6	49.0	49.2
		Number in thousands		
45–64 years	20,823	17,852	22,935	19,572
		Percent distribution		
45–54 years	54.8	52.9	53.7	53.4
55–64 years	45.2	47.1	46.3	46.6
		Number in thousands		
65 years and over	9,265	5,877	13,431	7,980
		Percent distribution		
65–74 years	65.2	84.7	58.7	83.9
75 years and over	34.8	15.3	41.3	16.1
1980		Number in thousands		
35–44 years	12,448	9,738	13,144	10,817
		Percent distribution		
35–39 years	54.5	53.0	54.5	52.0
40–44 years	45.5	47.0	45.5	48.0
		Number in thousands		
45–64 years	21,137	18,684	23,339	20,810
		Percent distribution		
45–54 years	51.8	50.1	50.4	49.9
55–64 years	48.2	49.9	49.6	50.1
		Number in thousands		
65 years and over	10,367	8,434	15,338	11,639
		Percent distribution		
65–74 years	65.5	73.0	57.8	67.9
75 years and over	34.5	27.0	42.2	32.1

Table 1. Number and percent distribution of men and women 35 years and older, by year and age group, according to sex and population: United States and NHANES I Epidemiologic Followup Study, selected years, 1971–87—Con.

Year and age group	Men		Women	
	U.S. civilian population	NHEFS[1]	U.S. civilian population	NHEFS[1]
1985		Number in thousands		
35–44 years	15,378	8,742	16,124	8,594
		Percent distribution		
35–39 years	55.7	44.0	55.6	39.0
40–44 years	44.3	56.0	44.4	61.0
		Number in thousands		
45–64 years	21,425	17,058	23,464	19,729
		Percent distribution		
45–54 years	51.0	49.4	49.6	47.8
55–64 years	49.0	50.6	50.4	52.2
		Number in thousands		
65 years and over	11,536	9,979	17,000	14,004
		Percent distribution		
65–74 years	64.8	69.1	56.1	61.2
75 years and over	35.2	30.9	43.9	38.8
1987		Number in thousands		
35–44 years	16,646	4,035	17,380	3,881
		Percent distribution		
35–39 years	54.7	24.6	54.4	17.3
40–44 years	45.3	75.4	45.6	82.7
		Number in thousands		
45–64 years	21,642	10,119	23,619	11,175
		Percent distribution		
45–54 years	52.1	50.2	50.7	50.4
55–64 years	47.9	49.8	49.3	49.6
		Number in thousands		
65 years and over	12,115	5,327	17,715	7,965
		Percent distribution		
65–74 years	64.6	66.0	55.6	59.9
75 years and over	35.4	34.0	44.4	40.1

– Quantity zero.

[1]NHEFS is NHANES I Epidemiologic Followup Study.

NOTE: Numbers for NHEFS have been weighted to represent the U.S. population in 1971–75.

Table 2. Mean, ratio of means, standard deviation, and correlation of hospital stays per 1,000 population by sex, age, and survey: National Hospital Discharge Survey and NHANES I Epidemiologic Followup Study, 1971–87

Sex, age, and survey	Mean	NHEFS/NHDS	Standard deviation	Correlation
Men				
35–44 years:				
NHDS	112	0.58	12.0	0.50
NHEFS	65		11.2	
45–64 years:				
NHDS	184	0.74	12.1	0.47
NHEFS	136		30.4	
65 years and over:				
NHDS	397	0.61	28.9	0.69
NHEFS	243		66.7	
Women				
35–44 years:				
NHDS	165	0.59	23.4	0.12
NHEFS	96		13.8	
45–64 years:				
NHDS	185	0.65	16.2	–0.15
NHEFS	119		18.0	
65 years and over:				
NHDS	352	0.56	29.8	0.77
NHEFS	200		58.3	

NOTE: NHDS is the National Hospital Discharge Survey and NHEFS is the NHANES I Epidemiologic Followup Study.

Table 3. Number of hospital stays per 1,000 population by time period, sex, age, and source of report: NHANES I Epidemiologic Followup Study, 1971–87

Sex, age, and source of report	1971–82	1983–87
Men		
35–44 years:		
Reported by hospital	67	57
All stays	91	70
45–64 years:		
Reported by hospital	141	147
All stays	187	174
65 years and over:		
Reported by hospital	237	313
All stays	318	396
Women		
35–44 years:		
Reported by hospital	103	91
All stays	142	112
45–64 years:		
Reported by hospital	120	122
All stays	157	146
65 years and over		
Reported by hospital	198	255
All stays	264	309

Table 4. Mean, ratio of means, standard deviation, and correlation of length of hospital stay by sex, age, and survey: National Hospital Discharge Survey and NHANES I Epidemiologic Followup Study, 1971–87

Age, sex, and survey	Mean	NHEFS/NHDS	Standard deviation	Correlation
Men				
35–44 years:				
NHDS	6.9	1.31	0.6	0.21
NHEFS	9.1		3.5	
45–64 years:				
NHDS	8.2	1.19	1.0	0.76
NHEFS	9.8		2.0	
65 years and over:				
NHDS	10.3	1.11	1.3	0.67
NHEFS	11.4		2.4	
Women				
35–44 years:				
NHDS	6.3	1.16	0.6	–0.25
NHEFS	7.2		1.5	
45–64 years:				
NHDS	8.2	1.15	0.8	0.45
NHEFS	9.5		3.1	
65 years and over:				
NHDS	10.9	1.04	1.4	0.84
NHEFS	11.3		2.2	

NOTE: NHDS is the National Hospital Discharge Survey and NHEFS is the NHANES I Epidemiologic Followup Study.

Table 5. Results of regressions of number of hospital stays per 1,000 population on time and the years of Medicare's prospective payment system by age, sex, and survey: National Hospital Discharge Survey and NHANES I Epidemiologic Followup Study, 1971–87

Age, sex, and survey	Intercept[1]	Annual change[2]	Medicare PPS[3]	*R*-squared[4]
Men				
35–44 years:				
NHDS	[5]120	–0.33	[5]–19.22	0.75
NHEFS	[5]74	–0.80	–5.56	0.31
45–64 years:				
NHDS	[5]175	[5]1.88	[5]–25.23	0.36
NHEFS	[5]80	[5]7.93	[5]–51.89	0.71
65 years and over:				
NHDS	[5]349	[5]6.79	[5]–44.07	0.58
NHEFS	[5]125	[5]13.94	–23.31	0.87
Women				
35–44 years:				
NHDS	[5]193	[5]–2.35	[5]–22.91	0.84
NHEFS	[5]79	[5]2.93	[5]–31.74	0.49
45–64 years:				
NHDS	[5]184	1.23	[5]–33.88	0.52
NHEFS	[5]90	[5]4.19	[5]–29.51	0.55
65 years and over:				
NHDS	[5]305	[5]6.58	[5]–42.23	0.51
NHEFS	[5]92	[5]13.06	–33.54	0.87

NOTE: NHDS is the National Hospital Discharge Survey and NHEFS is the NHANES I Epidemiologic Followup Study.

[1]Intercept is estimated hospital stays per 1,000 population in 1970.

[2]Annual change is the estimated average change in the number of stays per 1,000 persons per year.

[3]Medicare PPS is the Medicare prospective payment system, which is represented by a categorical variable with a value of 1 for stays occurring in the per 1983–87 and a value of zero otherwise.

[4]R-squared is the proportion of the variation in stays per 1,000 population explained by the regression.

[5]$p < 0.05$

Table 6. Results of regressions of average length of hospital stay on time and the years of Medicare's prospective payment system by age, sex, and survey: National Hospital Discharge Survey and NHANES I Epidemiologic Followup Study, 1971–87

Age, sex, and survey	Intercept[1]	Annual change[2]	Medicare PPS[3]	*R*-squared[4]
Men				
35–44 years:				
NHDS	[5]7.89	[5]–0.102	–0.116	0.91
NHEFS	[5]12.03	–0.403	2.397	0.14
45–64 years:				
NHDS	[5]9.79	[5]–0.172	[5]–0.217	0.99
NHEFS	[5]12.79	[5]–0.370	1.050	0.57
65 years and over:				
NHDS	[5]12.31	[5]–0.205	[5]–0.573	0.98
NHEFS	[5]15.28	[5]–0.508	2.313	0.59
Women				
35–44 years:				
NHDS	[5]7.39	[5]–0.121	–0.003	0.95
NHEFS	[5]6.82	–0.016	1.682	0.25
45–64 years:				
NHDS	[5]9.40	[5]–0.118	[5]–0.488	0.99
NHEFS	[5]12.37	–0.336	0.452	0.24
65 years and over:				
NHDS	[5]13.07	[5]–0.228	[5]–0.576	0.99
NHEFS	[5]14.83	[5]–0.422	0.786	0.72

[1]Intercept is the estimated average length of stay in 1970.

[2]Annual change is the estimated average change in the average length of stay per year.

[3]Medicare PPS is the Medicare prospective payment system, which is represented by a categorical variable with a value of 1 for stays occurring in the per 1983–87 and a value of zero otherwise.

[4]R-squared is the proportion of the variaton in the average length of stay explained by the regression.

[5]$p < 0.05$

NOTE: NHDS is the National Hospital Discharge Survey and NHEFS is the NHANES I Epidemiologic Followup Study.

Table 7. Results of regressions of number of hospital stays per 1,000 population and length of hospital stay on time, sex, and age categories: National Hospital Discharge Survey and NHANES I Epidemiologic Followup Study, 1971–87

Hospital stays	Intercept[1]	Time[2]	Male[3]	Age 35	Age 65[5]
Stays per 1,000					
NHDS	[6]186.13	0.003	–3.255	[6]–46.029	[6]189.912
NHEFS	[6]80.28	[6]4.722	9.784	[6]–47.294	[6]94.000
Average length of stay					
NHDS	[6]9.81	[6]–0.182	0.020	[6]–1.567	[6]2.385
NHEFS	[6]11.36	[6]–0.236	0.790	[6]–1.488	[6]1.700

[1]Intercept is estimated stays per 1,000, or length of stay, in 1970 for women 45–64 years of age.

[2]Results for time show the average change in stays per 1,000, or length of stay, each year.

[3]Results for males show the difference in stays per 1,000, or length of stay, between men and women.

[4]Results for the 35–44 years age group show the difference in stays per 1,000, or length of stay, for persons 35–44 years of age compared with persons 45–64 years of age.

[5]Results for the 65 and older age group show the difference in stays per 1,000, or length of stay, for persons 65 years and over compared with persons 45–64 years of age.

[6]$p < .05$

NOTE: NDHS is the National Hospital Discharge Survey and NHEFS is the NHANES I Epidemiologic Followup Study.

Vital and Health Statistics series descriptions

SERIES 1. **Programs and Collection Procedures**—These reports describe the data collection programs of the National Center for Health Statistics. They include descriptions of the methods used to collect and process the data, definitions, and other material necessary for understanding the data.

SERIES 2. **Data Evaluation and Methods Research**—These reports are studies of new statistical methods and include analytical techniques, objective evaluations of reliability of collected data, and contributions to statistical theory. These studies also include experimental tests of new survey methods and comparisons of U.S. methodology with those of other countries.

SERIES 3. **Analytical and Epidemiological Studies**—These reports present analytical or interpretive studies based on vital and health statistics. These reports carry the analyses further than the expository types of reports in the other series.

SERIES 4. **Documents and Committee Reports**—These are 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. **International Vital and Health Statistics Reports**—These reports are analytical or descriptive reports that compare U.S. vital and health statistics with those of other countries or present other international data of relevance to the health statistics system of the United States.

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

SERIES 10. **Data From the National Health Interview Survey**—These reports contain statistics on illness; unintentional injuries; disability; use of hospital, medical, and other health services; and a wide range of special current health topics covering many aspects of health behaviors, health status, and health care utilization. They are based on data collected in a continuing national household interview survey.

SERIES 11. **Data From the National Health Examination Survey, the National Health and Nutrition Examination Surveys, and the Hispanic Health and Nutrition Examination Survey**—Data from direct examination, testing, and measurement on representative samples of the civilian noninstitutionalized population provide the basis for (1) medically defined total prevalence of specific diseases or conditions in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics, and (2) analyses of trends and relationships among various measurements and between survey periods.

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

SERIES 13. **Data From the National Health Care Survey**—These reports contain statistics on health resources and the public's use of health care resources including ambulatory, hospital, and long-term care services based on data collected directly from health care providers and provider records.

SERIES 14. **Data on Health Resources: Manpower and Facilities**—Discontinued in 1990. Reports on the numbers, geographic distribution, and characteristics of health resources are now included in Series 13.

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

SERIES 16. **Compilations of Advance Data From Vital and Health Statistics**—Advance Data Reports provide early release of information from the National Center for Health Statistics' health and demographic surveys. They are compiled in the order in which they are published. Some of these releases may be followed by detailed reports in Series 10–13.

SERIES 20. **Data on Mortality**—These reports contain statistics on mortality that are not included in regular, annual, or monthly reports. Special analyses by cause of death, age, other demographic variables, and geographic and trend analyses are included.

SERIES 21. **Data on Natality, Marriage, and Divorce**—These reports contain statistics on natality, marriage, and divorce that are not included in regular, annual, or monthly reports. Special analyses by health and demographic variables and geographic and trend analyses are included.

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

SERIES 23. **Data From the National Survey of Family Growth**—These reports contain statistics on factors that affect birth rates, including contraception, infertility, cohabitation, marriage, divorce, and remarriage; adoption; use of medical care for family planning and infertility; and related maternal and infant health topics. These statistics are based on national surveys of women of childbearing age.

SERIES 24. **Compilations of Data on Natality, Mortality, Marriage, Divorce, and Induced Terminations of Pregnancy**—These include advance reports of births, deaths, marriages, and divorces based on final data from the National Vital Statistics System that were published as supplements to the *Monthly Vital Statistics Report* (MVSR). These reports provide highlights and summaries of detailed data subsequently published in *Vital Statistics of the United States*. Other supplements to the MVSR published here provide selected findings based on final data from the National Vital Statistics System and may be followed by detailed reports in Series 20 or 21.

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

Data Dissemination Branch
National Center for Health Statistics
Centers for Disease Control and Prevention
6525 Belcrest Road, Room 1064
Hyattsville, MD 20782

(301) 436–8500
E-mail: nchsquery@cdc.gov
Internet: http://www.cdc.gov/nchswww/nchshome.htm

DEPARTMENT OF
HEALTH & HUMAN SERVICES

Centers for Disease Control and Prevention
National Center for Health Statistics
6525 Belcrest Road
Hyattsville, Maryland 20782

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

STANDARD MAIL (A)
POSTAGE & FEES PAID
PHS/NCHS
PERMIT NO. G-281

DHHS Publication No. (PHS) 97-1397, Series 2, No. 123
7-0565 (9/97)