

Understanding and Interpreting the National Hospital Ambulatory Medical Care Survey (NHAMCS): Key Questions and Answers

Editor's note:

The National Hospital Ambulatory Medical Care Survey (NHAMCS) is widely used for medical research. Nearly 500 articles have been published based on this database, including 28 in *Annals of Emergency Medicine*. NHAMCS is a national probability sample survey of visits to emergency and outpatient departments in nonfederal, general, and short-stay hospitals conducted by the Centers for Disease Control and Prevention, National Center for Health Statistics (<http://www.cdc.gov/nchs/ahcd.htm>). Strengths of the NHAMCS survey are its rigorous methodology, nationally representative nature, large size, wide array of variables, and capacity to examine long-term trends. Investigators can freely download the database and test locally developed hypotheses. Challenges with NHAMCS are that, given its logistical and statistical complexity, resulting research can be difficult for readers to interpret and for editors and reviewers to critique. There are important limitations and caveats to this survey that, if overlooked, could lead to misleading or inaccurate conclusions.

In this article we have posed a series of questions to two NHAMCS statisticians, and we believe that their answers will be invaluable to both NHAMCS investigators and readers of its research.

ACCESSING AND MANIPULATING THE DATABASE

Q. What are the first steps in using the NHAMCS data?

The first step in using NHAMCS data is to understand what the data represent. NHAMCS is a records-based or encounter level survey, producing annual estimates of the number and attributes of visits to hospital emergency departments (EDs) in the U.S. Because the survey is visit based rather than population based, estimates of persons cannot be obtained and, therefore, incidence or prevalence rates of health conditions in the population cannot be calculated. Public use NHAMCS data cannot be used for state level analysis, although estimates can be made for four regions of the country. Because NHAMCS is a national probability sample survey and not a count of ED visits in the U.S., the estimation process involves adjustments for both survey and item nonresponse and makes several ratio adjustments within and across hospitals. Each record can represent many thousands of visits. Applying the visit sampling weight is a critical factor in deriving the survey estimates.

Additional variables (e.g., Federal Information Processing Standard state and county codes, emergency service area and hospital type, annual ED visit volume, teaching hospital, medical school affiliation, trauma level rating) are only available using the NHAMCS restricted data set available through the National Center for Health Statistics (NCHS) Research Data Center: <http://www.cdc.gov/rdc/>. In addition, better variance estimation can be attained by using

restricted data. Researchers must submit a proposal and pay a set-up fee of \$750 per day. There are four modes of access to hosted data: National Center for Health Statistics Research Data Center, Census Bureau Research Data Center, remote access, and staff-assisted access.

Whether you use the restricted files or public use files, it is important to read the public use file documentation and the documentation updates on the DC's National Center for Health Statistics (NCHS) web site: http://www.cdc.gov/nchs/ahcd/documentation_updates.htm.

Take advantage of the SAS, Stata, and SPSS input statements, variable and value labels, and format assignments that are available http://www.cdc.gov/nchs/ahcd/ahcd_questionnaires.htm#documentation.

A checklist to guide data analysis and manuscript preparation is shown in the FIGURE.

Q. The typical research sequence is for an investigator to pose a question and then design a study to answer it. From an editorial perspective, we perceive that at times investigators reverse the sequence, instead wondering what questions they can ask based on what's in the database. Do you see this as a problem and do you have any guidance for investigators, editors, or clinicians in this regard?

A. We don't see this as a problem. If an investigator has a research question, then specific hypotheses can be formulated based on a review of NHAMCS data. Further analysis can address whether the data support the hypothesis or not. As always the case, replicating the analysis using other data sets or data years is good research practice. On a related issue, occasionally researchers submit a manuscript thinking that their question can be answered by NHAMCS data when it clearly cannot be. That's why it is important to understand the limitations of the survey and read the documentation as suggested in the answer to the previous question.

Q. NHAMCS data can be difficult to retrieve and use. The raw files are not delimited and do not open automatically into statistical or database software. Manual variable selection and extraction from these files is tedious and error-prone. Combining multiple years of data is difficult and variables are coded differently from year to year. What plans are there for making the data easier to access and use?

A. Similar to other CDC National Center for Health Statistics' (NCHS) datasets (e.g., the National Health Interview Survey), NHAMCS public use files are provided in ASCII text format that allow the data to be used by many software products. NCHS provides sample SAS statements for use with NHAMCS data files from 1993 forward, which should make the process of using the data fairly straightforward. SPSS and Stata code are provided for some years, and efforts are underway to make them available for more years as staff resources permit. It is true that the survey undergoes periodic redesign, which can make the process of combining multiple years of data challenging. NCHS staff make every effort to point out important changes in the annual public use file documentation, both in the summary of changes at the beginning of the document and in the record format section. Staff members are also available to answer technical questions by calling 301-458-4600.

Q. Given that a frequent use of NHAMCS is simple summary data about discrete diagnoses, have you given consideration to having an interactive website version of the data

using a query page? Users could insert any ICD-9-CM diagnosis code or variable, the years of interest, and the summary data are printed in a table and/or graph, and then you can plug in stratifying variables, and it gives you graphs by age or geographic region, etc.

A. NCHS is currently evaluating a web-based tool that is specifically designed for a general user who has no knowledge of statistical software packages, and for any user who does not require in-depth statistical analysis. The tool will provide the public with the opportunity to perform data queries of analyses not typically included in standard NCHS reports. The ease, flexibility, and expediency of the tool will provide users with a range of functionalities, which include stratification of the data by: diagnosis group (International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes), expected primary source of payment, disposition, and geographic area (i.e., region and Metropolitan Statistical Area). This tool has the capability to perform a variety of analyses (i.e., crude rate, age-adjusted rate, etc.) and filter data by a number of demographic characteristics such as age, sex, race, and ethnicity. Users of the tool will be able to customize their query to obtain simple graphics and tables, print and save chart results, and suppress data for reasons of statistical reliability. At this time, no release date has been set.

ANALYZING THE DATABASE

Q. Does the sample include children's hospitals, teaching hospitals, and/or tertiary care centers, or is it all community emergency departments?

A. The sample includes nonfederal, general, and short-stay hospitals (< 30 days); therefore, it does include children's, teaching, and tertiary care hospitals as long as they meet the length of stay requirement.

Q. Who completes the original data forms at each hospital? Are there any minimum qualifications for these abstractors? NHAMCS has a high level of quality control once it receives data forms, but what quality control exists before the forms are given to the National Center for Health Statistics?

A. Approximately half of the time ED medical records are abstracted by local hospital staff onto a Patient Record form and the other half are completed by Census Bureau field representatives. In the former case, a clerk in the medical records department or less frequently a nurse or other ED staff member abstracts the record. NHAMCS permits hospitals to select personnel for this task and does not stipulate their minimum credentials. The Census Bureau field representatives train hospital staff on how to complete the Patient Record form and they are provided with an instruction booklet that contains definitions of the data items.

Census Bureau field representatives must have a high school diploma and either 6-12 months of general clerical experience or 1-2 years of college. Most of our NHAMCS field representatives have had experience working on other Census surveys or the decennial census. When first assigned to NHAMCS, they receive 2 days of classroom training which includes abstracting data from mock medical records. They then must pass written and oral exams. A senior field representative oversees their first hospital induction interview and their first chart abstraction. Subsequent annual training focuses on new survey procedures and new items on both the induction and Patient Record forms, as well as reminders about particular problems related to sampling and abstraction. It also includes a self-study exam and abstraction exercises. The field

representative manual contains a chapter on abstraction and each field representative is provided with a medical abbreviations book.

Budget constraints have precluded reabstraction studies to assess the repeatability of the chart review process, although one was conducted in the summer of 2012.

Q. For readers unfamiliar with NHAMCS this database is somewhat of a mysterious “black box”. Most authors do not report a methodological level of detail similar to what might be noted in a standard trial. What should ideally be reported in each study?

A. Authors may use a template for the Methods section specific for NHAMCS which may be found in Appendix E1. References are also provided.

Q. Is it possible to link individual ED data from one year to the next?

A. The initial NHAMCS sample of hospitals was selected in 1991 and it is updated every 3 years. Each included hospital is randomly assigned to one of 16 4-week reporting periods; therefore, the same hospital may not be in the public use data file in two consecutive years. Within the ED, there may be different administrative units (e.g., general, pediatric, fast-track) that in the NHAMCS are referred to as emergency service areas. When multiple emergency service areas exist within a single ED, the sampling includes all such areas. The hospital identifier is masked and does not carry over, either between EDs and outpatient departments of the same hospital in a given year or for hospitals across years; therefore, it is not possible to link ED data from one year to the next with NHAMCS public use files. This is possible, however, using restricted data available through the NCHS Research Data Center (RDC)

<http://www.cdc.gov/rdc/>.

Q. A frequent NHAMCS research format is that of temporal comparisons, i.e., does the frequency of a specific diagnosis, test, or scenario change from 2000 to 2009? Are there any particular caveats or concerns regarding applying NHAMCS in this fashion?

A. Since the inception of NHAMCS in 1992, various items were collected for a period of years, then dropped for a time, and then reinstated. When examining trends and/or combining multiple years of data, it is important to check that the variables of interest were collected in all of the years contained in your analysis and that the wording of the item and/or answer categories did not change in a way that might impact the analysis. You may view the ED Patient Record forms at http://www.cdc.gov/nchs/ahcd/ahcd_survey_instruments.htm#nhamcs.

There are caveats to be aware of when performing trend analyses. For example, the number of drugs that could be entered increased from 5 in 1994 to 6 in 1995 and to 8 in 2003. If an increase is observed in the number of medications prescribed or the prescribing rate, this could be due to the change in the way drugs were recorded and not an actual increase in prescribing. Therefore, a trend in prescribing from 2000-2009 should only include the 6 first-listed drugs in all years.

Slight changes in the diagnosis, external cause of injury, and procedure codes in the ICD-9-CM may occur from year to year.

Also, when combining multiple years of data, it is easier for the reader to comprehend an average annual estimate rather than the total number of visits that occurred during the study period. For example, it is preferable to state that from 1998-2007, the average annual number of ED visits for condition X was 3 million rather than from 1998-2007, there were 30 million ED visits for condition X.

Another pitfall to be avoided is that in an attempt to improve data quality, NCHS occasionally recodes variables. From 1997-2004, “primary expected source of payment” (PAYTYPE) was collected and then in 2005 this item was modified to capture multiple sources of payment. The variable PAYTYPE was created to be consistent with previous years and it used a hierarchical scheme to recode the multiple source data into primary source. From 2005-2007, Medicaid was at the top of the hierarchy followed by Medicare. It was later decided that Medicare should be at the top to be more consistent with insurance industry practices. It is planned that 2005-2007 NHAMCS public use files will be re-released to reflect this change. In the meantime, the analyst will have to recode it. For more information, see http://www.cdc.gov/nchs/data/ahcd/Expected_Sources_of_Payment.pdf

Another example of this is the triage level variable. Urgency (i.e., was visit urgent or nonurgent) was collected from 1992-1996. This item was changed to “Immediacy with which patient should be seen” with 4 response categories (e.g., >1 hour-2 hours) in 1997 and then in 2005, “Immediate” was added to the response categories increasing the number to 5. In 2009, the response categories for this item were changed from checkboxes to a write-in response of 1-5, although there are still checkboxes for “No triage” and “Unknown.” During the ED induction interview, the respondent is asked “How many levels are in this emergency service area’s triage system?” If a 3- or 4-level system is used, then the triage level responses on the Patient Record form are mapped to the best corresponding categories in the 5-level system during the editing process.

Q. How is the reliability of estimates determined? How specifically does one calculate this and on what variables?

A. For an estimate to be considered reliable, two criteria must be met. The first condition is that each estimate be based on at least 30 sample or unweighted records. If an estimate is based on < 30 cases it is considered unreliable. The second condition is making sure the weighted data have a relative standard error (RSE) of less than 30 percent. The RSE is equal to the standard error divided by the estimate (i.e., number of visits, percent, rate) and expressed as a percent. Both conditions must be met. For example, if an estimate is based on more than 30 cases, but the RSE is greater than 30 percent it is still considered unreliable. One way to deal with samples not meeting our criteria for reliability is to increase the sample size by combining multiple years of data, if this is consistent with the study question and design.

Part of the sampling error is a result of the survey not being a simple random sample. Due to the survey design, there is a clustering of hospitals within geographic Primary Sampling Units and also a clustering of visits within hospitals. Because of this, the user should always calculate the standard errors using software that takes the clustered nature of the data into account.

Appropriate variance estimates can be calculated using software such as SUDAAN, or survey procedures in SAS, Stata, or SPSS. Make sure all of the records in the data files were included in the analysis in order to obtain the correct sample variance estimate. For example, even if you are only interested in ED visits by children and adolescents under 19 years of age, all records must be included in the program you are running.

Q. Doesn't use of a standard error require an assumption of a normal distribution – when biomedical data are typically skewed?

A. The Central Limit Theorem states that given a sufficiently large sample size, the sample estimate approximates the population estimate, and upon repeated sampling, its distribution would be approximately normal.

Q. NHAMCS data has survey weights attached to each observation that allow one to obtain estimates that are representative of the U.S. population. Should one always use survey weights?

A. For the visit data, one should always use the weights when presenting results as the purpose of the NHAMCS is to produce national estimates. Unweighted visit data should be used only to determine the number of sample cases, because each record can represent thousands of visits. Keep in mind that for a reader who only looks at the findings, it can be confusing to see unweighted data in the text of the Results section. In order to show the number of sample cases upon which an estimate is based, unweighted data may be presented either in the Methods section or in a table in the Results section where both the number of sample cases and the weighted estimates are shown and the headings clearly labeled.

If the researcher is using hospital level data and is presenting the distribution of percentages or means within the ED, then the ED weight (EDWT) should be used and not the patient weight (PATWT) since the unit of analysis is the ED.

DATABASE QUALITY CONTROL

Q. How are variables selected? What is the process for making the decision of what variables to drop and which ones to add?

A. About every 2 years, a group of experts (including representatives of the American College of Emergency Physicians and Society for Academic Emergency Medicine) are sent the current version of the Patient Record form and are asked to review it for additions, deletions, and modifications. In addition, public health agencies of the U.S. Department of Health and Human Services may request that an item be included. Prior to 2012, NHAMCS data were collected on paper forms; therefore, there were space considerations. In 2012, NCHS initiated data collection via a computerized tool. While this alleviates the space problem, respondent burden is always a consideration.

Q. What efforts are made by NCHS to improve item nonresponse rates and to correct errors on the Patient Record forms?

A. Item nonresponse rates in the NHAMCS are generally low (5 percent or less). Census Bureau field representatives are trained to review the Patient Record forms completed by hospital staff for missing data and to obtain the data if possible. One specific check performed by field representatives is for hospital admissions. If an ED submitted no Patient Record forms that had a disposition of “Admit to this hospital”, then the field representative is to ask their ED contact for an explanation and retrieve the missing information if this is an error. NCHS has no control over items that are not documented in the chart, e.g., cause of injury. Data are imputed for some variables with high rates of missing data, e.g., race, ethnicity. Some errors may be corrected during the NCHS editing process.

Q. NHAMCS non-response rates for many variables are higher than the rates of missing data for most clinician-run prospective or retrospective research. NHAMCS recommends that data not be presented when the item nonresponse rate is greater than 30 percent. But are the missing data truly random, and might not missing data levels less than this bias the estimates?

A. Levels of item nonresponse can vary considerably in the survey. Most nonresponse occurs when the needed information is not available in the medical record or is unknown to the person completing the data instrument. Nonresponse can also result when the information is available, but survey procedures are not followed and the item is left blank. If nonresponse is random, the observed distribution for the reported item (i.e., excluding cases for which the information is unknown) would be close to the true distribution. However, if nonresponse is not random, the observed distribution could vary significantly from the actual distribution. Items with a nonresponse >30 percent should be treated with caution.

For variables with more than 10% missing data, an analysis should be performed to assess the sensitivity of results to alternative assumptions about the missing (missing at random compared to assumptions that all missing values take on particular non-missing values). Imputation of variables missing more than 10% may also be considered.

Q. How do you decide when to impute missing data? How can we know when imputation is present and how likely it is to be reliable?

A. Currently, five ED variables are imputed: birth year (only age is on the public use file), sex, race, ethnicity, and triage level. The decision is based on the importance of the variable and/or the percent missing. Each imputed record is flagged. The following variables indicate whether the data were imputed: BDATEFL, SEXFL, ETHNICFL, RACERFL, and IMMEDRFL. The suffix letters FL stand for “flag” and indicate that the missing data have been imputed. Non-imputed versions of these variables are also provided. In regard to reliability, after the data are imputed, edits are performed that may flag an incorrect assignment of a value, e.g., imputed records with a triage level of 1 (immediate) are reviewed if they have a long wait time.

Birth year, sex, race, and triage level (note: the way that triage level was collected changed several times since 1992, see answer to question on temporal comparisons) have been imputed since the inception of the survey in 1992. Ethnicity was first imputed in 2004. Most of the items chosen for imputation are related to the ability to calculate population rates of visits. If age, sex,

race, or ethnicity is missing, then a proper visit rate per population cannot be calculated. Investigators may also choose to impute other variables, but should clearly note this in their methodology.

Q. Annals published a study that found that NHAMCS significantly underestimated the measurement of pregnancy testing compared with chart review.¹ It also found pregnancy testing among only 55% of women with a diagnosis of ectopic pregnancy, when pregnancy testing is a mandatory element of this diagnosis and almost certainly occurred in all. Accurate coding was better with Census staff than hospital employees. How might such discrepancies be explained, and how should they impact future NHAMCS analyses?

A. In review of our training materials, during the survey years used in this study (2002-2006) Census Bureau field representatives were trained to look for “pregnancy test”, although they are directed to look up acronyms. For the 2005-2006 surveys, they were provided with a list of diagnostic tests, including “HCG”, and instructions regarding which checkbox to mark, e.g., pregnancy test”. After communication with the authors of this study, “HCG” was specifically added to the “pregnancy test” item on the ED Patient Record form beginning in 2010 and the field representatives were trained on it at that time. This example demonstrates that errors in the capture of data from medical records are possible with NHAMCS, and we are actively attempting to identify such problematic variables and improve our training. If researchers identify such an issue, NCHS would appreciate hearing about it.

In 2011, we identified an error in the processing of the WHOCOMP variable (i.e., who completed the Patient Record form) which affects years 1999 and 2001-2008. Data for the “Hospital staff” category were incorrectly switched with those for “Census field representative – Abstraction after reporting period.” These reversed values also reverse the finding of Schuur et al,¹ i.e., hospital staff actually completed this item more accurately than Census Bureau field representatives. The 2008 file has been corrected, but until the other files can be updated researchers must be aware of this error, as detailed on pages 5 and 73 of the 2009 public use file documentation.

REPORTING AND INTERPRETING NHAMCS RESEARCH

Q. What are the most common errors and reporting problems that you have seen investigators make with NHAMCS? Please cite some specific examples if possible.

A. Authors tend to make statements about the prevalence of certain conditions in ED patients. As mentioned previously, NHAMCS is a records-based survey. Theoretically, the same patient may visit the ED more than once during the reporting period for the same problem and be selected into the sample each time.

Researchers sometimes present the percentages of visits by age, race, sex, or geographic region. However, if you want to see whether one group uses the ED more than another, it is often preferred to compare visit rates among various population groups such as the number of visits per 100 females in the population. Percentages of ED visits may be of more interest when

comparing the mix of utilization of ED services, for example, what percentage of ED visits by children on Medicaid are for asthma compared to the percentage with private insurance.

If there is press coverage for your article, then you will need to make it clear to the reporter that a rate of 32 visits per 100 persons cannot be rewritten as 32% of the population.

Q. Have you seen any studies whose conclusions overreach the capabilities of NHAMCS? Please cite some specific examples if possible.

A. We have seen studies where the research question could not be answered by analyzing NHAMCS data. One manuscript attempted to report the frequency with which ED patients received a certain class of drugs after a specific procedure. The NHAMCS captures the procedure in question and whether that class of drugs were administered in the ED or prescribed at discharge, but the relative timing of either is not collected.

A second manuscript identified asthmatics by looking at visits where albuterol was administered, and then assessed how frequently steroids were also given as a measure of quality of care. Current National Heart Lung and Blood Institute guidelines recommend the use of systemic steroids when there is incomplete response to a single inhaled beta-agonist. Using NHAMCS data, visits where patients received a single bronchodilator (and would therefore not necessarily be expected to receive steroids) cannot be differentiated from those visits where patients required multiple doses. In addition, drugs that were administered prior to ED arrival are not captured in NHAMCS. An ED patient may have been referred from a primary care office or transferred from another ED, where they may already have received steroids.

Another manuscript exemplifies the potential risk of underreporting with NHAMCS. Entitled “Poor provider adherence to the CDC treatment guidelines in U.S. ED visits with a diagnosis of pelvic inflammatory disease,”² the conclusion of antibiotic undertreatment may be exaggerated by abstractors’ inability to locate the documentation of medications actually given or prescribed. Accordingly, studies of NHAMCS that identify an excess of something (e.g., inappropriate antibiotic prescribing)³ are likely to be more reliable than those reporting a paucity of something.

THE FUTURE

Q. NHAMCS, MEPS, and other governmental databases use different methods of sampling, and in the past with results showing variation about some variables that are in common (e.g., number of ED visits). Why are there multiple different national databases all using different small periodic samples for different areas/variables of interest, instead of one more comprehensive sample and database capturing more variables?

A. Different government agencies conduct ED surveys in order to collect information that will assist them in carrying out their missions. There are multiple ways of looking at the same basic information. The Medical Expenditure Panel Survey (MEPS) is a population-based survey that collects data on the use of specific health services and their associated cost. The National Emergency Department Sample (NEDS) is constructed using records from the State Emergency Department Databases and the State Inpatient Databases; however, only 29 states provide data.

The National Electronic Injury Surveillance System-All Injury Program (NEISS-AIP) collects data on initial non-fatal injury-related ED visits. As mentioned previously, NHAMCS uses a national probability sample of ED visits and sample data are weighted to produce annual national estimates. In addition, it is the most comprehensive of all of the ED databases. Efforts to combine surveys are currently underway (see answer to next question). These various ED-related research data sources have been compared and discussed in greater detail elsewhere.^{4,5}

Q. What is the future of NHAMCS? We might be able to capture many more visits and data points (with the same chart limitations of NHAMCS) in the not so distant future with electronic health records, and have fewer problems with small sample extrapolation if data could be anonymous and captured on a larger scale. The NHAMCS sampling strategy has not really fundamentally changed over time, has it? Should it?

A. The NHAMCS sampling strategy has not fundamentally changed since its inception in 1992, but it will be transformed in the near future. In 2013, NHAMCS will be integrated into the National Hospital Care Survey (NHCS) which was launched in 2011 to collect data on inpatient discharges. The Drug Abuse Warning Network (DAWN), which has collected data on drug-related ED visits since 1972 and was administered by the Substance Abuse and Mental Health Services Administration, will also be integrated into the NHCS in 2013. This coordination of survey activities will increase the wealth of data on health care utilization across episodes of care and allow for linkages to other data sources such as the National Death Index and data from the Centers for Medicare and Medicaid Services. In 2012, NCHS will conduct a pretest which will test various sampling and data collection methods with the intention of moving toward electronic data collection.

REFERENCES

1. Schuur JD, Tibbetts SA, Pines JM. Pregnancy testing in women of reproductive age in US emergency departments, 2002 to 2006: assessment of a national quality measure. *Ann Emerg Med* 2010 May;55(5):449–457.e2.
2. Shih TY, Gaydos CA, Rothman RE, Hsieh YH. Poor provider adherence to the Centers for Disease Control and Prevention treatment guidelines in US emergency department visits with a diagnosis of pelvic inflammatory disease. *Sex Transm Dis*. 2011 Apr;38(4):299-305.
3. Vanderweil SG, Tsai CL, Pelletier AJ, et al: Inappropriate use of antibiotics for acute asthma in United States emergency departments. *Acad Emerg Med* 2008; 15:736-743.
4. Owens PL, Barrett ML, Gibson TB, Andrews RM, Weinick RM, Mutter RL. Emergency Department Care in the United States: A Profile of National Data Sources. *Ann Emerg Med*. Elsevier Inc.; 2010 Aug. 1;56(2):150–165.
5. Hirshon JM, Warner M, Irvin CB et al. Research using emergency department-related data sets: Current status and future directions. *Acad Emerg Med*. 2009;16:1103-09.