Comparison of the reliability of nationally representative estimates from the full (2007) and half (2008) sample National Hospital Discharge Surveys - Implications for Trend Analyses

Margaret J. Hall, Ph.D. and Maria F. Owings, Ph.D.

Centers for Disease Control and Prevention, National Center for Health Statistics, Division of Health Care Statistics, 3311 Toledo Road, Room 3331, Hyattsville, MD 20782/mhall@cdc.gov;mowings@cdc.gov

Introduction

The National Center for Health Statistics (NCHS) is part of the Centers for Disease Control and Prevention (CDC). The NCHS collects, analyzes and disseminates information to monitor health and health care use. Within NCHS's Division of Health Care Statistics, data are gathered on health care utilization at physician's offices as well as outpatient and emergency departments, ambulatory surgery facilities and inpatient care at hospitals. This paper discusses the decision to reduce the sample size of the National Hospital Discharge Survey (NHDS), and the consequences this decision had on estimates from the survey.

Overview of the National Hospital Discharge Survey (NHDS)

NHDS was conducted from 1965 through 2010, and was the longest, continuously-fielded annual survey of inpatient care in US hospitals. Its objective was to provide information on the utilization of the Nation's hospitals and the nature and treatment of illness among hospitalized patients. The original survey was conducted from 1965 to 1987, and utilized a 2-stage, stratified, cluster design. In 1988, the survey was redesigned and a new sample of hospitals was drawn using a modified three-stage, stratified, cluster design. The hospital sample was updated every three years to account for changes in the hospital universe over time. Beginning in 2011, a new National Hospital Care Survey was implemented and so trend data on hospital care will continue to be available.

The following types of hospitals were in scope for NHDS: short-stay hospitals (average length of stay < 30 days), as well as general (medical or surgical) and children's general hospitals regardless of average length of stay. Hospitals had to have at least 6 beds staffed for inpatient use, and they had to be located in the 50 states or the District of Columbia. The following types of hospitals were out-of-scope for NHDS: Federal, military, Veterans Administration hospitals and hospital units in institutions, such as prisons.

Beginning in 1988, NHDS had a modified three-stage probability design with the 1st stage comprised of Primary Sampling Units (PSUs) which were hospitals or geographic areas (e.g. counties, groups of counties, or metropolitan statistical areas). At the second stage hospitals were selected within area PSUs, using PPS sampling, that is, probabilities proportional to size, where size was measured by the hospitals' patient volume. The third stage involved selecting inpatient discharges within hospitals according to a systematic random sampling technique.

Data were collected for NHDS using two systems. One was a manual system in which sampling and abstraction of discharge records were performed at the sampled hospitals either by Census Bureau field staff or by hospital staff themselves. In 2007, this system was used for approximately 55 percent of the responding hospitals. An automated system was also used. Computerized files of discharge records were obtained from commercial abstracting services, states, or from the hospitals themselves. The automated system was used for the remaining 45 percent of responding hospitals.

Since NHDS gathers data from a sample, rather than the entire universe, estimates obtained from the survey have sampling error. The standard error (SE) is a measure of the sampling variability that occurs by chance when only a sample is surveyed. To derive standard errors for NHDS estimates, statistical software must be used that accounts for the complex survey design (e.g. SUDAAN, STATA, SAS). The relative standard error (RSE) of an estimate is obtained by dividing the standard error by the estimate itself. When multiplied by 100, the RSE is expressed as a percent of the estimate. For example, to calculate the relative standard error for pneumonia hospitalizations in 2007 the standard error of the estimate, 56,000, would be divided by the estimate of 1,056,000. The resulting .053 would

then be multiplied by 100 to convert this number to a percent. That would mean that the RSE for the pneumonia estimate was 5.3%.

In NHDS reports, estimates are not presented if they are based on fewer than 30 sampled cases or have a relative standard error (RSE) greater than 30 percent. If a researcher chooses to report an estimate that does not meet these reliability guidelines, it is recommended that this be noted in the presentation or publication of the results.

Decision to draw a half sample

Before 2008, data were collected from 400 to 500 hospitals, yielding approximately 300,000 inpatient discharges annually. Beginning in 2008, it was decided that the full sample could no longer be funded.

NCHS made the decision to halve the NHDS sample for the years 2008, 2009 and 2010 due primarily to budgetary constraints and competing priorities. It was estimated that the half sample would yield data from an average of about 200 hospitals and 150,000 inpatient discharges annually. It was understood that some loss of precision in the estimates would result from reducing the sample, but cost savings were anticipated. The resulting expected savings were to be used to develop the new National Hospital Care Survey which began data collection in 2011.

Implementing the half sample

Two major options for implementing the half sample were considered. In the first option, data would be gathered from the same number of hospitals, but from half the number of sampled records in each hospital. Savings from this option would result because the field staff would abstract fewer records and fewer data elements would have to be keyed and coded. This option would not appreciably affect the precision of estimates, since the same number of independent sampling units (hospitals) would be retained. The downside of this option was that all manual hospitals where field staff previously performed data collection would still require a visit. So while there would be a decrease in the costs of abstracting, keying and coding data, there would be no savings in travel costs to the hospitals.

The second option was to halve the number of hospitals, but gather the same number of records from the remaining hospitals. Cutting the number of hospitals meant a savings in travel costs, as well as in the costs of abstracting, keying, and coding the data. The tradeoff was that standard errors would be higher since the estimates would rely on fewer independent sampling units (hospitals). Because this option would free up more funds for the redesign of the survey, it was implemented.

In implementing the half sample, the goal was to adopt a very simple, efficient method of hospital selection using hospitals that were in-scope for the survey in 2006 as the “frame”. An attempt was made to preserve the region and hospital sampling strata of the full sample. The description of the general characteristics of the full (2007) and the half (2008) samples can be found in the following table.

	2007 Full Sample	2008 Half Sample
# of in-scope hospitals	477	238
# of hospitals providing data	422	207
# of discharges sampled	365,648	165,630
Unweighted response rate	88%	87%
Weighted response rate	82%	79%

Because a large number of NHDS estimates are stable from one year to the next, changing no more than about 10%, and typically not significantly different, we expected this to continue after the implementation of the half sample. Projections developed while planning for the half sample led us to expect that standard errors from the half sample would be larger, with RSEs perhaps as much as doubling from the prior year.

Unforeseen circumstances affecting the shift to the half sample

Critical programming staff who wrote and executed the NHDS weighting programs for many years left NCHS, just as we were transitioning to the half sample. In addition, the weighting programs in prior years were all written for, and executed on, the CDC mainframe which was being phased out. Thus, all weighting programs had to be adapted to run on PC's, validated, and then modified to reflect the shift to the half sample. These additional steps and the training of new staff led to delays in producing half sample data. In addition to losing time, the resources required to enact this shift also increased.

Federal employees', and on-site contractors', time were not taken into account in the calculation of the costs of this project. If they were, the considerable in-house resources needed to adapt the weighting programs for the half sample would have reduced the overall cost savings discussed below.

Effects of drawing a half sample

As was expected, due to the reduced sample size, error estimates for statistics produced by the half sample survey generally increased, and in some cases, the relative standard errors (RSEs) doubled. The following table illustrates the effects of the half sample on the pneumonia estimates. The difference between the 2007 and 2008 pneumonia estimates was not statistically significant.

	2007 Full Sample	2008 Half Sample
Estimated number of pneumonia inpatients	1,056,000	1,183,000
Standard error	56,000	140,000
Relative standard error	5%	12%

Since a decrease in sample size decreases the statistical power of a test, the effects must be larger in order to detect significant differences. Acceptance of the null hypothesis (i.e., that there is no statistically significant difference between estimates when in fact a true difference exists -- a Type II error) is more likely.

Differential impact of the half sample

Even in years when the NHDS had the full sample of records, there were some estimates which were more likely to be unreliable. As expected these less stable estimates became even more problematic with the half sample. Included among these were:

> Estimates of children's hospitalizations. Children are not hospitalized often, and when they are their hospitalizations tend to cluster in children's hospitals or in large, academic medical centers. This clustering led to higher standard errors for children's estimates. Another reason for less reliable estimates is the varying response rates of hospitals with a high volume of children's discharges. These historic problems resulted in additional unreliable estimates with the half sample.
>
> Hospitalizations for mental disorders. Psychiatric discharges tend to cluster in certain hospitals leading to high standard errors even when NHDS had a full sample. Moving to a half sample exacerbated these problems leading to fewer reliable estimates.
>
> Estimates for the West Census region. In the West, both sampling issues and patterns of nonresponse, created estimation problems for many categories for the half sample years.

Cost savings

Although the NHDS sample size was cut in half, the cost of the survey did not decrease by half. Prior to the implementation of the half sample, NHDS data collection and processing costs were over $1.5 million dollars per year – including costs for data abstraction as well as data keying and coding. The half sample resulted in a 33% reduction of costs. As was noted above, these costs did not count the salaries of federal employees and on-site contractors who increased their time on this project, in order to implement the half sample.

The reasons why half of the budget was not saved when the sample of hospitals and the previous number of records were reduced by half have to do with the methods by which the NHDS data were gathered in the past. That is, many hospitals (45%) had provided NHDS data electronically. For these hospitals there were no cost savings, since there had been no travel to these hospitals, and the data provided to NHDS were already abstracted, keyed, and coded. The remaining NHDS hospitals that provided manual data numbered about 240 in 2007. Halving that number meant that the costs of travel were eliminated for only 120 hospitals. But the costs of abstraction were not cut in half because some of these hospitals had abstracted data for the survey at either no, or minimal, cost to NHDS.

The cost savings, although less than expected, were nevertheless instrumental in funding the design of the new hospital care survey which began data collection in 2011. This redesign was long overdue and may not have happened without the resources saved by switching to the half sample.

Pluses resulting from half sample implementation

On the positive side, most of the demographic estimates, e.g. age, sex, and region of hospitalized patients continued to have reliable estimates and relatively low standard errors for the years of the half sample. Also, the vast majority of specific diagnostic and procedure categories annually reported in NHDS publications and data releases were still reliable and could be published. The days of care and lengths of stay for these categories were also mostly reliable.

Implications

Using estimates within large categories like the age category, sex, and region remains possible because standard errors, although increased, are still relatively low. Most common diagnoses and procedures crossed with demographic categories still yield reliable estimates. However, researchers accustomed to using data based on the full sample must take additional care in selecting research topics for years of the half sample. It is not advisable to assume that analyses can be supported by one year of data for the half sample years – 2008, 2009, and 2010.

Even in NHDS full sample years, preliminary steps have been an important part of the research process. But they become even more important when data are based only on the half sample. Before proceeding with their analysis, researchers should check sample sizes for estimates, and check published SE's, and calculate approximate errors based on formulas in the documentation.

As stated above, it is inadvisable to assume that an estimate (e.g. a particular diagnosis or procedure) that was reliable in previous years is still reliable. It is best to examine how the reliability of estimates was affected for the subcategories of interest. Application to the NCHS or CDC Research Data Center may be necessary to access data that is needed to obtain specific rather than generalized error estimates not included in published or posted tables.

If reliability problems occur when analyzing one year of data, data for multiple years can be combined. Since this increases the sample size it lowers the standard error. For example, an analysis of trend data could compare 2000-2001 to 2004-2005 and to 2008-2009. This technique has been used frequently for NCHS surveys with smaller sample sizes.

Data for the third year (2010) of the half sample, and the final year of the NHDS, are expected to be released to the public in March 2012.

New directions

The process of redesigning NHDS has been completed and the National Hospital Care Survey is underway. Data for 2011 are being gathered from hospitals. Instead of abstracting inpatient data from medical records, this survey is using claims data. Future plans are to add data from outpatient and emergency departments, and from ambulatory hospital-based and freestanding surgery facilities, to this new hospital care survey.