# COMPARISON OF ESTIMATES FROM THE NATIONAL HOSPITAL DISCHARGE SURVEY USING TWO DIFFERENT SURVEY DESIGNS

Barbara J. Haupt, National Center for Health Statistics
6526 Belcrest Road, Room 950, Hyattsville, Maryland 20782

## INTRODUCTION

The National Hospital Discharge Survey (NHDS), one of a number of surveys conducted by the National Center for Health Statistics (NCHS), has been done continuously since 1965 and is the principal source of data on inpatient utilization of short-stay non-Federal hospitals in the United States. The NHDS provides nationally representative estimates on characteristics of patients, lengths of stay, diagnoses, and surgical and nonsurgical procedures in U.S. hospitals.

Beginning with the 1988 NHDS, a new sample of hospitals was selected and several data collection and estimation procedures were revised. Since many hospitals had begun to submit information in machine-readable form to commercial abstract services or state systems, one objective of the new design was to collect data via these sources. Another objective of the new design was to cluster the hospital sample in the same geographic areas used for other surveys conducted by NCHS, allowing analyses combining data on the use of hospitals with selected population and health resource utilization characteristics available from other NCHS surveys.

In order to evaluate the new survey design, data were collected and estimates obtained for the first 3 months (January - March) of 1988 using both the old and the new survey methods. This paper presents major findings of a comparison of these estimates. Because both sets of data were collected during the same period, differences in estimates based on the old and new survey methods should be primarily the result of sampling error, differences in the survey methods, or bias introduced over time as the original sample became less representative of the universe of hospitals.

The estimates compared are similar to those produced for the NHDS annual summaries: frequencies of discharges, days of care, average length of stay, all-listed diagnoses, and all-listed procedures by age, sex, and race of the patient discharged and by geographic region, bedsize, and ownership of hospitals.

Confidence intervals at the 95% level (1.96 standard errors) were computed for each estimate obtained using the old survey methods. The corresponding estimate from the new survey was then examined. Differences between the estimates were not statistically significant if the "new" estimate fell within the confidence interval of the "old" estimate. If the new estimate did not fall within the interval, a comparable confidence interval was computed for the new estimate. Differences between the two estimates were considered to be significantly different if the two confidence intervals did not overlap. No testswere done for estimates considered unreliable -- that is, estimates for which the relative standard error was 30% or greater or estimates that were based on fewer than 30 sample records.

## DESCRIPTION OF THE NHDS

NHDS collects data on patient discharges from non-Federal hospitals located in the 50 States and the District of Columbia. Only hospitals with 6 or more beds and an average length of stay of less than 30 days for all patients were included in the old sample survey.

The universe for the new sample survey includes, in addition, hospitals whose specialty is general (medical or surgical), maternity, or children's general, regardless of the average length of stay. Hospital units of institutions are excluded from both surveys.

DATA COLLECTION

Prior to 1985 all data were collected manually. Starting in 1985 2 data collection procedures were used for the survey: the traditional manual system and an automated system that involved purchasing data from abstract services. The automated method was used in less than 20% of the sample hospitals under the old survey design. Under the new survey design, almost 40% of the sample hospitals submitted 70% of all sampled records through abstract services during 1988.

In the manual procedure hospitals, sample discharges are selected using lists of discharges as the sampling frame. Data are then abstracted from the medical records. Completed forms are forwarded to NCHS for processing.

For the automated procedure hospitals, data tapes are purchased from abstracting services. The data, which have already been coded and edited according to each abstracting service's criteria, are sorted by NCHS. Sample discharges are selected and the data are edited and weighted by NCHS.

SAMPLE DESIGN

Both the old and the new designs for the NHDS sample involve multistage procedures in which the hospitals in the universe are placed into different strata and are arranged within the strata according to different types of hospital characteristics. A sample of hospitals is selected from this universe of hospitals. Patient discharges are then selected from the sample hospitals.

Old design - The original universe for the survey consisted of hospitals obtained from the 1963 National Master Facility Inventory (NMFI) (6). This universe was updated periodically using the NMFI and data from the American Hospital Association (AHA) to reflect the new hospitals entering the universe. The sample of discharges was selected using a stratified two-stage design.

The first stage of the sample design was the selection of hospitals from the universe. Hospitals with 1,000 beds or more were selected with certainty. Hospitals with less than 1,000 beds were stratified by size within region. Within each stratum, hospitals were distributed by type of ownership and geographic division. Sample hospitals were drawn with probabilities ranging from certainty for the largest to 1 in 40 for the smallest hospitals.

The second stage of the design involved selecting discharges from the sampled hospitals. The within-hospital sampling ratio for selecting these discharges varied inversely with the probability of hospital selection, so the overall probability of selecting a discharge was approximately the same in each size class. Details on the original design have been published (7).

New design - The universe for the redesigned NHDS consisted of hospitals listed in the April 1987 SMG Hospital Market Data Tape (9) that were accepting inpatients by August 1987. All hospitals with 1,000 beds or more or with 40,000 or more annual discharges were selected with certainty in the NHDS sample.

The non-certainty hospitals were selected using a stratified three-stage design. The first stage was the selection of 112 primary sampling units (PSU's). Hospitals within the selected PSU's were stratified based on geographic region, PSU size, and, for the 12 largest PSU's, abstracting status (whether or not

the hospital subscribed to an abstracting service). Within each of these strata the hospitals were ordered by PSU number, abstracting service status, and specialty-size groups. The second stage involved the random selection of hospitals from each stratum by probability proportional to size, with at least 3 hospitals being selected from each PSU (or all of the hospitals in PSU's with fewer than 3 hospitals).

At the third stage a sample of discharges from each hospital was randomly selected. The sampling rate was determined by the hospital's sampling stratum and method of data collection (manual or automated). The target sample size was 250 discharges from hospitals using the manual system of data collection. For hospitals whose data were collected through abstracting services, 250 discharges were sought from each hospital with less than 4,000 annual discharges and 2,000 discharges were selected from each of the remaining hospitals. The new NHDS design has been presented in more detail elsewhere (8).

## FINDINGS

Most of the estimates obtained using the new survey methods were not significantly different from those using the old methods. Although differences occurred in about 95% of the estimates, less than 10% of these differences were significant. In general, the new survey estimates of discharges, days of care, all-listed diagnoses, and all-listed procedures were lower than the comparable old survey estimates, while estimates of average length of stay more often were the same or higher using the new survey methodology than those obtained using the old survey methodology.

Significant differences were found between total estimates of discharges, days of care, all-listed diagnoses, or all-listed procedures for very few characteristics (Figure 1). Cross-tabulations contained a larger number of significant differences. This paper, however, is mainly concerned with differences between the "total" estimates.

## RACE

New survey estimates for discharges and for all-listed diagnoses were significantly lower than the comparable estimates obtained using the old survey methods for the category race not stated. Examination of cross-tabulations of the data revealed that this pattern was consistent for several estimates of discharges, days of care, all-listed diagnoses, and all-listed procedures. An analysis of the hospitals submitting data to the NHDS indicated that the method of data collection was a critical factor. Hospitals that submitted data through abstract services reported race not stated for 5.6% of their discharges, whereas hospitals that submitted data via the manual method reported race not stated for 10.8% of their discharges. The proportion of hospitals submitting data through abstracting services increased from 14% under the old survey design to 37% in the new survey design. More complete reporting of race from abstracting service hospitals thus resulted in a decrease in the estimates for the race not stated category using the new survey methods.

## GEOGRAPHIC REGION

The estimate of all-listed diagnoses obtained using the new survey methods for the West region was significantly lower than that obtained from the old methodology. A comparison of NHDS data with data from the AHA (1-5) was done in order to assess the quality of the NHDS regional estimates. AHA conducts an annual census of hospitals and publishes data each year for community hospitals; the hospitals in scope for the NHDS are similar to these community hospitals.

As shown in Figure 2, the 1975 NHDS estimate of discharges in the West was virtually identical to the number of admissions reported by AHA for 1975 but, over time, NHDS estimates diverged from the AHA numbers. The 1988 NHDS estimate (from the new survey) is almost identical to the AHA figure. This suggests that the hospitals from the West region in the old NHDS sample were becoming less representative of all short-stay non-Federal hospitals in that region; the redesigned sample seems to better represent these hospitals.

HOSPITAL OWNERSHIP

Estimates of discharges and days of care were significantly lower under the new NHDS design as compared with those obtained under the old NHDS design for government (non-Federal) hospitals. Several reasons for these differences were explored: the change in the universe for each sample, definitions of ownership categories in the 2 universes, and differences in response rates for government hospitals between survey methodologies. None of these provided an explanation for the differences in the estimates.

NHDS estimates for hospital ownership categories were then compared to utilization data from the AHA for ownership categories of community hospitals (1-5). The finding were generally similar to those seen in the comparison of NHDS and AHA regional data. These comparisons suggest that the old survey methods had been producing overestimates of the use of government hospitals and that the estimates obtained using the new methods may be more representative of the use of all short-stay non-Federal government hospitals.

HOSPITAL BEDSIZE

Estimates obtained from the new NHDS survey design were significantly higher than those obtained from the old survey design for discharges, days of care, all-listed

diagnoses, and all-listed procedures for hospitals with 100-199 beds. Conversely, new survey estimates were significantly lower than old survey estimates for discharges, all-listed diagnoses, and all-listed procedures for hospitals with 200-299 beds and for discharges, days of care, all-listed diagnoses, and all-listed procedures for hospitals with 500 or more beds. The changes in estimates for bedsize categories are not surprising since there are differences in the definition of hospitals between the universes used for the old and new survey designs. As stated previously, the universe for the old survey design was derived from NMFI and AHA data, but the new universe was based on the SMG Hospital Market Data Tape. Groups of hospitals under an umbrella organization are sometimes listed as a single hospital by AHA but are more likely to be listed as separate entities on the SMG data tape. For example, a medical center consisting of 3 separate facilities in different locations but under 1 umbrella organization could be listed as a single large entity by AHA but as 3 smaller entities by SMG. The result is that there are fewer large hospitals (and more small hospitals) in the universe for the new survey design than in the old universe.

The definition of beds used to construct the bedsize categories was also changed as part of the new methodology. In the old survey, bedsize was based on the total number of beds staffed and available for use. In the new design, the number of short-term beds was used, resulting in fewer hospitals in larger bedsize categories and more in smaller size categories.

Thus, although the estimates for hospital bedsize categories obtained using the new survey methods varied from the estimates based on the old methods (and from 1988 AHA data), they were generally similar to the bedsize data for the universe on

which they were based.

## DIAGNOSES

Significant differences were found between estimates using the 2 survey methodologies for 2 diagnostic categories. Estimates were higher for cataract but lower for alcohol dependence syndrome using the new survey methods.

The increases in the cataract estimates seem to be related to increases in the number of discharges with a diagnosis of cataract and a length of stay of less than 1 day. This may reflect more hospitals in the new sample formally admitting patients who would be treated in outpatient surgery programs at other hospitals. It is also possible that these hospitals reported data on outpatients to the NHDS.

The changes in the estimates of alcohol dependence syndrome seem to be influenced by a change in the participation of specialty hospitals in the NHDS. The old sample included 2 hospitals that specialize in the treatment of alcoholism and other chemical dependencies. These 2 hospitals accounted for 11% of all discharges with a first-listed diagnosis of alcohol dependence syndrome. Although the new sample also includes 2 alcoholism and other chemical dependency hospitals, neither of these hospitals participated in the survey during the first 3 months of 1988.

## SUMMARY

The methodology for the National Hospital Discharge Survey (NHDS) has been revised in several ways. These revisions, implemented in 1988, include use of a different hospital sampling frame, changes in survey design, and increased use of data purchased from abstracting service organizations. To investigate the effects of these revisions on NHDS estimates of hospital use, data were collected from January through March 1988 using both the old and the new survey methods. This paper compared estimates based on both survey methods for a variety of hospital and patient characteristics.

Few estimates were identical between the 2 survey methodologies, but most of the variations could be attributed to sampling error. Estimates from 2 different samples of the same population would be expected to vary by chance even if exactly the same methods were used to collect and process the data from the 2 samples. Where there were statistically significant differences in nonmedical data, the new methods appeared to produce more accurate estimates than the old methods. Race is more likely to be reported using the new methods. "New" estimates for hospitals in the West and government-owned hospitals are more similar than the corresponding "old" estimates compared to data from the census of hospitals conducted by the American Hospital Association. Numerous significant differences in estimates for bedsize categories between the 2 survey methodologies reflect changes in the universe and definitions for the new survey.

Few statistically significant differences were found in the medical data using the old and new survey methods. Two main differences, estimates for cataract and alcohol dependence syndrome, may have resulted from problems with the new survey. A measurement error, reporting outpatients to the NHDS, is one possible explanation of the higher estimates for cataract using the new survey methods. The decrease in estimates of alcohol dependence syndrome may be due to the nonresponse of hospitals that specialize in the treatment of alcoholics. In general, however, estimates of the utilization of short-stay non-Federal hospitals in the United States were not found to be greatly affected by the change from the old to the new survey methods.

REFERENCES

1. American Hospital Association. Hospital Statistics, 1976 edition. Chicago. 1976.

2. American Hospital Association. Hospital Statistics, 1981 edition. Chicago. 1981.

3. American Hospital Association. Hospital Statistics, 1986 edition. Chicago. 1986.

4. American Hospital Association. Hospital Statistics, 1988 edition. Chicago. 1988.

5. American Hospital Association. Hospital Statistics, 1989 edition. Chicago. 1989.

6. National Center for Health Statistics. Development and Maintenance of a National Inventory of Hospitals and Institutions.

National Center for Health Statistics. Vital Health Stat 1(3). 1965.

7. Simmons WR. Development of the Design of the NCHS Hospital Discharge Survey. National Center for Health Statistics. Vital Health Stat 2(39). 1970.

8. Graves EJ. National Hospital Discharge Survey: Annual Summary, 1988. National Center for Health Statistics. Vital Health Stat 13(106). 1991.

9. SMG Marketing Group, Inc. Hospital Market Data Base. Healthcare Information Specialists, 1342 N. LaSalle Drive, Chicago, ILL 06010. 1989.

FIGURE 1

CHARACTERISTICS FOR WHICH SIGNIFICANT DIFFERENCES OCCURRED
TOTAL ESTIMATES FROM OLD AND NEW SURVEY METHODOLOGIES
NATIONAL HOSPITAL DISCHARGE SURVEY

RACE

Not stated
  Discharges      ↓
  All-listed diagnoses    ↓

REGION

West
  All-listed diagnoses    ↓

HOSPITAL OWNERSHIP

Government
  Discharges      ↓
  Days of care      ↓

HOSPITAL BED SIZE

100-199 beds
  Discharges      ↑
  Days of care      ↑
  All-listed diagnoses    ↑
  All-listed procedures    ↑
200-299 beds
  Discharges      ↓
  All-listed diagnoses    ↓
  All-listed procedures    ↓
500 beds or more
  Discharges      ↓
  Days of care      ↓
  All-listed diagnoses    ↓
  All-listed procedures    ↓

DIAGNOSES

Alcohol dependence syndrome
  Discharges      ↓
  All-listed diagnoses    ↓

Cataract
  Discharges      ↑
  All-listed diagnoses    ↑

FIGURE 2

DISCHARGES, NATIONAL HOSPITAL DISCHARGE SURVEY
ADMISSIONS, AMERICAN HOSPITAL ASSOCIATION
WEST REGION, 1975-88

| YEAR | NHDS | AHA | NHDS-AHA |
|------|------|-----|----------|
| | Number in thousands | | |
| 1975 | 5,454 | 5,432 | 22 |
| 1980 | 6,103 | 5,769 | 334 |
| 1985 | 6,502 | 5,510 | 992 |
| 1987 | 6,678 | 5,369 | 1,309 |
| 1988 | 5,391 | 5,378 | 13 |

FIGURE 3

DISCHARGES, NATIONAL HOSPITAL DISCHARGE SURVEY
ADMISSIONS, AMERICAN HOSPITAL ASSOCIATION
GOVERNMENT (NON-FEDERAL) HOSPITALS, 1975-88

| YEAR | NHDS | AHA | NHDS-AHA |
|------|------|-----|----------|
| | Number in thousands | | |
| 1975 | 6,836 | 7,067 | - 231 |
| 1980 | 7,773 | 7,413 | 360 |
| 1985 | 7,776 | 6,028 | 1,748 |
| 1987 | 6,860 | 5,507 | 1,353 |
| 1988 | 4,894 | 5,424 | - 530 |