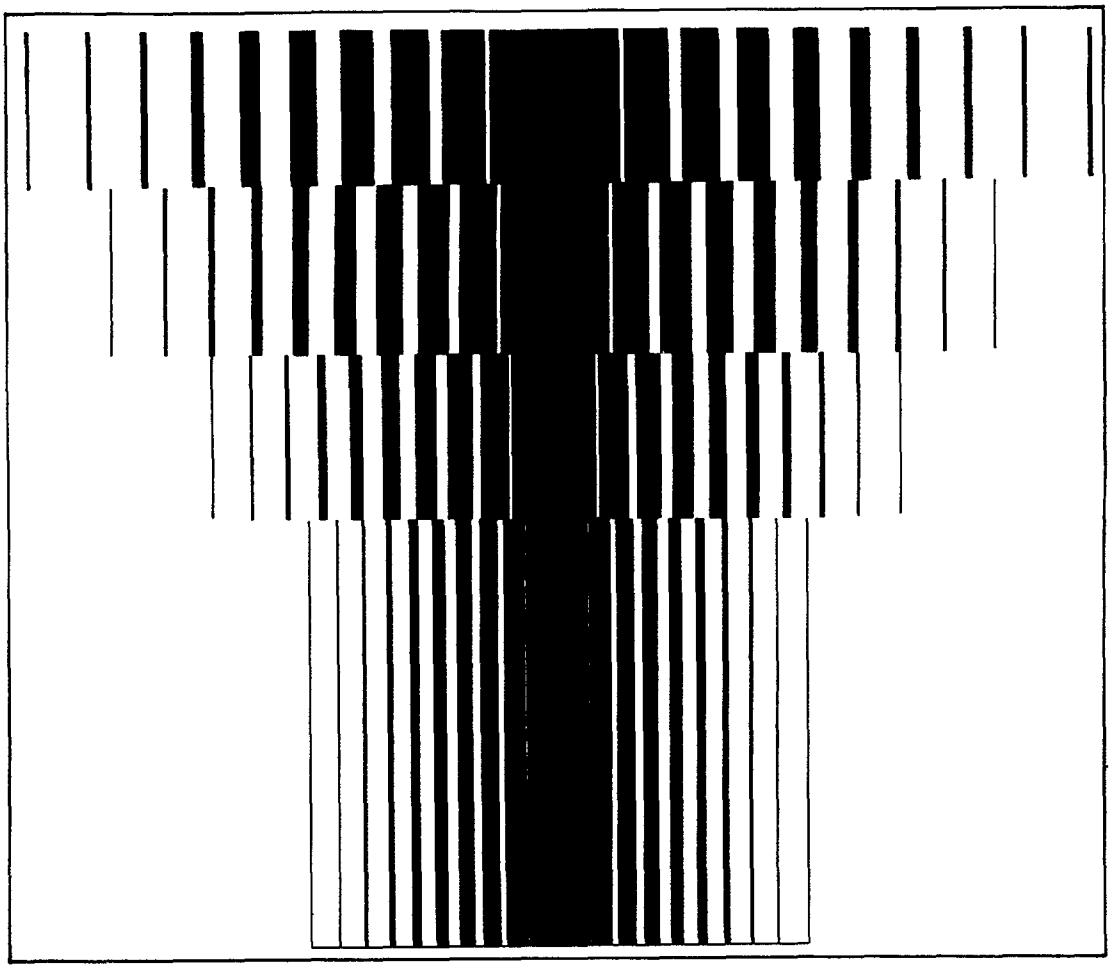
Evaluation of Data Collection and Coding for Medical Conditions in the National Medical Care Utilization and Expenditure Survey

Series A, Methodological Report No. 5



U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

Published by
Public Health Service
Centers for Disease Control
National Center for Health Statistics

February 1988

National Medical Care Utilization and Expenditure Survey

The National Medical Care Utilization and Expenditure Survey (NMCUES) is a unique source of detailed national estimates on the utilization of and expenditures for various types of medical care. NMCUES is designed to be directly responsive to the continuing need for statistical information on health care expenditures associated with health services utilization for the entire U.S. population.

NMCUES will produce comparable estimates over time for evaluation of the impact of legislation and programs on health status, costs, utilization, and illness-related behavior in the medical care delivery system. In addition to national estimates for the civilian noninstitutionalized population, it will also provide separate estimates for the Medicaid-eligible populations in four States.

The first cycle of NMCUES, which covers calendar year 1980, was designed and conducted as a collaborative effort between the National Center for Health Statistics, Public Health Service, and the Office of Research and Demonstrations, Health Care Financing Administration. Data were obtained from three survey components. The first was a national household survey and the second was a survey of Medicaid enrollees in four States (California, Michigan, Texas, and New York). Both of these components involved five interviews over a period of 15 months to obtain information on medical care utilization and expenditures and other health-related information. The third component was an administrative records survey that verified the eligibility status of respondents for the Medicare and Medicaid programs and supplemented the household data with claims data for the Medicare and Medicaid populations.

Data collection was accomplished by Research Triangle Institute, Research Triangle Park, N.C., and its subcontractors, the National Opinion Research Center of the University of Chicago, Ill., and SysteMetrics, Inc., Berkeley, Calif., under Contract No. 233–79–2032.

Co-Project Officers for the Survey were Robert R. Fuchsberg of the National Center for Health Statistics (NCHS) and Allen Dobson of the Health Care Financing Administration (HCFA). Robert A. Wright of NCHS and Larry Corder of HCFA also had major responsibilities. Daniel G. Horvitz of Research Triangle Institute was the Project Director primarily responsible for data collection, along with Associate Project Directors Esther Fleishman of the National Opinion Research Center, Robert H. Thornton of Research Triangle Institute, and James S. Lubalin of SysteMetrics, Inc. Barbara Moser of Research Triangle Institute was the Project Director primarily responsible for data processing.

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402

Suggested Citation

Gans, J. E. Evaluation of Data Collection and Coding for Medical Conditions in the National Medical Care Utilization and Expenditure Survey. *National Medical Care Utilization and Expenditure Survey.* Series A, Methodological Report No. 5. DHHS Pub. No. 88–20005. National Center for Health Statistics, Public Health Service. Washington. U.S. Government Printing Office, Feb. 1988.

Library of Congress Cataloging-in-Publication Data

Gans, Janet.
Evaluation of Data Collection and Coding for Medical Conditions in the National Medical Care Utilization and Expenditure Survey.
(National Medical Care Utilization and Expenditure Survey. Series A, Methodological report ; no. 5) (DHHS pub. ; no. 88–20005)
Bibliography: p.
1. National Medical Care Utilization and Expenditure Survey (U.S.)—Evaluation. 2. National Health Interview Survey (U.S.)—Evaluation. 3. Medical care surveys—Statistical methods—Evaluation. 4. Health surveys—Statistical methods—Evaluation. 5. Nosology.
I. National Center for Health Statistics (U.S.)
II. Title. III. Series: National Medical Care Utilization and Expenditure Survey (Series). Series A, Methodological report ; no. 5. IV. Series: DHHS publication ; no. 88–20005. [DNLM: 1. National Medical Care Utilization and Expenditure Survey (U.S.) 2. Data Collection—methods. 3. Health Surveys—United States. W2 A N2235n no. 5]
RA410.7.G36 1987 362.1'0723 87–600341
ISBN 0–8406–0384–3

Contents

List of Text Tables

Symbols

- - -	Data not available
. . .	Category not applicable
-	Quantity zero
0.0	Quantity more than zero but less than 0.05
*	Test statistic is significant at 0.05 level
**	Test statistic is significant at 0.01 level

Evaluation of Data Collection and Coding for Medical Conditions in the National Medical Care Utilization and Expenditure Survey

By Janet E. Gans, Ph.D., of NORC (formerly the National Opinion Research Center)

Executive Summary

This is one of five reports that evaluates the National Medical Care Utilization and Expenditure Survey. That survey was designed for the collection of data about the U.S. civilian noninstitutionalized population during 1980. During the course of the survey, information was obtained on health, access to and use of medical services, associated charges and sources of payment, and health insurance coverage. This report evaluates procedures used in the National Medical Care Utilization and Expenditure Survey to collect and code medical conditions. Estimates derived from the National Medical Care Utilization and Expenditure Survey are compared with estimates derived from the National Health Interview Survey. Several of the procedures used in the two surveys were identical or similar.

In addition to estimates for the national population, the National Medical Care Utilization and Expenditure Survey was designed to produce a data base to help analyze expenditures and health services provided by the Medicare and Medicaid programs. A national household probability sample was augmented by samples drawn from the Medicaid eligibility rolls of California, Michigan, New York, and Texas. Information from State eligibility and claims files was obtained for people in the household samples with reported Medicaid coverage. Information from Federal Medicare files was obtained for people reported to be covered by Medicare. Estimates of morbidity in this report are based solely on the national household sample, and exclude conditions reported by respondents in the four-State Medicaid sample.

For each of the aspects of medical condition collection and coding included in the evaluation, procedures are examined that could affect estimates of incidence and prevalence. These include the length of the recall period, the multiple coding of a single condition, definitional differences between condition codes used to construct a morbidity category, miscoding conditions as chronic, adaptations of the *Health Interview Survey Medical Coding Manual and Short Index,* and the absence of a checklist.

Differences between lengths of the recall periods in the two surveys and the use of a checklist in one but not the other seem to account for most differences between the morbidity estimates in the two surveys. Overall, the 3-month recall period used in the National Medical Care Utilization and Expenditure Survey resulted in decreased reporting of minor, acute illnesses compared with that from the National Health Interview Survey, with its 2-week reference period. The use of repeated interviewing in the National Medical Care Utilization and Expenditure Survey did not fully compensate for the checklists used in the National Health Interview Survey for reporting chronic conditions.

Only a small proportion of conditions were multiply coded (assigned more than one ICD–9 code) in the National Medical Care Utilization and Expenditure Survey, and there is little evidence of systematic bias in the use of this practice. The practice of multiple coding should be abandoned in the future unless a compelling reason for its continued use is found. Definitional differences, too, had minimal impact on incidence and prevalence estimates, though the development of standardized definitions for morbidity categories across major health surveys is recommended. The tendency to miscode conditions as chronic in the National Medical Care Utilization and Expenditure Surveys was also minimal but could be avoided entirely in future longitudinal health surveys by finding out the duration or termination of an illness.

NOTE: This report was prepared by NORC, formerly the National Opinion Research Center, by contractual arrangement with the National Center for Health Statistics (Contract No. 282–84–2109). The author is grateful for the support received during all stages of the preparation of this document, from colleagues at NORC, the University of Chicago, and from the staff of the National Center for Health Statistics. At NORC, Dr. Martin R. Frankel and Dr. Roger Tourangeau provided valuable consultation on statistical issues. Harrison Greene and Hyman Bern provided quality programming support; Hyman Bern performed the analysis that generated sampling errors for national estimates. Ms. Pearl Zinner offered much-needed background information on NMCUES medical condition coding; the author appreciates her close scrutiny of each report draft. Dr. Ronald Andersen of the Center for Health Administration Studies at the University of Chicago helped identify the issues addressed in the report. The author also is grateful for the careful editing of the report by Ms. Sofi Ravin and Ms. Susan Campbell of NORC, who enhanced its conceptual clarity.

Continual support was provided by the National Center for Health Statistics and our project officer, Mr. Robert Wright, Chief, Utilization and Expenditure Statistics Branch. When questions or potential errors in the data were identified during the analysis, Ms. Michele Chyba of the Division of Health Interview Statistics quickly and patiently solved the problems. Dr. Andrew White encouraged the inclusion of sampling errors for national estimates in the analysis and provided suggestions for revisions to the technical appendix. Editors in the Publications Branch provided valuable assistance during the preparation of the final report.

The adaptations made by the National Medical Care Utilization and Expenditure Survey to the National Health Interview Survey Medical Coding Manual also had minimal impact on estimates, with the possible exception of the absence of merging procedures in the National Medical Care Utilization and Expenditure Survey.

Introduction

The primary purpose of the National Medical Care Utilization and Expenditure Survey (NMCUES) was the collection of accurate information on the use and costs of health-care services in the United States. Many of the instruments and procedures used to elicit and record the reporting of morbidity conditions in NMCUES paralleled those used in the National Health Interview Survey (NHIS), a key source for estimates of acute and chronic illness in the United States. Given some similarities in methodology, it would be expected that selected estimates of types of illness and disease produced through NMCUES and NHIS would be similar for the year 1980, the year for which information on the Nation's health was collected through both surveys.

However, there were important differences in NMCUES and NHIS that affected incidence and prevalence estimates. For example, NHIS respondents were interviewed only once during the year and were asked to report medical events during the 2-week period prior to the interview. NMCUES respondents were interviewed several times during the course of the year and asked to report utilization and expenditure events that occurred during a 3-month interval prior to each interview. As noted in the report, some of these methodological differences reflect differences between the primary purposes of the two studies. NHIS was designed primarily to estimate the incidence and prevalence of morbidity conditions, whereas the primary goal of NMCUES was the collection of information about the cost and utilization of health-care services necessitated by medical conditions.

This report identifies and examines differences between NMCUES and NHIS—particularly in data collection and coding procedures—that could plausibly affect morbidity estimates. Specifically, the study examines the following:

1. *Instruments and strategies used to elicit reporting of chronic conditions*—A checklist was used in NHIS, whereas a combination of repeated interviews, a 3-month recall period, and an expanded questionnaire were employed in NMCUES.
2. *Definitions of illness and disease*—Different sets of ICD–9 codes were used in the two studies to construct some condition categories, such as cerebrovascular and heart disease.
3. *The miscoding of conditions as chronic in NMCUES*—This is based upon the 3-month convention.
4. *Conventions in coding a single medical condition*—In NHIS only one ICD–9 code could be assigned to a condition whereas in NMCUES up to three codes could be assigned to a condition.
5. *Coding instructions*—For NMCUES, a total of 58 annotations were added to the original coding instructions in the *Health Interview Survey Medical Coding Manual and Short Index* (National Center for Health Statistics, 1979); 25 of these affected the type of ICD–9 code assigned.

Based on the results of these comparisons, recommendations as to which of the methodological procedures developed for NMCUES should be retained, further adapted, or abandoned in future surveys similar to NMCUES in purpose and design are included in this report.

In both NMCUES and NHIS, a morbidity condition, or more simply a "condition," is defined as any entry on the questionnaire that describes a departure from a state of physical or mental well-being. The number and types of conditions reported in each survey are used to create prevalence and incidence estimates, which describe the presence of disease in a population. "Prevalence" refers to the total number of cases of a morbidity condition present during a particular period in time, regardless of date of onset. Prevalence estimates usually refer to the presence of chronic conditions. "Incidence" measures reflect the number of new cases of an illness or disease that appear in a specific population over a specified time interval. Incidence estimates usually refer to acute conditions. Unless otherwise specified, "prevalence" as used in this report refers to chronic conditions, and "incidence" refers to acute conditions.

This report is divided into seven additional sections. The first section provides background information about the two survey designs. Subsequent sections describe and analyze the methodological differences between NMCUES and NHIS. Specifically, the report addresses differences on the five dimensions listed above. The last section discusses the major findings from the study and offers particular recommendations about the potential application of NMCUES methodology to other longitudinal health surveys.

Survey Designs and Procedures: Parallels Between NMCUES and NHIS

Samples

NHIS is a nationwide, cross-sectional survey of the civilian noninstitutionalized population of the United States. The sampling plan follows a multistage probability design that permits continuous sampling of households. Each week of the year households in 1 of the 52 NHIS subsamples are interviewed and the sample is additive over time. Each household is interviewed one time in a face-to-face interview.

Primary sampling units (PSU's), which consist of a county, a small group of contiguous counties, or a standard metropolitan statistical area, constitute the first stage of the sample design. These include area segments, which are defined geographically; list segments, for which 1970 census registers were used as the frame; and permit segments, that is, updated lists of building permits issued in sample PSU's since 1970. Each segment contains an expected four households, the ultimate sample unit. In all, the sample included 376 primary sampling units and 12,000 segments, yielding approximately 39,000 eligible households containing about 103,000 persons. (Detailed descriptive materials on sample design, estimation procedures, questionnaire development, data collection, and field procedures in NHIS are found in Bean, 1970; National Center for Health Statistics, 1964; National Center for Health Statistics, 1975; Simmons, 1975).

NMCUES, too, was a national multistage probability sample of the civilian noninstitutionalized population in the United States. The first stage consisted of primary sampling units, and the second stage consisted of census enumeration districts or block groups. Smaller area segments constituted the third stage. All together, the sample included 135 PSU's, 809 second-stage units, and 809 segments.

The NMCUES survey consisted of three major components: a national household sample; a four-State sample of Medicaid households; and the collection of records from two sources—State records for respondents in the four-State sample and Federal Medicare records for the national household survey respondents. In all, approximately 6,000 households were interviewed in the national household survey, and data were obtained for 17,600 individuals.

Respondents in the national household survey were interviewed over a 14-month period during 1980–81, with data being collected for the calendar year 1980. Most households were interviewed five times at approximately 10–12 week intervals, though some households were interviewed only four times. The first round of interviewing began in early February and ended in late April of 1980. The second round of interviewing was conducted between early May and the end of July. Round 3 interviews took place between early August and October. Round 4 interviews were conducted during a 9-week period from November 1 through mid-December, and Round 5 interviews were held between the second week of January and the end of March 1981 (Bonham, 1983).

Data Collection Instruments and Procedures

NHIS and NMCUES interviewers followed similar procedures to trigger the reporting of conditions and used the same general flow of questions to collect condition data. Interviewers in both surveys first asked respondents whether or not a particular event had occurred (for example, a visit to the doctor) and, if it had, the number of times it had occurred. Respondents next were asked to provide details about each reported medical event, including the condition associated with it. Once information on all events had been collected, interviewers filled out a separate condition section, asking respondents for detailed information about each condition reported during the interview. Through the condition sections of both survey instruments, information was gathered about the name and cause of the condition, the part(s) of the body affected, and the date when the condition first was noticed.

Data collection instruments used in NMCUES included a core questionnaire and supplements to the questionnaire, a computerized summary of responses, and a control card. Whenever a respondent reported a condition in NMCUES, the interviewer recorded the condition name (or key descriptive words) in the person's column on the control card and assigned a unique two-digit number to the condition. Interviewers assigned the two-digit condition numbers sequentially during the entire data collection period. If a condition name or description had been entered previously onto the control card for a respondent, the interviewer asked, "Is this the same condition you told me about (earlier today/in a previous

interview)?" If the answer was "yes," nothing new was recorded on the control card. The interviewer entered the two-digit number associated with the condition in the questionnaire under "COND #" and continued with the interview. The list of conditions on the control card became a cumulative master list of all unique conditions reported for a respondent and was the source for linking conditions from one section of the questionnaire to another, as well as between rounds.

Coding Procedures

In both NHIS and NMCUES, a four-digit condition code was assigned to each condition reported for a respondent. Coding procedures in both surveys were based on the *Health Interview Survey Medical Coding Manual and Short Index* (National Center for Health Statistics, 1979) as the primary source, though in NMCUES some procedures described in the NHIS coding manual were modified (NMCUES, 1980). The secondary source for coding conditions used in both surveys was the Ninth Revision of the *International Classification of Disease, 1975 Revision, Manual of the International Statistical Classifications of Diseases, Injuries, and Causes of Death* (ICD–9), Volumes 1 and 2 (World Health Organization, 1977 and 1978, respectively). In NMCUES, National Opinion Research Center (NORC) and Research Triangle Institute (RTI) staff trained in condition coding assigned ICD–9 codes on an ongoing basis at the end of each round.

In both NHIS and NMCUES a prefix was assigned to the four-digit ICD–9 code, indicating whether a condition was chronic or acute. The key difference between chronic and acute conditions was the date of onset. Acute conditions in both NMCUES and NHIS were defined as having lasted less than 3 months. Similarly, chronic conditions in both surveys were defined as having lasted 3 months or more or as belonging to certain classes of diseases always considered chronic regardless of date of onset (for example, arthritis or impairments such as a missing hand). Chronic conditions were assigned a prefix "1," and acute conditions were assigned a prefix "2." When date of onset was unknown or could not be determined from available information, NMCUES coders assigned a prefix code "3" to the condition; NHIS coders were instructed to code such conditions as acute.

Data Tapes

The two data bases used in this analysis are the 1980 public use data tapes from NHIS and the 12-month files from NMCUES. Data from NHIS were cleaned, edited, and checked for consistency in the course of preparing a public use tape. Two NHIS files, the condition file and the hospital file, are included in the analysis. Both are event-level files; that is, each file contains a record for every hospital stay or condition. The weights used in the analysis are included in the NHIS data.

The NMCUES 12-month files are an "intermediate" set of files. The 12-month files were chosen for this analysis because the acute-versus-chronic prefix, critical to the analysis, appears in the file. In the NMCUES public use tapes, the prefix had been edited out. The 12-month files have undergone minimal cleaning to remove inconsistencies and out-of-range values that appear in the originally keyed data. In addition, codes used to describe missing data ("98's" and "99's") were keyed, taking the place of blanks. The variables included in the analysis were drawn from the hospital-stay file and the condition file, both of which are event-level files. The condition file contains no imputed conditions. The 72 dummy records in the condition file are excluded from the analysis.

Weighting Procedures

In NHIS, quarterly and annual prevalence estimates for chronic conditions were calculated. The chronic conditions included in NHIS prevalence estimates appeared on one of the study's six checklists. Using prevalence estimates of conditions not specified on the six lists is not recommended in NHIS (National Center for Health Statistics, 1983). Prevalence estimates for each calendar quarter were calculated by averaging estimates of "checklist" conditions for all weeks of interviewing in a quarter or 13-week period. Annual prevalence data were derived by averaging the four quarterly figures.

Statistics for the incidence of acute conditions, which reflect the number of occurrences during a specified time period, also were averaged for each quarter or 13-week period to adjust for the 2-week length of recall. The estimated quarterly total for the condition is 6.5 times the average 2-week estimate produced by the 13 successive samples taken during the period. The annual total is the sum of the four quarters. Thus the morbidity experience of persons interviewed during a year—experience that actually occurred for each person during a 2-calendar-week interval prior to the week of interview—is treated as though it measured the total of such experience during the year.

Unlike in NHIS, no adjustment for length of recall was made in NMCUES. Each acute and chronic condition in this report received the weight assigned to the respondent (the basic person weight).

Differences in Data Collection Methods

Methods of Data Collection in the Two Surveys

Despite similarities between the NMCUES and NHIS questionnaires, there were important differences between the study designs that could have affected incidence and prevalence estimates. These included the number of interviews, the length of the reference period, and the approach to instrumentation.

As was mentioned before, NHIS respondents were interviewed only once during the course of the year, whereas NMCUES respondents were interviewed as many as five times over a 14-month period. Also, different time intervals were used in NHIS and NMCUES when asking respondents to report medical events. NHIS respondents were asked to report only medical events that had occurred during a 2-week period prior to the week of the interview, though a 12-month reference period was used for hospitalizations and for chronic conditions. In NMCUES, the reference period varied (between rounds), but its average length was 10 to 12 weeks, or about 3 months.

Overall, because of the longer reference period in NMCUES, it would seem logical that the level of forgetting would be greater among NMCUES respondents than among NHIS respondents, especially for relatively minor acute conditions. The findings discussed by Tourangeau and Rasinski (1987) generally support this hypothesis, although they suggest that the rate of forgetting over time may be slow. In short, though it is impossible to prove an underreporting of conditions by NMCUES respondents, there is some evidence that the increased reference period may have decreased the reporting of conditions.

Whereas the 3-month reference period in NMCUES may have created problems of respondent recall, past research has demonstrated that the 2-week recall period in NHIS is too restrictive a time period to allow for the manifestation of events associated with chronic conditions. To avoid an underreporting of chronic conditions resulting from the 2-week recall period, a checklist was developed for NHIS to elicit the reporting of specific chronic conditions, even when they are not associated with a medical event, limitation, or disability. The use of a checklist boosts reporting levels for existing conditions on the list because the items provide a vehicle for respondents to report conditions even when they do not entail disability days or other medical events. By using a checklist, the primary aim of NHIS was served—to collect accurate estimates of morbidity conditions.

The NMCUES questionnaire, in contrast, reflects that survey's primary emphasis on utilization and expenditures. In seeking information about respondents' contacts with health-care services, the NMCUES design altered the NHIS questionnaire from which they had so heavily borrowed. Whereas in NHIS there was an allowance for conditions to be reported in six sections (bed-disability days, school- and work-loss days, dental visits, medical visits, hospital stays, and functions limitations), in NMCUES the number of sections in the questionnaire that would elicit the reporting of acute and chronic conditions was increased. The NMCUES questionnaire contained nine sections. Respondents were asked to report conditions associated with hospital outpatient department visits, hospital stays, and emergency room visits; it included inquiries about contacts with providers outside the hospital setting. NMCUES respondents were questioned about prescribed medicines, other medical expenses, and disability days. Conditions also were recorded in the Limitations and Background Information sections (Supplement 1), and the Barriers to Care section (Supplement 5). Compared with the NHIS instrument, the NMCUES questionnaire provided previously unexplored opportunities for respondents to recall and report medical conditions.

One additional alteration should be noted briefly: the omission in the NMCUES survey of three questions about the termination of illnesses. As will become apparent later, that omission affected the designation of a condition as chronic or acute in NMCUES.

Analysis: The NHIS Checklist Versus Three NMCUES Data Collection Procedures

Currently two techniques are used in NHIS to collect medical conditions: a "person approach," in which respondents are asked a series of questions about health-related actions, such as visits to the doctor; and a "condition approach," in which they are led through a checklist of specific conditions that might otherwise remain unreported.

In NMCUES only the "person approach" was used to collect information about conditions. For this reason

prevalence estimates for chronic conditions that appear on the NHIS checklist would probably be lower in NMCUES than in NHIS. However, this liability may have been offset by the extended reference period in NMCUES or by the expanded number of sections in the NMCUES questionnaire. Earlier NCHS-sponsored studies also suggest that the repeated interviewing done in NMCUES might boost prevalence estimates; those studies have shown that repeated interviews substantially increase the reporting of chronic conditions. The question addressed in this section is whether the 3-month reference period, additional sections in the questionnaire used to collect condition data, and repeated interviewing during the year compensated for the lack of a checklist in NMCUES in producing estimates of chronic conditions.

The NHIS sample was divided into sixths, and respondents in each sixth of the sample were asked explicitly if they had any condition that appeared on a list of about 20 conditions. Six different checklists were used; one for each sixth of the sample. (Prevalence estimates published by NCHS are restricted to the responses of the one-sixth sample who were asked explicitly about conditions on a specific checklist.) The analysis below examines the effects of the checklist, comparing prevalence estimates for the sample in NHIS who were asked about specific conditions and for a sample of respondents who were not. The difference between the two reporting levels can be interpreted as the gain resulting from the checklist. The analysis examines next whether NMCUES prevalence estimates approximate more closely the estimate from the sample that was asked explicitly about the conditions on the checklist or from the sample that was not.

As shown in Table A, estimates derived from NHIS respondents who were asked directly about conditions were substantially greater than estimates from the subsample who were not asked explicitly about those conditions. Differences between the NHIS checklist estimates and the NMCUES estimates were significant for 7 of the 11 conditions on the checklist. The NMCUES estimate (based on identical ICD–9 codes to define the category) represents an improvement over the nonchecklist sample but does not fully compensate for the use of the checklist.

The three exceptions to this pattern were estimates for upper gastrointestinal conditions not elsewhere classified (NEC), enteritis, and diseases of the gallbladder. More than twice as many cases of upper gastrointestinal conditions, fewer than one-fifth as many cases of enteritis, and slightly more cases of gallbladder conditions were estimated through NMCUES than through NHIS. It appears that NMCUES medical condition coders assigned a disproportionate number of digestive conditions into an "other" category (that is, upper gastrointestinal conditions NEC), adopting a conservative coding strategy for ill-defined digestive conditions. Thus the relatively low number of cases of enteritis might have been categorized as a gastrointestinal condition NEC.

Overall, the effect of the checklist on prevalence estimates was greater than the combined effects of repeated interviews, an expanded questionnaire, and an increased reference period. However, for most categories, NMCUES estimates were considerably closer to the NHIS checklist estimates than to the nonchecklist estimates, an indication of the substantial compensatory effects of the NMCUES procedures.

Table A

NHIS and NMCUES prevalence estimates for digestive conditions and the standard error of difference between the NHIS and NMCUES estimates, by selected condition categories: United States, 1980

Selected condition category	Digestive checklist (NHIS)[1]	Other checklist (NHIS)[2]	NMCUES	NHIS-NMCUES difference	*t* of difference (NHIS checklist-NMCUES)
	Estimates in thousands				
Ulcer of stomach	3,615	1,004	3,187	441	1.36
Frequent constipation	3,579	63	697	2,893	*10.13
Hernia of abdominal cavity	3,888	1,013	3,082	818	*2.28
Upper gastrointestinal conditions NEC	3,720	251	8,413	4,517	*7.76
Gallbladder	1,217	344	1,455	212	.88
Gastritis	1,706	168	723	1,020	*4.22
Diverticula of intestine	1,380	122	556	824	*4.12
Enteritis	2,293	502	89	2,203	*10.50
Intestinal trouble	1,649	131	1,275	374	1.51
Stomach trouble	824	255	486	338	*2.35
Liver trouble and other unspecified diseases of liver	415	96	371	54	.46

[1]Prevalence estimates are based upon a one-sixth subsample of the population who responded to a checklist of digestive diseases.
[2]Prevalence estimates represent the one-sixth subsample who were asked about conditions on the skin and musculoskeletal diseases checklists.

Definitions of Morbidity Categories

Definitions

Through both NHIS and NMCUES, data were collected on illness (for example, cardiovascular disease and emphysema), but the ICD–9 codes used to construct morbidity categories in the two studies were not always the same. Both NHIS and NMCUES relied upon the *Health Interview Survey Medical Coding Manual and Short Index* as revised in January (National Center for Health Statistics, 1979) as the primary source for assigning diagnostic codes to conditions. However, in NHIS those guidelines were modified somewhat: Its diagnostic categories and their constituent codes appear in RECODE 1 (all conditions), RECODE 2 (acute conditions), and RECODE 3 (chronic conditions) of the NHIS Public Use Data Files for 1980. In NMCUES certain morbidity categories defined in the NHIS coding manual also were modified, but the "Basic Tabulation List" in ICD–9 (World Health Organization, 1975, pp. 746–755) was used in the survey.

The morbidity categories examined in this report are illnesses reported most often in the major health surveys conducted in the United States: fractures and dislocations, cerebrovascular disease, heart diseases, influenza, hypertensive disease, acute respiratory diseases, and the common cold. Two other categories, malignant neoplasms and arthritis and rheumatism, were also among the most-reported diseases but were excluded from the analysis because the NHIS questionnaire was not designed specifically to measure malignant neoplasms, and the combination of arthritis and rheumatism was not defined as a single condition category in NMCUES.

Analysis: Definitions Versus Other Sources of Difference

The analysis in this section of the report examines whether differences between estimates of incidence and prevalence in NHIS and NMCUES are due to differences between the diagnostic codes used to define a given morbidity category or to some other procedural difference between the two studies. The analysis, using unpublished NCHS figures, first examines differences between NHIS and NMCUES estimates when the definition used in each study is imposed on its own data base. The second part of the analysis compares estimates when the two definitions are imposed first on NHIS data and then on NMCUES data. The question addressed throughout is whether observed differences or similarities should be attributed to the definition used or to other features of the two studies.

If the NHIS and NMCUES definitions yield similar estimates when applied to a single data set, then differences between the original NHIS and NMCUES estimates can be attributed to sources other than definition, because all factors other than the definitions would have been controlled for. By the same reasoning, definitional effects could be inferred if differences between estimates persist after all other factors have been taken into account. Of course, the possibility that persisting differences might be due to other methodological differences could not be ruled out. Although not conclusive, this analysis offers suggestive evidence about whether differences between definitions of a morbidity category substantially affect incidence and prevalence estimates, or whether differences between estimates more likely result from procedural differences such as the length of the reference period.

Table B presents selected condition categories and ICD–9 codes used to define them in NHIS and NMCUES. As can be seen, the ICD–9 codes used to define each condition are different in each survey.

Table C shows estimates derived from imposing NHIS and NMCUES definitions first on NHIS data and then on NMCUES data. For some diseases, NHIS estimates consistently exceeded those of NMCUES (compare columns A and D). Such was the case for fractures and dislocations, influenza, hypertensive disease, and colds. Differences between estimates were significant for all conditions except fractures and dislocations. The differences remained significant when NHIS and NMCUES definitions were imposed on the NMCUES data. The only exception was the common cold, where the two definitions produced identical estimates.

Definitional effects are evident and significant for heart disease. However, this may result in part from NHIS rules that permitted information about heart disease and hypertension to be "merged"; that is, when information about those conditions was reported on separate condition pages of the questionnaire, they could be consolidated in the diagnostic coding. NMCUES, in contrast,

left the two conditions as distinct diagnoses. As a probable result of this departure from NHIS coding procedures, NMCUES reports a lower rate of hypertension compared with NHIS and has a higher prevalence than NHIS of heart disease.

It is likely that both definitional and other methodological factors affected estimates of acute respiratory conditions. The incidence of acute respiratory ailments including common cold estimated by NHIS was 124 million cases (Table C), one-third more than the NMCUES estimate of 88 million. Those differences disappeared when the NMCUES definition was imposed on NHIS data, an indication that the NMCUES definition was no more restrictive than that used in NHIS. When the NHIS definition was imposed on NMCUES data, however, the incidence estimate dropped substantially below the NMCUES estimate, to 65 million (Table C). It could be that the NHIS definition for acute respiratory conditions is more restrictive than the NMCUES definition but that other methodological differences, such as a shorter recall period, enabled more cases to actually be recorded in NHIS.

In sum, several points should be noted. First, the definition used can affect estimates of the incidence and prevalence of disease in some condition categories. Second, by ruling out the effects of definitional factors in certain cases, this analysis has pointed to other sources that may have produced differences between estimates. After definitional differences have been taken into account, significantly more instances of acute respiratory ailments, influenza, and the common cold still are recorded through NHIS. Though the exact reason for this difference is not clear, it could be that the more lengthy recall period in NMCUES and the relatively low level of severity of these conditions led to their being underreported in NMCUES compared with NHIS. In other words, it would not be surprising if the level of forgetting is high for relatively unimportant, acute events such as the common cold. If this is the case, the 3-month recall period used in NMCUES may be sufficiently protracted as to miss 50 percent of all common colds and, possibly, their associated costs.

The impact of definitional differences may have been diminished by the number and kind of conditions used in this analysis. Definitional differences might have been more powerful if a greater number of conditions had been examined. This would have produced more opportunities to observe significant changes in estimates in cases where two definitions were imposed on the same data set. Also, had the conditions chosen been of a kind less often reported—and thus less conventionally

Table B

ICD–9 diagnostic codes for selected condition categories as defined by NHIS and NMCUES

Condition category	NHIS	NMCUES
Fractures and dislocations	733.8; 800–839	800–839
Cerebrovascular	348.5; 430–435; 437.0–.2, .4–.6, .8, .9	430–438, 343
Heart disease	390; 392; 393–398; 402.1, .9; 404.1, .9; 413–415.0; 416; 417 (except 417.1); 420–424; 425 (except 425.3, .5); 426; 427.0–.6, .8, .9; 428.0, .1, .9; 429.0–.3, .5, .8, .9; 785.0–.3	390–429
Influenza	487.0, .1, .8	487
Hypertension	401; 402 (except 402.1, .9); 403; 404 (except 404.1, .9); 405; 796.2	401–405
Acute upper respiratory disease	461–465; 470; 471; 475; 478.0–.7, .9	460–465 470–478
Common cold	079.3; 460	460

SOURCES: National Center for Health Statistics: *National Health Interview Survey 1980 Public Use Data Tape Documentation—Part I.* (Includes annotations to the *Medical Coding Manual and Short Index.*) Public Health Service. Hyattsville, Md., June 1983.

World Health Organization: *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death.* Vol. 1. Based on the Recommendations of the Ninth Revision Conference, 1975. Geneva. World Health Organization, 1977.

World Health Organization: *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death.* Vol. 2. Based on the Recommendations of the Ninth Revision Conference, 1978.

Table C

Estimates for selected condition categories derived from the NHIS and NMCUES samples by the disease definition used in each study

Condition category	NHIS Sample		NMCUES Sample		*t* of the difference	
	NHIS definition A	NMCUES definition B	NHIS definition C	NMCUES definition D	NHIS on NHIS vs. NMCUES on NMCUES E	NHIS and NMCUES on NMCUES F
	Estimates in thousands					
Fractures and dislocations	7,941	7,941	6,995	6,995	1.24	...
Cerebrovascular	2,101	2,083	2,157	2,197	0.25	1.69
Heart disease	16,434	36,887	11,650	32,503	*10.68	*32.62
Influenza	113,799	113,799	69,590	69,669	*10.94	*2.62
Hypertension	24,919	25,003	21,649	21,937	*2.03	*4.27
Acute upper respiratory disease	31,076	124,218	22,287	88,383	*17.05	*33.91
Common cold	93,143	93,143	43,486	43,486	*15.41	...

defined across surveys—there might have been more noteworthy differences between estimates due to definitional factors.

Though the findings are equivocal, they point to the need for a standardized definitional scheme across major household-based health surveys. This would eliminate inadvertent effects of definitional differences on incidence and prevalence estimates.

Conditions Miscoded as Chronic

Coding of Chronic Conditions in the Two Surveys

Although according to both NMCUES and NHIS, chronic conditions are defined, in part, as those having lasted 3 months or more, there were some important differences between the two studies. In particular, NMCUES used two procedures for coding conditions as chronic. Some conditions were coded chronic "by definition"; that is, they were considered chronic regardless of date of onset. Conditions coded chronic by definition include asthma, cancer, cirrhosis of the liver, and diabetes. (See Appendix I for the complete list of conditions coded chronic regardless of date of onset.)

Other conditions were coded chronic if the respondent reported the date of onset as having occurred on or before the 3-month reference period prior to the date of interview. The references to these conditions as chronic are with respect to the 3-month convention; that is, if the date of onset subtracted from the date of interview was equal to or greater than 90 days (3 months), the condition was coded chronic.

This coding tactic produced a miscoding of conditions as chronic because some of the conditions that began on or before the first day of the reference period also *ended* prior to the interview. However, the NMCUES questionnaire did not contain any item which would pinpoint whether the condition was arrested, inactive, or cured. So, for example, it is impossible to differentiate between a respiratory ailment that began a day prior to the reference period and lasted for its entirety and one that began a day prior to the reference period and lasted 3 days. The respiratory ailment in the first example is clearly chronic and appropriately coded using the 3-month convention. The respiratory ailment in the second example was acute but miscoded as chronic. In short, conditions coded chronic by the 3-month convention include conditions that are truly chronic as well as conditions that were actually acute but began 3 months prior to the interview.

Analysis: The 3-Month Convention

Because it is impossible to determine with precision the number of acute and miscoded chronic conditions, this analysis was designed to determine the maximum level of distortion that could have been introduced by the coding of conditions as chronic using the 3-month convention. The approach is a process of elimination; that is, the systematic exclusion of conditions that would not be subject to this error.

The analysis is limited to the 16,207 (94.6 percent) key people for whom there is information from all rounds in which they were eligible, and who reported one or more conditions during the survey. Altogether, the 16,207 respondents had a total of 47,680 acute and chronic conditions, excluding dummy records. After the exclusion of conditions that were not assigned an ICD–9 code and conditions assigned a combination of acute and chronic codes, a total of 47,523 (99.7 percent) conditions remained.

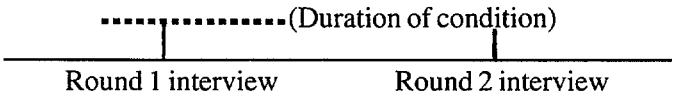
The first step in the analysis was to eliminate all acute conditions, including pregnancies (conditions with a prefix "2"), because, by definition, these conditions were not falsely classified as chronic. This eliminated another 23,994 conditions. Excluded next were conditions defined as impairments or chronic regardless of date of onset, because these would not be subject to miscoding. Subtracting acute conditions and conditions considered chronic by definition eliminated a total of 24,525 (51.6 percent) conditions.

Of the remaining 22,998 conditions (48.2 percent), conditions with a date of onset prior to November 1, 1979 were deleted. This deletion was done because conditions that began 3 months or more before February 4, 1980 (the date of the first interview) and entailed medical utilization or expenditures during the round 1 data-collection period between February and April were likely to be genuinely chronic. The exclusion of conditions that began 3 months or more prior to February 1980 eliminated another 15,829 conditions. This exclusion left 7,169 (15 percent) conditions.

Finally, conditions reported in more than one round (recurring conditions) also were considered to be legitimately coded chronic, and those conditions were deleted as well. This procedure eliminated another 6,040 conditions (12.7 percent), leaving 1,129 conditions, or approximately 2.4 percent of all conditions coded chronic by the 3-month convention.

It must be noted that some of the conditions reported in two adjacent rounds in fact might have been acute conditions miscoded as chronic. A condition could have had its date of onset prior to a round—for example round 1. It then would be reported in round 1, continue

into part but not all of round 2, and be reported in round 2 as well, as shown by the illustration below.



Thus, some portion of the conditions that appear in more than one round might be acute conditions miscoded as chronic. However, the assumption is that the majority of conditions reported in adjacent rounds were genuinely chronic. Of the 6,040 conditions that were reported in more than one round, 2,836 (46.9 percent) were reported in two (though not necessarily adjacent) rounds. If it is assumed that all conditions reported in two rounds were falsely classified as chronic, still only 12 percent of all chronic conditions would have been miscoded.

In sum, even if all conditions reported in one or two rounds really were acute conditions miscoded as chronic, still no more than 15 percent of all chronic conditions would have been coded falsely. Again, it is important to remember that the 15-percent figure represents the maximum level of error; that is, the estimate includes conditions that were legitimately coded as chronic using the 3-month convention as well as acute conditions miscoded as chronic.

Nevertheless, it is critical that future studies similar to NMCUES include a question about the termination of an illness, paralleling NHIS procedure. In the condition section, NHIS respondents are asked, "Does (PERSON) still have this condition?"; "Is this condition completely cured or is it under control?"; and "About how long did (PERSON) have this condition before it was cured?" (See items 16 c, d, and e in section AA of the 1980 NHIS questionnaire, Appendix IV.) The addition of these or similar questions would reduce substantially the miscoding of conditions as chronic because they would provide the analyst with information that could be used to determine the duration of an illness or disease. The analyst then could assess more accurately the impact of the 3-month recall period and differentiate between acute and chronic conditions.

Multiple Assignment of ICD–9 Codes

Assignment of Codes in the Two Surveys

In NHIS only one ICD–9 code was allowed to be assigned to a single condition, whereas in NMCUES up to three codes were allowed. This section examines the proportion of conditions that were multiply coded in NMCUES and the possibility that there is a pattern to the multiple coding of conditions. In other words, if certain types of morbidity conditions received more than one condition code, then incidence and prevalence estimates for some categories would be inflated. All else equal, NHIS and NMCUES incidence and prevalence estimates should be comparable for conditions that received only one code.

Analysis: Patterns in Multiple Coding

Of the 49,137 conditions included in the analysis of the NMCUES data, 47,163 (96.0 percent) received one diagnostic code; 1,580 (3.2 percent) conditions were assigned two diagnostic codes; and 394 (0.8 percent) conditions received three diagnostic codes. In short, fewer than 5 percent of all conditions in NMCUES received more than one diagnostic code, and conditions that were multiply coded tended to be assigned two ICD codes rather than three.

As shown in Table D, a greater proportion of chronic than acute conditions were multiply coded, though the differences are rather small. Of the 24,494 chronic conditions, 1,110 (4.5 percent) were multiply coded; this figure probably reflects the disproportionate number of impairments that were multiply coded. The 550 multiply coded acute conditions represent 2.3 percent of all acute conditions.

Table D

Number and percent distribution of NMCUES morbidity conditions, according to number of diagnostic codes assigned to each type of condition

Number of diagnostic codes	Type of condition			
	Chronic	Acute	Unknown	Total
	Number			
Total	24,494	24,073	570	49,137
1	23,384	25,523	256	47,163
2	984	484	112	1,580
3	126	66	202	394
	Percent distribution			
Total	100.0	100.0	100.0	100.0
1	95.5	97.7	44.9	96.0
2	4.0	2.0	19.6	3.2
3	.5	.3	35.4	.8

There is no clear pattern to the multiple assignment of codes, when spread across the 17 major morbidity categories of the ICD–9 (World Health Organization, 1977; see Table E). One-third (620) of the multiple codes were impairments (a category unique to NHIS and NMCUES, and a departure from standard ICD–9 coding conventions), and another 20 percent involved injuries and poisonings. It is not clear, however, why these two categories received more than one code. The other half of the multiply coded conditions were spread across the remaining 16 morbidity categories. Of the 1,974 conditions assigned more than one diagnostic code, 149 (7.5 percent) were treated as missing data because they were assigned a single digit or alphanumeric code not used in NHIS or ICD–9 coding procedures. In short, approximately half of all multiply coded conditions involved impairments, and injuries and poisonings. Except for those two categories, the overall impact of multiple coding on incidence and prevalence estimates was negligible. In general, only 1 to 2 percent of all conditions were multiply coded; 7.6 percent of injuries and poisonings, and 16.8 percent of impairments received a multiple code.

The NMCUES hospital file was examined for a possible relationship between hospital stays and morbidity conditions involving complications that would require the use of multiple codes. Only 231 (12.6 percent) of the 1,835 multiply coded conditions were in the hospital file; an average number of multiply coded conditions were spread evenly across each of the nine sections of the questionnaire where conditions were reported. Only 142 conditions involved both impairments and injuries and poisonings, indicating that the two categories generally were not coded together for the same condition.

The infrequent use of multiple coding indicates that it was not intended to supersede the assignment of a single diagnostic code and should perhaps be abandoned until a rationale for its continued use is developed. At this point, it only obscures comparisons between estimates from NMCUES and other health surveys.

Table E

Number and percent distribution of multiply coded conditions by the major morbidity categories

Morbidity category	Number of codes	Percent distribution of multiply coded conditions in each category	Percent of each category that was multiply coded
Total	1,825	100.0	...
Infectious and parasitic diseases	51	2.8	1.7
Neoplasms	23	1.3	3.5
Endocrine, nutritional, and metabolic diseases, and immunity disorders	62	3.4	4.6
Diseases of the blood and blood-forming organs	4	0.2	1.5
Mental disorders	40	2.2	3.9
Diseases of the nervous system and sense organs	145	7.9	2.8
Diseases of the circulatory system	75	4.1	2.2
Diseases of the respiratory system	164	9.0	1.1
Diseases of the digestive system	58	3.3	2.1
Diseases of the genitourinary system	126	6.9	5.4
Complications of pregnancy, childbirth, and the puerperium	4	0.2	1.5
Diseases of the skin and subcutaneous tissue	36	2.0	1.9
Diseases of the musculoskeletal system and connective tissue	41	2.2	1.7
Congenital anomalies	5	0.2	6.1
Conditions originating in the perinatal period	([1])	([1])	([1])
Symptoms, signs, and other ill-defined conditions	20	1.1	0.6
Injury and poisonings	351	19.2	7.6
Impairments	620	34.0	16.8

[1]No multiple codes appear in this category.

Adaptations of the *Medical Coding Manual*

Adaptations

Although most coding procedures used in NHIS were followed in NMCUES, certain other procedures were adapted, listed in the annotations to the *Health Interview Survey Medical Coding Manual and Short Index* (NCHS, 1983). In all, a total of 58 annotations to the NHIS medical coding manual (October 1979) were developed in NMCUES. Of the 58 annotations, 33 were excluded from the analysis because they did not affect the type of ICD–9 code assigned, or the morbidity estimates, which are the main concern of this comparative analysis.

This section examines the impact of the NMCUES annotations on estimates of morbidity conditions. The discussion should provide interested users of NMCUES condition data with a frame of reference for comparisons between NMCUES and NHIS incidence and prevalence estimates and with recommendations as to the usefulness of particular annotations.

Analysis: The NMCUES Annotations

The NMCUES annotations introduced three new kinds of coding procedures.

1. They altered the rules for coding conditions of unknown duration and for those vaguely reported.
2. They coded specific diseases differently.
3. They treated multiple conditions differently in two cases (for hospital admissions and in the rules for combining and merging conditions).

A detailed discussion of specific instances of these kinds of annotations follows.

1. Alterations in Rules for Coding Conditions of Unknown Duration and for Those Vaguely Reported

a. Use of the prefix "3," indicating a condition with unknown onset—Three of the eighteen annotations introduce the use of a prefix "3" to a diagnostic code and instruct coders on its use. As mentioned earlier, in both NHIS and NMCUES a prefix indicating whether a condition was chronic or acute was used. Chronic conditions received a prefix of "1," and acute conditions were assigned a prefix "2." Sometimes, however, it was not clear when a condition began, and, therefore, conditions could not be designated as chronic or acute with certainty.

This problem was resolved in NHIS by coding conditions of unknown onset as acute, and assigning a prefix "2." In NMCUES, however, the prefix "3" was used to indicate ambiguity about the onset of the condition. In all, 268 or 0.5 percent of all conditions were designated as having an unknown date of onset. Two-thirds of these conditions were injuries and poisonings, diseases of the respiratory system, or part of the category defined by the ICD–9 as symptoms, signs, and ill-defined conditions. Still, so few conditions were coded as having an unknown onset that the effects of the procedure on estimating incidence and prevalence, even for these categories, are negligible. Overall, the infrequent use of the prefix "3" can be interpreted as evidence that NMCUES interviewers had little difficulty ascertaining accurate information about the date of onset of a condition. Little is gained by the use of the prefix "3," and its absence is not likely to distort analyses of acute and chronic conditions.

b. Use of additional codes for symptoms, signs, and ill-defined conditions—ICD–9 codes 780–796 in NHIS and 780–799 in NMCUES were used to code symptoms, signs, and ill-defined conditions. The additions in NMCUES include codes for senility without mention of psychosis, sudden death where cause is unknown, and other ill-defined and unknown causes of morbidity and mortality. Overall, the difference between the NHIS and NMCUES estimates was substantial. The NHIS estimate for ill-defined conditions was 17 million, whereas the NMCUES estimate, based on the expanded code range, was 38 million conditions. Even when the same code range (780–796) was used, the NMCUES estimate of 26 million remained substantially greater than the NHIS estimate. Although the reason for this discrepancy is not clear, it could be that NHIS merged such ill-defined conditions with ones that were less ambiguously described.

2. Coding Differences for Specific Conditions

a. Tonsils and adenoids—Different codes were used in NHIS and NMCUES to designate whether tonsil and adenoid conditions were chronic or acute. Chronic tonsil and adenoid conditions were coded 474.0–474.9 in both NHIS and NMCUES, and 465 when they were acute. However, tonsils and adenoids were included on an NHIS

checklist, and the chronic codes were used if the condition required hospitalization, regardless of whether surgery was performed. In NMCUES chronic codes were used only if surgery was performed. As a result, the prevalence of chronic diseases of tonsils and adenoids is larger in NHIS (2.3 million cases) than in NMCUES (1.5 million cases).

b. Pregnancies and deliveries—Three of the five annotations to NHIS coding procedures in this area pertained to pregnancy, and two referred to the coding of deliveries. In NMCUES pregnancies were coded differently than in NHIS in several ways. First, in NHIS pregnancies were not coded as conditions; they were coded as "AAA" in NMCUES to collect information on costs associated with pregnancy. In all, 399, or less than 1 percent, of the 49,137 conditions in NMCUES were pregnancies.

The second and third annotations pertaining to pregnancy allowed NMCUES coders to assign more than one code within the 630–676 range. NHIS coders were instructed to use only one code within this range, preferring to describe "delivery" or "abortion" instead of other conditions present. In addition, NHIS coders were directed to avoid coding complications of pregnancy (conditions within the 630–676 range) unless the complication existed prior to the pregnancy. In such cases, the condition complicating the pregnancy was coded as it was for nonpregnant women. In NMCUES coders were allowed to use the 630–676 code range, and multiple coding was permitted. Together, these conventions should produce an increased use of these codes in NMCUES compared with NHIS. However, the NMCUES estimate for complications related to pregnancy was 3.3 million, comparable with the NHIS estimate of 3.4 million.

Estimates for complications involving delivery also were affected by the different coding instructions in the two surveys. In NHIS an allowance was made for only one diagnostic code to be used per admission for parturition; the survey design gave preference to the application of the "delivery" (650–669) or "abortion" (630–639) codes over others. Because in NMCUES multiple coding was permitted and coders were allowed to use codes that NHIS coders did not, the NMCUES estimate for complications of delivery (code 670–676) should be greater than the NHIS estimate. As expected, the NMCUES estimate (676,000 cases) was larger than the NHIS estimate (209,000 cases).

The fifth annotation also referred to differences between NHIS and NMCUES coding of deliveries with complications. In both surveys complications of deliveries were coded 660–669. The NMCUES estimate for deliveries with complications was 423,000; the NHIS estimate was 128,000.

c. Epilepsy—Three of the eighteen annotations modified the coding of epilepsy related or due to injury. NHIS coders were instructed to assign one of nine digits that appear after the decimal point in the diagnostic code 345 (that is, 345.0–345.9). The three annotations instructed NMCUES coders to use only the 345.9 code for unspecified forms of epileptic convulsions, fits, or seizures not elsewhere classified. Although this change would not affect the number of cases of epilepsy reported in each study, the prevalence of epilepsy in NMCUES (1 million cases) is half that of the NHIS estimate (2.6 million cases). The greater prevalence of epilepsy in NHIS can probably be explained by its presence on an NHIS checklist that includes conditions affecting the nervous system.

d. Circumcision—No codes for circumcision appear in either NHIS or NMCUES data. This is rather curious because specific instructions for coding circumcision were given in both surveys. NHIS coders were instructed to code circumcision only if the patient underwent other surgical procedures at the same time. In NMCUES those instructions were deleted so circumcision was not coded at all.

3. Treatments of Multiple Conditions

a. Use of multiple diagnoses for a single hospital admission—Because in NHIS only one diagnosis was allowed to be coded per hospital admission, the number of hospitalizations equals the number of conditions associated with a hospital stay, and the number of diagnostic codes equals the number of conditions associated with a hospital stay. Surgical operations performed in the hospital are an exception, because in NHIS up to three conditions were allowed to be coded for each operation.

By comparison, in NMCUES up to four conditions were allowed to be entered for a single admission to the hospital, and three codes per condition. Thus a greater number of conditions associated with each hospital stay would be expected in NMCUES than in NHIS. In all, there were 3,150 conditions associated with the 2,671 hospital stays in NMCUES. Of the 2,659 hospital stays with legitimate ICD–9 codes, 2,268 (85.3 percent) were associated with one condition; 308 (11.6 percent) were associated with two conditions; and 83 (3.1 percent) were associated with three or four conditions. Thus, in NMCUES as in NHIS, most hospital stays received a single diagnostic code, but a fair proportion were associated with multiple conditions.

b. Use of combining and merging—Procedures to consolidate closely related conditions were developed for NHIS. The procedures are used to curtail overcoding, which occurs if symptoms or diseases are treated as separate conditions when, from a clinical standpoint, they should be considered part of a single condition. For example, if a respondent reported having fever, diarrhea, and nausea, a single diagnostic code for stomach flu might be assigned to the separately reported symptoms.

In both NHIS and NMCUES information about conditions was allowed to be consolidated *if that information*

was reported on the same condition page of the questionnaire. Such consolidation is called "combining."

However, only NHIS permitted conditions to be "merged"; that is, only in NHIS was information about conditions allowed to be consolidated *when the data were reported on separate condition pages of the questionnaire*. For example, in NHIS a sore throat and cold reported on separate condition pages could be assigned a single diagnostic code. Also the "merging" of disability days was permitted in NHIS; that is, if both conditions entailed disability days, the condition with a larger number of reported disability days became the figure associated with the single "merged" condition. Similarly, the earlier of the two dates of onset associated with each condition became the date of onset for the "merged" condition, and a medical visit associated with one of the conditions became an additional datum associated with the merged condition.

By comparison, in NMCUES the two conditions were left as distinct because they were reported on separate condition pages; they were not collapsed into a single diagnostic code, nor was ancillary information associated with each condition consolidated.

The decision not to merge conditions in NMCUES reflects differences between the primary purposes of the two studies. The key goal in NHIS was to collect information that was as accurate as possible about *medical conditions*, whereas in NMCUES this was secondary to collecting information on charges and care associated with medical events, including conditions. In other words, to preserve a maximum amount of information associated with *each* reported medical condition, the consolidation of information about conditions and events associated with them was avoided in NMCUES.

The question, however, is whether the advantages of repeated interviewing are lost by failing to implement merging procedures, especially across rounds. Unlike cross-sectional surveys, longitudinal surveys offer the opportunity to consolidate conditions and information about them that through time are shown to be genuinely part of other conditions (for example, the stomach ulcer that subsequently proves to be cancer of the stomach). One way to check the extent to which conditions were overcoded by deleting merging procedures in NMCUES would be to take certain morbidity conditions that generally are merged, such as coronary and hypertensive diseases, and merge them. The level of merging within rounds should be less than or equal to the level in NHIS, and the level of merging between rounds would indicate the degree to which "undermerging" occurs in NHIS because interviews are conducted only once.

In failing to implement merging procedures, especially across rounds, the NMCUES design failed to take advantage of the unique opportunity provided by multiple interviews to see the development of information over time. Again, a stomach ulcer could develop into cancer of the stomach, and in NMCUES, ultimately, the more accurate diagnosis as well as the total charges involved in treating that condition could be identified.

The most practical and prudent time to merge conditions is during the final interview at the end of the study. At that point, the interviewer could ask respondents to review a computer-generated summary of all conditions reported for each member of the household during the survey period and ask whether any of the conditions reported during earlier interviews were shown subsequently to be part of the same condition. Merging conditions at that time minimizes the risks of consolidating two conditions that prove later to be unrelated or finding out that a condition is really related to a different condition reported in a later round of the survey.

In sum, the overall impact of deleting merging procedures in NMCUES is likely to be small, given that only 3 percent of all conditions were combined or merged in NHIS. It is possible that some conditions were more likely than others to be combined or merged, and larger estimates of incidence and prevalence for those conditions in NMCUES would be expected relative to NHIS. However, there is no theoretical reason to delete merging rules in future studies similar to NMCUES, and, in fact, the incorporation of merging techniques at the end of the study could improve the accuracy of estimates of illness and disease and the associated charges.

Summary and Recommendations

The findings of this study confirm the utility of a number of procedures and instruments used respectively in NHIS and NMCUES. The results also point to possible ways of integrating and adapting some of those methodologies in subsequent health surveys of a longitudinal nature such as NMCUES. This summary of findings provides a rationale for the proposed recommendations.

NMCUES and NHIS differed in their estimates of different types of conditions. In most cases, NMCUES estimates were lower than NHIS estimates. This trend is surprising in view of the features of NMCUES known to boost the reporting of conditions; that is, repeated interviewing and an expanded questionnaire. As noted earlier, repeated interviewing did substantially boost estimates yielded from single interviews but did not fully offset the use of a checklist.

Length of recall, too, seems to have affected the reporting of conditions and involved a trade-off in the two studies between the reporting of acute and chronic conditions. The relatively short, 2-week recall period in NHIS tended to produce a high level of reporting of acute conditions and an underreporting of chronic conditions not on the checklist. Extension of the recall period to 3 months in NMCUES produced an underreporting of relatively minor, acute conditions and a relatively high level of reporting chronic conditions. Through use of a checklist, underreporting of some chronic conditions is adjusted for in NHIS. On the other hand, underreporting of acute conditions is not adjusted for in NMCUES. By implication, the charges associated with these conditions are also missing from the findings of the NMCUES survey. A checklist of commonly underreported acute conditions could compensate for the longer recall period in longitudinal studies similar to NMCUES.

The ability to accurately designate a condition as chronic or acute in NMCUES was undermined by the inability to establish the duration of an illness. Possibly 15 percent of conditions in NMCUES could have been miscoded as chronic when they were actually acute. The findings suggest that in the future explicit questions should determine whether a particular condition is ongoing, brought under control, or cured. Nevertheless, the findings suggest that standard definitions for morbidity categories across major health surveys (or at a minimum, between NHIS and NMCUES) would provide an especially useful check against relative overreporting and underreporting of conditions between surveys.

The analysis of the 58 annotations made to NHIS Medical Coding Manual (National Center for Health Statistics, 1979) revealed that the majority of annotations had little impact on the number or kinds of diagnostic codes assigned. Three types of changes, however, are worth commenting on: use of the prefix "3," assignment of more than one diagnostic code to a condition, and deletion of merging rules.

The innovation in NMCUES that designated conditions with an unknown date of onset by a special prefix "3" seems not to have made any difference. Fewer than 1 percent of all conditions were coded "3," and therefore the practice of such coding could be abandoned in the future. Adoption of the NHIS practice of coding these conditions as acute would introduce little if any distortion and could be substituted for the new NMCUES procedure.

The NMCUES design also changed the NHIS coding framework by permitting up to three ICD–9 codes to be assigned to each condition. Because the multiple coding had no clear rationale nor a significant effect on estimates, this practice could also be abandoned, except for the multiple coding of operations, a practice already followed in NHIS.

The absence of merging rules in NMCUES probably inflated the incidence and prevalence of conditions reported, and produced a 3- to 5-percent level of overcoding. Though these differences may not be major, there are other reasons for suggesting that merging rules could be applied appropriately to NMCUES. Given the opportunity provided by repeated interviews over time, merging in NMCUES could enhance the accuracy of condition data, especially if merging were implemented after the completion of the last interview.

Recommendations that flow from this evaluation of NMCUES medical condition data collection and coding methods could be summarized as follows.

Abandon:

- Special code for chronic conditions of undetermined origin.
- Multiple coding of conditions, except in the case of hospital admissions and surgical operations.

Retain:

- Expanded questionnaire developed in NMCUES.
- Repeated interviews.

Add:

- Checklist of generally underreported acute conditions.
- Checklist of generally underreported chronic conditions.
- Questions about the termination of an illness.
- Merging procedures, as specified above.
- Standard definitions of morbidity categories, usable across health surveys.
- Clearly delineated rationale for any and all procedures undertaken and instruments used.

References

Bean, J. A.: Estimation and sampling variance in the Health Interview Survey. *Vital and Health Statistics*. Series 2, No. 38. PHS Pub. No. 1000. National Center for Health Statistics, Public Health Service. Washington. U.S. Government Printing Office, June 1970.

Bonham, G. S.: Procedures and questionnaires of the National Medical Care Utilization and Expenditure Survey. *National Medical Care Utilization and Expenditure Survey*. Series A, Methodological Report No. 1. DHHS Pub. No. 83–20001. National Center for Health Statistics, Public Health Service. Washington. U.S. Government Printing Office, Mar. 1983.

Deming, W.: On simplification of sampling design through replication with equal probabilities and without stages. *JASA* 51:24–53, 1956.

Frankel, M.: *Inference from Survey Samples, An Empirical Investigation*. Ann Arbor. Institute for Social Research. University of Michigan, 1971.

Jack, S. S.: Current estimates from the National Health Interview Survey, United States, 1980. *Vital and Health Statistics*, Series 10, No. 139. DHHS Pub. No. (PHS) 82–1567. National Center for Health Statistics, Public Health Service. Washington. U.S. Government Printing Office, Dec. 1981.

Kovar, M. G., and Poe, G. S.: The National Health Interview Survey design, 1973–84, and procedures, 1975–83. *Vital and Health Statistics*. Series 1, No. 18. DHHS Pub. No. (PHS) 85–1320. National Center for Health Statistics, Public Health Service. Washington. U.S. Government Printing Office, Aug. 1985.

National Center for Health Statistics: Health survey procedure, concepts, questionnaire development, and definitions in the Health Interview Survey. *Vital and Health Statistics*. Series 1, No. 2. PHS Pub. No. 1000. Public Health Service. Washington. U.S. Government Printing Office, May 1964.

National Center for Health Statistics: Health interview procedure, 1957–74. *Vital and Health Statistics*. Series 1, No. 11. DHEW Pub. No. (HRA) 75–1311. Health Resources Administration. Washington. U.S. Government Printing Office, Apr. 1975.

National Center for Health Statistics: *Health Interview Survey Medical Coding Manual and Short Index*. Public Health Service. Hyattsville, Md., Jan. 1979.

National Center for Health Statistics: *National Health Interview Survey 1980 Public Use Data Tape Documentation—Part I.* (Includes annotations to the *Health Interview Survey Medical Coding Manual and Short Index*). Public Health Service. Hyattsville, Md., June 1983.

National Medical Care Utilization and Expenditure Survey: *Household Survey and State Medicaid Household Survey Coding Procedures Manual, Round 1*. Research Triangle Institute. Research Triangle Park, N.C., 1980.

Simmons, W. R.: The statistical design of the health household interview survey. *Health Statistics*. PHS Pub. No. (HRA) 75–1311. Health Resources Administration, Public Health Service. Washington. U.S. Government Printing Office, Apr. 1975.

Tourangeau, R., McWilliams, H., Jones, C., et al.: *High School and Beyond: First Follow-Up Sample Design Report*. Chicago. NORC, May 1983.

Tourangeau, R., and Rasinski, K. A.: Evaluation of data collection frequency and the use of a summary in the National Medical Care Utilization and Expenditure Survey. *National Medical Care Utilization and Expenditure Survey*. Series A, Methodological Report No. 2. DHHS Pub. No. 87–20002. National Center for Health Statistics, Public Health Service. Washington. U.S. Government Printing Office, May 1987.

World Health Organization: *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death*. Vol. 1. Based on the Recommendations of the Ninth Revision Conference, 1975. Geneva. World Health Organization, 1977.

World Health Organization: *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death*. Vol. 2. Based on the Recommendations of the Ninth Revision Conference, 1975. Geneva. World Health Organization, 1978.

Appendixes

List of Appendix Figures

Appendix I. Selected Chronic Conditions

Selected chronic conditions regardless of date of onset:

Condition	Code
Absence	X32, X32.9, X33, X90, X26, X27, X28, X29, X35, X23, X20, X25, X22, X24, X21, X92, X34, X31, X11, X30, X92.9, X11.9, 757.4, 703.8, 525.1, 873
Alcoholism	303
Allergy (*except* cases with onset in past 3 months *and* due to drugs (995.2), bee sting (989.5), venomous bites (such as snake and spider) (989.5), chemicals (989.9), procedures in 996–999, to contactants (including sunburn) in 692 or substances taken internally (693) or radiation (990))	477 478.8 691, 692 693.1
Arteriosclerosis	440
Arthritis	711–716, 721
Asthma	493
Bronchiectasis	494
Calculi	592, 594
Cancer	140–239
Cardiac conditions	402, 410–429
Cataract	366
Cerebral palsy	X50, 343, X50.9, X51–X64
Cerebrovascular disease	430–438
Cirrhosis of liver	571.2, 571.5
Clawfoot	X78.9
Cleft palate	X91.9
Clubfoot	X78.9
Color blindness	368.5, 368.55, 368.59
Congenital condition	744.0–744.3
Coronary condition	410–414
Cyst	733.2, 706.2, 610, 611.5, 685
Deafmutism	X05
Detachment of retina	361
Diabetes	250
Drug addiction or dependence	304
Emphysema	492
Epilepsy	345
Flatfoot	X77
Glaucoma	365
Goiter	240–242
Gout	274
"Growth"	140–239
Harelip	X91.9
Hay fever	477
Heart or cardiac disease	402, 410–429
Hemeralopia	368.1
Hemorrhoids	455
Hernia	550–553
Hypertension	401–405
Loss	X11, X12
Mental deficiency	X19
Mental disorders	300–306 (except 305.0), 310, 312–316
Mole	236.1, 631
Mongolism	X19.9
Multiple sclerosis	340
Neoplasm	140–239
Neuroses	300
Nyctalopia	368.6
Optic nerve disorders	377.0–377.2, 377.4–377.9
Paralysis agitans	332
Personality disorders	301
Polyps	622.7, 385.3, 478.4, 471.0, 471.9, 471.8, 620.8, 569.0, 621.0, 478.4
Prostate condition	600–602
Psychosis	290–299
Refractive errors	367
Retardation	X19
Retinal conditions	361, 362, 363.3
Retrolental fibroplasia	362.2
Rheumatic fever	390–398
Rheumatism	725, 726.0, 726.2, 729.0, 729.1
Rupture	550–553
Specific	X14
Stones	592–594
Stroke	430–435, 437
Thyroid	240–246
Trick knee	X86
Tuberculosis	010–019
Tumor	140–239
Ulcer of stomach	531–534
Varicose veins	454–456

ICD–9 codes that are included in the NHIS chronic condition edit but are not included in the chronic condition list in the NHIS *Medical Coding Manual and Short Index* (National Center for Health Statistics, 1979) are listed below.

Condition	Code
Deficiency of cell-mediated immunity	279.1
Fragments of torsion dystonia	333.8
Migraine	346
Eye disorders	360.1, 360.2, 360.4, 363.4,

Condition	Code
Eye disorders—con.	363.5, 364.7, 369.9, 370.2, 370.3, 371.3, 372.0–372.3, 372.6–372.7, 373.1, 373.3, 376.4, 378.7, 379.2, 379.5
Diseases of esophagus	530
Gastritis	535
Other noninfective gastro-enteritis and colitis	558
Disease of musculoskeletal system	710.3–710.4
Neck contracture	723.5, X79

Appendix II. NMCUES Annotations to the NHIS Medical Condition Coding Manual

NMCUES annotations to the NHIS Medical Condition Coding Manual that do not affect ICD–9 code assignment are as follows:

Page	Item	Change	Effect
		Section I: Introduction and Orientation Guide	
3	F.3	Change "Hosp. Page" to Hospital Stays Section.	In NMCUES a Hospital Stays Section was used rather than a Hospital Page, as was used in NHIS.
		Section III. General Coding Principles and Problems	
10	—	First paragraph, line 3, delete "for hospital stays and."	This instruction refers to rules for ICD–9 code assignment. The deletion of this phrase does not itself produce changes; rather, it is consistent with deletions listed in Section VIII, "Hospitalization and Surgery."
12	D	First paragraph, lines 1 (pages 31–33) and 2, delete "and merging."	Merging rules were not used in NMCUES. See text for discussion of the effects of the absence of merging rules.
12	D	First paragraph, lines 3 and 4, after "code" insert a period (.) and delete "and one Condition Page."	The phrase refers to merging rules, which were not used in NMCUES. See text for discussion of the effects of the absence of merging rules.
13	F	Paragraph 3, delete "if" and substitute "is."	Typographical error (with no coding implications).
13	"Active" F.1	Line 1, change "3a" to "condition."	This instruction refers to the place on the questionnaire where the condition is entered. On NHIS it appears in 3a; in NMCUES, it appears on the space labeled "condition."
		Lines 4 and 5, delete "AA column filled... in item 16d;".	There was no NMCUES question in reference to whether a condition was completely cured or under control, whereas such a question was included in NHIS in Q.16d. See text for discussion of this deletion.
		Line 5, change "(2)" to "(1)."	This renumbers NHIS items used in NMCUES.
		Line 7, change "(3)" to "(2)."	This renumbers NHIS items used in NMCUES.
	"Active" F.2	Line 1, change "3b or 3c" to "2 and 3."	This renumbers NHIS items used in NMCUES.
	"Inactive" F.1	Line 1, change "3a" to "condition."	This renumbers NHIS items used in NMCUES.
	F.2	Line 1, change "3b or 3c" to "2 and 3."	This renumbers NHIS items used in NMCUES.
15	G.6	Delete paragraph 3, "Interviewer writes a... multiple symptoms."	This deleted instruction refers to procedures used in NHIS for merging conditions. In NMCUES conditions were not merged. See text for discussion of the effects of the absence of merging rules.
		Paragraph 4, line 1, change "Questionnaire" to "Condition Page."	This instruction refers to changes in the location of multiple symptoms of unknown cause between the NHIS and NMCUES questionnaires.
18	I.1	In the title, line 2, change *"deleted"* to "coded BD (bad data)."	In both NMCUES and NHIS this information is excluded from analysis. It is deleted altogether in NHIS; it is coded BD (bad data) in NMCUES.
	I.1.a	Paragraph 1, line 1, change "Delete the Condition Page" to "code bad data BD."	This makes the change consistent with the instruction specified in Section III, p. 18, and Item I.1 above.
18	I.2	First paragraph under title, line 4, delete "A Hospital Page... is not required."	In NMCUES, like in NHIS, hospital information is not collected under these circumstances.
22	K.4	Paragraph 3, line 3, change *"not* be coded," to "be coded BD (bad data)."	In both NMCUES and NHIS, this information is excluded from analysis. It is deleted altogether in NHIS; it is coded BD (bad data) in NMCUES.

Page	Item	Change	Effect
		Section III: General Coding Principles and Problems—Con.	
23	0	Line 3, change "do not code," to "code BD (bad data)."	This follows same exclusionary rules observed in NHIS.
		Section IV: "Combining and Merging"	
24	A.1	Paragraph 2, lines 2–4, delete "; or it may involve... becomes in order," and insert a period (.) after "Page."	This deletion refers to merging rules. See text (pages 31–33) for discussion of the effects of the absence of merging rules.
	A.2	Line 4, delete "or on separate Condition Pages."	This deletion refers to merging rules. See text (pages 31–33) for discussion of the effects of the absence of merging rules.
25	A.2.b	Lines 4–6, delete "The conditions may also... being 'same as'."	This deletion refers to merging rules. See text (pages 31–33) for discussion of the effects of the absence of merging rules.
		Lines 8–12, delete "However, if heart... also (440.9)."	This deletion refers to an example of merging rules. See text (pages 31–33) for discussion of the effects of the absence of merging rules.
	A.2.d	Last paragraph, lines 9 and 10, delete "These cannot be... separate Condition Pages."	This instruction orders merging rules not to be used to combine dagger-asterisk codes.
26	B.1–4	Delete.	This section describes merging procedures used in NHIS. See text for discussion of the effects of the absence of merging rules.
		Section V: Impairments and Their Causes	
36	G.1.b	Last paragraph, last sentence, delete "Prefer the entry...."	Deletion refers to determining whether one or both ears are impaired. In Q.4A of the NMCUES Condition Section are questions regarding whether right, left, or both ears are affected.
		Section VI: Injuries, Accidents, and Their Effects	
51	B	Delete.	NHIS codes indicating whether a condition was first injury, required hospitalization, or involved an adverse reaction were not used in NMCUES.
	D.2	Delete.	In both NMCUES and NHIS chronic conditions exacerbated by an accident are recorded. In NMCUES the condition is recorded the first time and only the first time it is reported by the respondent. The most recent manifestation of the condition is recorded in NHIS. The only difference is *which* accident gets coded as having exacerbated the condition.
		Section VII: Conditions Related to Childbearing	
52	—	First paragraph, lines 4–7, delete "These conditions may be.... the Hospital Page."	This eliminates an instruction used in NHIS but not used in NMCUES. These conditions of the newborn are all located in the Hospital Stay Section in NMCUES, whereas in NHIS the conditions appear on different pages.
	C	Delete paragraph 2.	In both NHIS and NMCUES the operation is recorded as reported by the respondent.
53	F	Lines 11–13, delete "*See also* Section... other toxemias."	This annotation is consistent with the deletion of Item M in Section III, p. 22. The instruction to delete Item M is a paragraph concerning the coding of pregnancy with hypertension and other toxemias. NHIS coders are instructed to code hypertension and other toxemias when associated with pregnancy the same way they would for nonpregnant women. Also NHIS coders are directed by the instruction not to code hypertension or other toxemias for nonpregnant women who had reported such problems during pregnancy, but who no longer have the condition.
		Section VIII: Hospitalization and Surgery	
54	A	Line 3, delete "diagnostic and."	Diagnostic code assignments differ between NHIS and NMCUES; see text (pages 30–31) for effects. This annotation is consistent with deletion of NHIS procedures for assigning diagnostic codes.
54	A	Line 4, change "Page" to "Stay Section."	In NHIS a Hospital Page is used; in NMCUES a Hospital Stay Section is used.
56	E	In the title, delete "*Multiple Diagnoses.*"	The deletion maintains consistency with changes in Section VIII that delete diagnostic coding-assignment procedures.

Page	Item	Change	Effect
		Appendix III	
1	—	Line 3, change "Page" to "Stay Section."	A Hospital Stay Section is used in NMCUES whereas in NHIS a Hospital Page is used.
		Short Index	
2	—	Delete "First Injury Codes" paragraph.	NHIS codes that indicate whether a condition was a "first injury" are not used in NMCUES.
	—	Delete "Hospitalized for *this* accident" paragraph.	NHIS codes that indicate whether a condition required hospitalization are not used in NMCUES.
	—	Delete "Code for Hospital Page" paragraph.	In NMCUES, there is no assignment of a 1-digit code to indicate whether or not, and for what, surgery was performed during a hospital stay.

NMCUES annotations to the *Health Interview Study Medical Coding Manual and Short Index* that did change the proportion or type of diagnostic code assigned to a condition are as follows.

Page	Item	Change
		Section II: Classes of Chronic and Acute Conditions
6	A.2	Add "3—Unknown onset—."
9	E.2.b	Delete.
	E.3	Delete last paragraph.
		Section III: General Coding Principles and Problems
15	G.6.B	Line 2, change "796" to "799."
18	2.a.2	Change "Delivery" to "BD (bad data) unless there is a cause."
22	M	Delete entire paragraph.
		Section IV: "Combining and Merging"
24	A.1	Delete paragraph 3.
		Paragraph 2, lines 2–4, delete "; or it may involve... becomes in order." and insert a period (.) after "Page."
		Delete paragraph 3.
	A.2	Line 4, delete "or on separate Condition Pages."
25	A.2.b	Lines 4–6, delete "The conditions may also... being 'same as.'"
		Lines 8–12, delete "However, if heart... also 440.9)."
	A.2.d	Last paragraph, lines 9–10, delete "These cannot be... separate Condition Pages."
26	B.1–4	Delete.

Page	Item	Change
		Section VII: Conditions Relating to Childbearing
52	A	Delete paragraph 2.
53	2 and 3	Delete.
		Section VIII: Hospitalization and Surgery
54–55	B and C	Delete.
56	E.1.a and b	Delete.
		Appendix III
13	—	Change "EPILEPSY (345)" to "EPILEPSY (345.9)."
	—	Under "EPILEPSY," line 6, delete "by type in 345.0."
	—	Under "EPILEPSY," second paragraph, lines 4, 5, and 6, delete "345" and substitute "345.9."
25	—	Under "CHRONIC DISEASE OF TONSILS AND ADENOIDS," lines 2–3, delete "on Hospital Page only or on Condition Page and also on Hospital Page, with or without surgery" and substitute "tonsillectomy on Condition Page."
31	5	Delete.
		Short Index
2	—	Under "Acute/Chronic Code" add "3 = unknown (DK)."
38	—	After "Pregnancy-normal:" delete "Do not code" and substitute "AAA."

CONDITION SECTION - ASK ONLY ABOUT CONDITIONS ENTERED ON CONTROL CARD IN THIS ROUND -- (BELOW LAST REF. DATE)

PERSON NAME: _____ # _____ NAME OF CONDITION: _____ COND.#: _____ [ ] - [ ][ ][ ] . [ ]

You said earlier that (PERSON) had (CONDITION).

	CODE ONE AND FOLLOW INSTRUCTIONS	
A	Accident or Injury.	01 (7)
BOX	On Card K	02 (6)
	Neither	03 (1)

1. What did the doctor or other medical person say it was -- did he give (CONDITION) a medical name?

_____

Didn't see doctor. 01

2. What was the cause of (CONDITION)?

_____

Accident or injury 01 (7)

3. DO ANY RESPONSES IN Q's. 1 OR 2 INCLUDE AN ENTRY BELOW?
Yes. . . . . .01(A)
No . . . . . .02(4)

Ailment	Attack	Defect	Growth	Trouble
Anemia	Cancer	Disease	Measles	Tumor
Asthma	Condition	Disorder	Problem	Ulcer
	Cyst		Rupture	

A. What kind of (WORD) is it?

_____

4. ARE ANY RESPONSES IN Q's. 1-3 ALLERGY OR STROKE?
Yes. . . . . .01(A)
No . . . . . .02(5)

A. How does the [allergy/stroke] affect (PERSON)?

_____

5. DO ANY RESPONSES TO Q's. 1-4 INCLUDE AN IMPAIRMENT, PART OF BODY, OR ANY ENTRY BELOW?

Yes . . . . . 01(A)
No. . . . . 02(6)

Abcess	Cancer	Hemorrhage	Palsy	Tumor
Ache (except head or ear)	Cramps (except menstrual)	Infection	Paralysis	Ulcer
Bleeding	Cyst	Inflammation	Rupture	Varicose veins
Blood Clot	Damage	Neuralgia	Sore	Weak
Boil	Growth	Neuritis	Soreness	Weakness
		Pain	Stiff(ness)	

A. What part of the body is affected?

_____

_____

SHOW DETAIL IN Q.5A

HEAD.	SKULL, SCALP, FACE	LEG.	RIGHT, LEFT, OR BOTH; HIP, UPPER, KNEE, LOWER, ANKLE
BACK, SPINE, OR VERTEBRA . .	UPPER, MIDDLE, LOWER	HAND	ENTIRE HAND OR FINGERS ONLY; RIGHT, LEFT OR BOTH
EAR	RIGHT, LEFT, OR BOTH; OUTER, MIDDLE, INNER	FOOT	ENTIRE FOOT, ARCH, OR TOES ONLY; RIGHT, LEFT OR BOTH.
ARM	RIGHT, LEFT, OR BOTH, SHOULDER, UPPER, ELBOW, LOWER, WRIST	SIDE	RIGHT OR LEFT

6. When was the (CONDITION) first noticed by (PERSON) or a medical person?

_____ / _____ (10)
MONTH YEAR

Over 1 year ago. 01 (10)

IF ACCIDENT OR INJURY, ASK Q.'s. 7 THROUGH 9.

7. At the time of the accident, what part of the body was hurt? Any other part?

A. What kind of injury was it? Anything else?

7 Part(s) of body	A Kind of Injury

8. What part of the body is affected now? Any other part?

A. How is (PERSON'S PART OF BODY) affected? Is (PERSON) affected in any other way?

8 Part(s) of body	A Current Effect

9. When did the accident or injury occur?

MONTH / YEAR

Over 1 year ago 01

10. IS CONDITION AN EYE CONDITION? Yes . . . . . .01(A)
No. . . . . .02(NC)

A. Can (PERSON) see well enough to read ordinary newspaper print with glasses with [his/her] . . .

(1) left eye?

Yes. 01

No 02

(2) right eye?

Yes. 01

No 02

AFTER LAST CONDITION IS COMPLETED, GO TO HEALTH INSURANCE SECTION.

Appendix IV. NHIS 1980 Questionnaire: Condition Section

CONDITION 1

1. Person number | Name of condition

2. When did -- last see or talk to a doctor about his . . .?

1 ☐ In interview week (*Reask 2*)
1 ☐ Past 2 wks. (*Item C*)
2 ☐ 2 wks.–6 mos.
3 ☐ Over 6–12 mos.
4 ☐ 1 yr.
5 ☐ 2–4 yrs.
6 ☐ 5+ yrs.
7 ☐ Never
8 ☐ DK if Dr. seen
9 ☐ DK when Dr. seen

A1 Examine "Name of condition" entry and mark
☐ Color blindness (*NC*) ☐ On Card C (*A2*)
☐ Accident or injury (*A2*) ☐ Neither (*3a*)

If "Doctor not talked to," transcribe entry from item 1.
If "Doctor talked to," ask:

3a. What did the doctor say it was? – Did he give it a medical name?

Do not ask for Cancer ☐ On Card C (*A2*)

b. What was the cause of . . .?
☐ Accident or injury (*A2*)

If the entry in 3a or 3b includes the words:

Ailment	Condition	Disorder	Rupture	Ask c:
Anemia	Cyst	Growth	Trouble	
Asthma	Defect	Measles	Tumor	
Attack	Disease	Problem	Ulcer	

c. What kind of . . . is it?

For allergy or stroke, ask:

d. How does the allergy (stroke) affect him?

If in 3a–d there is an impairment or any of the following entries:

Abscess	Damage	Paralysis	Ask e:
Ache (except head or ear)	Growth	Rupture	
Bleeding	Hemorrhage	Sore(ness)	
Blood clot	Infection	Stiff(ness)	
Boil	Inflammation	Tumor	
Cancer	Neuralgia	Ulcer	
Cramps (except menstrual)	Neuritis	Varicose veins	
Cyst	Pain	Weak(ness)	
	Palsy		

e. What part of the body is affected? _____ *Specify*

Show the following detail:

Head . . . . . skull, scalp, face
Back/spine/vertebrae . . . . . upper, middle, lower
Side . . . . . left or right
Ear . . . . . inner or outer; left, right, or both
Eye . . . . . left, right, or both
Arm . . . . . shoulder, upper, elbow, lower or wrist; left, right, or both
Hand . . . . . entire hand or fingers only; left, right, or both
Leg . . . . . hip, upper, knee, lower, or ankle; left, right, or both
Foot . . . . . entire foot, arch, or toes only; left, right, or both

Except for eyes, ears, or internal organs, ask if there are any of the following entries in 3a–d:

Infection Sore Soreness

f. What part of the (part of body in 3e) is affected by the (infection/sore/soreness) – the skin, muscle, bone, or some other part? *Specify*

Ask if there are any of the following entries in 3a–d:

Tumor Cyst Growth

g. Is this (tumor/cyst/growth) malignant or benign?

1 ☐ Malignant 2 ☐ Benign 9 ☐ DK

A2 Ask remaining questions as appropriate for the condition entered in:
1 ☐ Item 1 3 ☐ Q. 3b 5 ☐ Q. 3d
2 ☐ Q. 3a 4 ☐ Q. 3c 6 ☐ Q. 3e

4. During the past 2 weeks, did his . . . cause him to cut down on the things he usually does? 1 Y 2 N (9)

5. During that period, how many days did he cut down for as much as a day? ____ Days
00 ☐ None (9)

6. During that 2-week period, how many days did his . . . keep him in bed all or most of the day? ____ Days
00 ☐ None

Ask if 17+ years:
7. How many days did his . . . keep him from work during that 2-week period? (For females): not counting work around the house? ____ Days (9)
00 ☐ None (9)

Ask if 6–16 years:
8. How many days did his . . . keep him from school during that 2-week period? ____ Days
00 ☐ None

9. When did -- first notice his . . .?
1 ☐ Last week
2 ☐ Week before
3 ☐ Past 2 weeks–DK which
4 ☐ 2 weeks–3 months
5 ☐ Over 3–12 months
6 ☐ More than 12 months ago
(Was it during the past 12 months or before that time?)
(Was it during the past 3 months or before that time?)
(Was it during the past 2 weeks or before that time?)

A3
1 ☐ Not an eye cond. (*AA*)
2 ☐ First eye cond. (under 6) (*AA*)
3 ☐ First eye cond. (6+ yrs.) (*10*)
4 ☐ Not first eye cond. (*AA*)

10. Can -- see well enough to read ordinary newspaper print WITH GLASSES with his {left / right} eye? . . . 1 Y 2 N

FOOTNOTES

AA

1 ☐ Missing extremity *(A4)*
2 ☐ Condition in C2 does not have a letter as source *(A4)*
3 ☐ Condition in C2 has a letter as source, Doctor seen *(11)*
4 ☐ Condition in C2 has a letter as source, Doctor not seen *(15)*

11a. Does -- NOW take any medicine or treatment for his . . .? 1 Y 2 N *(12)*

b. Was any of this medicine or treatment recommended by a doctor? 1 Y 2 N

12. Has he ever had surgery for this condition? 1 Y 2 N

13. Was he ever hospitalized for this condition? 1 Y 2 N

14. During the past 12 months, about how many times has -- seen or talked to a doctor about his . . .? ____ Times
(Do not count visits while a patient in a hospital.) 000 ☐ None

15a. About how many days during the past 12 months has this condition kept him in bed all or most of the day? ____ Days
000 ☐ None

Ask if 17+ years:

b. About how many days during the past 12 months has this condition kept him from work? ____ Days
For females: **Not counting work around the house?** 000 ☐ None

16a. How often does his . . . bother him -- all of the time, often, once in a while, or never?
1 ☐ All the time 2 ☐ Often 3 ☐ Once in a while
0 ☐ Never *(16c)* 8 ☐ Other – *Specify* _____

b. When it does bother him, is he bothered a great deal, some, or very little?
1 ☐ Great deal 2 ☐ Some 3 ☐ Very little
4 ☐ Other – *Specify* _____

☐ All the time in 16a OR condition list 4 asked *(A4)*

c. Does -- still have this condition?
1 Y *(A4)* N

d. Is this condition completely cured or is it under control?
2 ☐ Cured 3 ☐ Under control *(A4)*
4 ☐ Other – *Specify* _____ *(A4)*

e. About how long did -- have this condition before it was cured?
0 ☐ Less than one month _____ Months _____ Years

A4 ☐ Accident or injury ☐ Other *(NC)*

17a. Did the accident happen during the past 2 years or before that time?
☐ During the past 2 years ☐ Before 2 years *(18a)*

b. When did the accident happen?
☐ Last week ☐ Over 3–12 months
☐ Week before ☐ 1–2 years
☐ 2 weeks–3 months

18a. At the time of the accident what part of the body was hurt?
What kind of injury was it? Anything else?

Part(s) of body	Kind of injury

If accident happened more than 3 months ago, ask:

b. What part of the body is affected now?
How is his -- affected? Is he affected in any other way?

Part(s) of body	Present effects

19. Where did the accident happen?
1 ☐ At home (inside house)
2 ☐ At home (adjacent premises)
3 ☐ Street and highway (includes roadway and public sidewalk)
4 ☐ Farm
5 ☐ Industrial place (includes premises)
6 ☐ School (includes premises)
7 ☐ Place of recreation and sports, except at school
8 ☐ Other – *Specify* _____

20. Was -- at work at his job or business when the accident happened?
1 Y 3 ☐ While in Armed Services
2 N 4 ☐ Under 17 at time of accident

21a. Was a car, truck, bus, or other motor vehicle involved in the accident in any way? 1 Y 2 N *(NC)*

b. Was more than one vehicle involved? Y N

c. Was it (either one) moving at the time? 1 Y 2 N

Appendix V. NHIS 1980
Chronic Condition Checklists

1

32a. **DURING THE PAST 12 MONTHS, did anyone in the family (you, your —, etc.) have –**

If "Yes," ask 32b and c.

b. **Who was this?** Enter name of condition and letter of line where reported in appropriate person's column in item C.

c. **During the past 12 months, did anyone else have . . .?**

Conditions affecting the digestive system,

Make no entry in item C for cold, flu, or grippe even if reported in question 32.

A. Gallstones?		I. Any disease of the *pancreas*?	
B. Any other *gallbladder* trouble?		J. Ulcer?	
C. Cirrhosis of the liver?		K. Hernia or rupture?	
D. Fatty liver?		L. A disease of the *esophagus*?	
E. Hepatitis?		M. Gastritis?	
F. Yellow jaundice?		N. FREQUENT indigestion?	
G. Any other *liver* trouble?		O. Any other *stomach* trouble?	
H. Diabetes?		P. Enteritis?	

2

32a. **Does anyone in the family (you, your —, etc.) NOW have –** If "Yes," ask 32b and c.

b. **Who is this?** Enter name of condition and letter of line where reported in appropriate person's column in item C.

c. **Does anyone else have . . .?**

A. Permanent stiffness or any deformity of the foot, leg, fingers, arm or back? (Permanent stiffness – joints will not move at all)

B. Paralysis of any kind?

32d. **DURING THE PAST 12 MONTHS, did anyone in the family (you, your —, etc.) have –** If "Yes," ask 32e and f.

e. **Who was this?** Enter name of condition and letter of line where reported in appropriate person's column in item C.

f. **During the past 12 months, did anyone else have . . .?**

Conditions C–N and V are conditions affecting the bone and muscle.

C. Arthritis of any kind or Rheumatism?		I. Trick knee?	
D. Gout?		J. A slipped or ruptured disc?	
E. Lumbago?		K. Curvature of the spine?	
F. Osteomyelitis? (os-tee-oh-my-uh-lite-iss)		L. REPEATED trouble with neck, back, or spine?	
G. A bone cyst or bone spur?		M. Bursitis or Synovitis? (sin-uh-vite-iss)	
H. Any other disease of the *bone or cartilage*?		N. Any disease of the *muscles or tendons*?	

3

32a. **DURING THE PAST 12 MONTHS, did anyone in the family (you, your —, etc.) have –**

If "Yes," ask 32b and c.

b. **Who was this?** Enter name of condition and letter of line where reported in appropriate person's column in item C.

c. **During the past 12 months, did anyone else have . . .?**

A. Goiter or other *thyroid* trouble?		Glandular disorders
B. Diabetes?		
C. Cystic fibrosis?		
D. Anemia?		Blood disorder
E. Epilepsy?		Conditions affecting the nervous system
F. Multiple sclerosis?		
G. Migraine?		

1

32a. DURING THE PAST 12 MONTHS, did anyone in the family have –

If "Yes," ask 32b and c.

b. Who was this? Enter in item C.

c. During the past 12 months, did anyone else have . . . ?

Conditions affecting the digestive system.

Make no entry in item C for cold, flu, or grippe even if reported in question 32.

Q. Diverticulitis?

R. Colitis?

S. Spastic colon?

T. FREQUENT constipation?

U. Any other *bowel* trouble?

V. Any other *intestinal* trouble?

W. Cancer of the stomach, colon or rectum?

X. During the past 12 months, did anyone in the family have any other condition of the digestive system? If "Yes," ask: Who was this? – What was the condition? *(Enter in item C)*

2

32d. DURING THE PAST 12 MONTHS, did anyone in the family have –

If "Yes," ask 32e and f.

e. Who was this? Enter in item C.

f. During the past 12 months, did anyone else have . . . ?

Conditions O–U and W–Z are conditions affecting the skin.

O. A tumor, cyst or growth of the skin?

P. Eczema or psoriasis? (so-rye-uh-sis)

Q. TROUBLE with dry or itching skin?

R. TROUBLE with acne?

S. A skin ulcer?

T. Any kind of skin allergy?

U. Dermatitis or any other *skin* trouble?

V. TROUBLE with fallen arches, flatfeet or clubfoot?

W. TROUBLE with ingrown toenails or fingernails?

X. TROUBLE with bunions, corns, or calluses?

Y. A disease of the *hair* or *scalp*?

Z. Any disease of the *lymph* or *sweat glands*?

3

32a. DURING THE PAST 12 MONTHS, did anyone in the family have –

If "Yes," ask 32b and c.

b. Who was this? Enter in item C.

c. During the past 12 months, did anyone else have . . . ?

H. Neuralgia or neuritis?

I. Sciatica?

Conditions affecting the nervous system

J. Nephritis?

K. Kidney stones?

L. Any other *kidney* trouble?

M. *Bladder* trouble?

N. *Prostate* trouble?

O. Disease of the *uterus* or *ovary*?

P. Any other female trouble?

Genito-urinary conditions

4

32a. Does anyone in the family (you, your —, etc.) NOW have –

If "Yes," ask 32b and c.

b. Who is this? – Enter name of condition and letter of line where reported in appropriate person's column in item C.

c. Does anyone else have . . .?

A–L are conditions affecting { hearing / vision / speech }

A. Deafness in one or both ears?		H. A detached retina or any other condition of the *retina*?	
B. Any other trouble hearing with one or both ears?		I. Any other trouble seeing with one or both eyes even when wearing glasses?	
C. Tinnitus or ringing in the ears?		J. A cleft palate or harelip?	
D. Blindness in one or both eyes?		K. Stammering or stuttering?	
E. Cataracts?		L. Any other speech defect?	
F. Glaucoma?		M. A missing finger, hand, or arm, toe, foot, or leg?	
G. Color blindness?		N. A missing (breast), kidney or lung?	

5

32a. Has anyone in the family (you, your —, etc.) EVER had –

If "Yes," ask 32b and c.

b. Who was this? – Enter name of condition and letter of line where reported in appropriate person's column in item C.

c. Has anyone else ever had. . .?

Conditions affecting the heart and circulatory system.

A. Rheumatic fever?		G. Stroke or a cerebrovascular accident?	
B. Rheumatic heart disease?		H. Hemorrhage of the brain?	
C. Hardening of the arteries or arteriosclerosis?		I. Angina pectoris?	
D. Congenital heart disease?		J. Myocardial infarction?	
E. Coronary heart disease?		K. Any other heart attack?	
F. High blood pressure?			

6

32a. DURING THE PAST 12 MONTHS, did anyone in the family (you, your —, etc.) have –

If "Yes," ask 32b and c.

b. Who was this? – Enter name of condition and letter of line where reported in appropriate person's column in item C.

c. During the past 12 months did anyone else have . . .?

Conditions affecting the respiratory system.

A. Bronchitis?		F. Sinus trouble?	
B. Bronchiectasis? (brong ke-ek tah-sis)		G. Deflected or deviated nasal septum?	
C. Asthma?		H. *Tonsillitis or enlargement of the tonsils or adenoids?	
D. Hay fever?			
E. Nasal polyp?		I. *Laryngitis?	

*If reported in question 32 only, ask:

1. How many times did — have . . . in the past 12 months? – If 2+ enter in item C.

If only 1 time, ask:

2. How long did it last? – If 1 month or longer, enter in item C. If less than 1 month, do not record.

If tonsils or adenoids removed during the past 12 months, enter condition causing removal in item C.

Make no entry in item C for cold; flu; red, sore, or strep throat; or "virus" reported in answer to question 32.

4

32a. Does anyone in the family NOW have –

If "Yes," ask 32b and c.

b. Who is this? Enter in item C.

c. Does anyone else have . . . ?

Conditions O–W are impairments.

Conditions Y and Z affect the nervous system.

O. Palsy or cerebral palsy?		**U. PERMANENT stiffness or any deformity of the back, foot, or leg?** (Permanent stiffness – joints will not move at all)	
P. Paralysis of any kind?		**V. PERMANENT stiffness or any deformity of the fingers, hand, or arm?**	
Q. Curvature of the spine?		**W. Mental retardation?**	
R. REPEATED trouble with back or spine?		**X. Any condition caused by an old accident or injury?** If "Yes," ask: **What is the condition?**	
S. Any TROUBLE with fallen arches or flatfeet?		**Y. Epilepsy?**	
T. A clubfoot?		**Z. REPEATED convulsions, seizures, or blackouts?**	

5

32a. DURING THE PAST 12 MONTHS, did anyone in the family (you, your ––, etc.) have –

If "Yes," ask 32b and c.

b. Who was this? Enter in item C.

c. During the past 12 months did anyone else have . . .?

Conditions affecting the heart and circulatory system.

L. Damaged heart valves?		**R. Gangrene?**	
M. Tachycardia or rapid heart?		**S. Varicose veins?**	
N. Heart murmur?		**T. Hemorrhoids or piles?**	
O. Any other *heart* trouble?		**U. Phlebitis or thrombophlebitis?**	
P. Aneurysm?		**V. Any other condition affecting blood circulation?**	
Q. Any blood clots?			

6

32a. DURING THE PAST 12 MONTHS, did anyone in the family have –

If "Yes," ask 32b and c.

b. Who was this? Enter in item C.

c. During the past 12 months, did anyone else have . . . ?

Make no entry in item C for cold; flu; red, sore, or strep throat; or "virus" reported in answer to question 32.

Conditions affecting the respiratory system.

J. Tumor, cyst, or growth of the bronchial tube or lung?		**O. Tumor, cyst, or growth of the throat, larynx, or trachea?**	
K. Emphysema?		**P. Any work-related respiratory condition such as dust on the lungs, silicosis or pneu-mo-co-ni-o-sis?**	
L. Pleurisy?			
M. Tuberculosis?		**Q. During the past 12 months did anyone in the family have any other respiratory, lung, or pulmonary condition?** If "Yes," ask: **Who was this ? – What was the condition?** *(Enter in item C)*	
N. Abscess of the lung?			

Appendix VI. Definitions of Terms Used in This Report

Acute condition—In both NHIS and NMCUES an acute condition was defined as a condition that has lasted less than 3 months and that has involved either medical attention or restricted activity. In NHIS, acute conditions had their onset during the 2 weeks prior to the interview week and involved either medical attention or restricted activity during the 2-week period. In NMCUES, acute conditions had their onset between the reference date (see definition) and the date of interview. However, excluded in both NHIS and NMCUES are some conditions that always are classified as chronic even though the onset occurred within the 3 months prior to the week of the interview. The codes in this report refer to the ninth revision of the International Classification of Diseases, as modified by the NHIS *Medical Coding Manual and Short Index* (NCHS, 1979).

Barriers to care—A section included in the round 5 supplement of NMCUES through which information was collected regarding whether a respondent had a health condition about which the respondent would have liked to see a doctor but did not. Data were also collected about the conditions involved and the reasons why a doctor or other medical person was not seen.

Basic person weight—The weight assigned to a person in NMCUES (and the events and conditions of that person) to adjust for the potential biasing effects of differences in the initial selection probabilities of reporting units and the effects of systematic nonsampling errors related to nonresponse and sampling frame undercoverage. The weight can be viewed as an inflation factor to account for the number of persons or visits in the survey population that the sample unit represents. The basic person weight represents a two-step adjustment: one for the reporting unit (RU), and the other for the person within the RU.

Chronic condition—A condition is considered chronic (1) if the condition is described by the respondent as having been noticed first more than 3 months before the week of the interview (NHIS) or the date of the interview (NMCUES), or (2) if it is one of the conditions always classified as chronic regardless of the onset, listed in Appendix I.

Combining—The procedure for assigning a single diagnostic code when certain of two closely related conditions or more are reported for the same person. Procedures for combining conditions used in NHIS appear in Appendix III of the *NHIS Medical Coding Manual and Short Index*. Procedures for combining conditions in NMCUES can be found in the annotations to the *NHIS Medical Coding Manual and Short Index*.

Condition—A morbidity condition, or simply a condition, is any entry on the questionnaire that describes a departure from a state of physical or mental well-being. It results from a positive response to one of a series of "medical-disability" impact or "illness-recall" questions. It is any illness, injury, complaint, impairment, or problem perceived by the respondent as inhibiting usual activities or requiring medical treatment. Pregnancy, vasectomy, and tubal ligation were not considered to be conditions; however, related medical care was recorded as if they were conditions. Neoplasms were classified without regard to site. Conditions, except impairments, are classified by type according to the *Ninth Revision of the International Classification of Diseases* (World Health Organization, 1977) as modified by the *National Health Interview Survey Medical Coding Manual* (NCHS, 1979); these modifications make the code more suitable for a household interview survey. Impairments are chronic or permanent defects, usually static in nature, that result from disease, injury, or congenital malformation. They represent decrease or loss of ability to perform various functions, particularly those of the musculoskeletal system and the sense organs. Impairments are classified by using a supplementary code specified in the coding manual. In the supplementary code, impairments are grouped according to type of functional impairment and etiology.

Condition code—A diagnostic code assigned to a condition in both NHIS and NMCUES, based upon the ninth revision of the *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death* (1975).

Condition number—A two-digit number associated throughout the data-collection period with a particular condition of a particular person.

Core questionnaire—The basic interview instrument used in NMCUES during each interview to obtain data about health, health care, charges for health care, source of payment, and health insurance coverage.

Disability—The general term used to describe any temporary or long-term reduction of a person's activity (that which restricted the person to bed or from work, or caused the person to cut down on usual activities) as a result of an acute or chronic condition.

Duration of illness—Duration of illness is defined in NHIS as the length of time that a person had the illness prior to the week of interview.

Household—Occupants of a housing unit or group quarters that was included in the NMCUES sample. A household may have been one person, a family of related people, a number of unrelated people, or a combination of related and unrelated people.

Impairment—Impairments are chronic or permanent defects, usually static in nature, that result from disease, injury, or congenital malformation. They are characterized by a decrease in or loss of ability to perform various functions, particularly those of the musculoskeletal system and the sense organs. All impairments are classified by means of a special supplementary code (x). Hence code numbers for impairments in the *Manual of the International Classification of Diseases, Injuries, and Causes of Death* (1975) are not used in NMCUES or NHIS. In the supplementary code, impairments are grouped according to type of functional impairment and etiology.

Incidence of conditions—The estimated number of conditions that have their onset within a specified time period.

Injury—A condition of the type that is classified according to the nature of injury code numbers (800–999) in the *Manual of the International Classification of Diseases, Injury, and Causes of Death* (1975). In addition to fractures, lacerations, contusions, burns, and so forth, which commonly are thought of as injuries, this group of codes includes effects of exposure, such as sunburn; adverse reactions to immunization and other medical procedures; and poisonings. Unless otherwise specified, the term injury is used to cover all of these.

Key person—A key person was (1) an occupant of a national household sample housing unit or group quarters at the time of the first interview; (2) a person related to and living with a State Medicaid household case member at the time of the first interview; (3) an unmarried student 17–22 years of age living away from home and related to a person in one of the first two groups; (4) a related person who had lived with a person in the first two groups between January 1, 1980, and the round 1 interview, but was deceased or had been institutionalized; (5) a baby born to a key person during 1980; or (6) a person who was living outside the United States, was in the Armed Forces, or was in an institution at the time of the round 1 interview but who had joined a related key person.

Limitation—The specific activity and extent to which the person partially can perform an activity, or can do it fully only part of the time, or cannot do it all. Specific questions about limitations are included in the respective questionnaires of both NHIS and NMCUES.

Merging—The consolidating of data about medical care, date of onset, disability, and so forth, when conditions or sites that can be combined are described in separate condition sections. Thus, combining results in the consolidation of conditions, and merging can result in the consolidation of events, such as medical visits, associated with each of those conditions. Merging rules were employed in NHIS but not in NMCUES.

NHIS—The National Health Interview Survey is a national, cross-sectional survey conducted annually since 1957. Through NHIS information on illness, disability, and the use of medical care services is obtained.

NMCUES—The National Medical Care Utilization and Expenditure Survey was a national longitudinal survey conducted in 1980–81 whose primary purpose was the collection of information on health, access to and use of medical services, associated charges and sources of payment, and health insurance coverage. The survey was conducted in five rounds at approximately 3-month intervals.

Onset of condition—A condition is considered to have had its onset when it was first noticed. Onset could be the time the person first felt sick or became injured, or it could be the time when the person or family first was told by a physician that the person had a condition of which he or she had been previously unaware.

Prevalence of conditions—Prevalence refers to all known cases of a disease in a specified population during a particular period of time regardless of their date of onset.

Recall or reference period—The period of time for which respondents were asked to report medical events. In NMCUES, the recall period was calculated as the date between the last interview and the present interview, except for round 1, where the beginning of the recall period was January 1, 1980, and round 5, where the recall period ended December 31, 1980. In NMCUES, the recall period averaged 90 days. In NHIS the recall period is defined as the 2-week period prior to the week of the interview.

Reference date—In NMCUES, the reference date is the date of the previous interview except (1) for the first interview, it was January 1, 1980, and (2) for a new person, it was the date the person joined the reporting unit. In NHIS, the reference period was the 2-week period ending prior to the Sunday of the week of the interview.

Round—The administrative term used to designate all interviews that occurred within a given period and that used the same instruments and procedures.

Three-month convention—During the data collection period, a condition was considered to be chronic if it was first noticed 3 months or more prior to the week of the interview (NHIS) or the date of the interview (NMCUES.)

Appendix VII. Method for Estimating Standard Errors

Estimating NMCUES Standard Errors

This appendix describes the procedures used to estimate standard errors of statistics derived from the NMCUES (Tourangeau and Rasinski, 1987) and the NHIS (Jack, 1981; Kovar and Poe, 1985) data. Most researchers are familiar with the use of standard errors to assess the variability of estimates based on simple random samples. Complex sample designs, such as the NMCUES household sample design, require computational procedures different from those used for a simple random sample to estimate variances. The NMCUES household sample departs from simple random sampling in three respects. First, the NMCUES data are clustered by geographical units. Because the sample was selected in stages (for example, the selection of counties and SMSA's constituted the initial step in sample selection), the respondents are not completely independent. By contrast, in a simple random sample, the selection probabilities for each unit are independent. Second, the NMCUES sample is stratified. For example, the selection of primary sampling units (PSU's) assured proportionate representation of each region in the country. Simple random samples lack such controls. Finally, the NMCUES data are weighted; these weights compensate for differences in the selection probabilities of individual responses. With a simple random sample, no weights are needed because all respondents have an equal chance of selection.

Each of these departures from the assumptions of simple random sampling affects the variability of sample estimates. The net impact of these departures is that, in general, estimates derived from complex samples vary considerably more than do similar estimates derived from a simple random sample with the same number of cases. Standard statistical packages—such as SPSS and SAS—assume simple random sampling or closely related designs; consequently, the results from such packages can be seriously misleading when estimates are derived from complex sample designs.

The standard errors of the NMCUES data presented in this report were estimated by the method of balanced repeated replication (BRR).

Balanced Repeated Replication

The replication approach originally was developed by Deming (1956). The principle underlying replicated sampling is quite simple. If a sample of size n is desired, g independent replicate samples are selected, each of size n/g. The variation among estimates from each replicate can be used to estimate the variance of estimates based on the entire sample. In fact, the NMCUES household sample is such a replicated sample, consisting of two independent national samples, each selected from the general-purpose national samples of the two contractors, NORC and Research Triangle Institute. The replication approach is limited, however, in that the precision of the standard error estimates depends on the number of replicates in the design; with fewer replicates, the standard errors are less accurate.

Balanced repeated replication (BRR), which extends the principle of replication, usually is applied to stratified designs with two primary selections per stratum. When one primary selection is chosen from each stratum, a half sample is created; the unselected primary units form another half sample. In a design with h strata, a total of $2^{(h-1)}$ different pairs of half samples can be formed in this fashion. Each pair is referred to as a replicate. It is customary to form only a portion of the possible replicates using an orthogonal balanced design.

For any given replicate, estimates such as the ratio means, r_1 and r_2, can be computed from each half sample. The sampling variance for the overall statistic (r) then can be estimated in any of several ways (Frankel, 1971). One method compares the estimate from one half sample with the overall estimate:

$$\mathrm{Var}_k\,(r) = (r_{1k} - r)^2$$

where

$\mathrm{Var}_k\,(r)$ = the variance estimate based on replicate k,
r = an estimate based on the entire sample, and
r_{1k} = an estimate based on one of the half samples from replicate k.

The final estimate for the variance of r is the average of Var_k across all the replicates. The estimate r need not be a ratio mean; the logic of BRR applies to any type of estimate, giving the method its broad generality.

Through BRR a number of practical advantages are offered. First, the program had been used to analyze NMCUES data in a previous NMCUES evaluation report (Tourangeau and Rasinski, 1987); thus, it was familiar and would minimize startup costs. Second, the program was known to have given accurate results (Tourangeau et al., 1983). Finally, the program was designed to be embedded in SAS, allowing analysts to use results from other SAS procedures while simultaneously calculating accurate standard error estimates.

The program—balanced repeated replication variance (BRRVAR)—was developed originally under contract to the National Center for Education Statistics (NCES). It has been used by NORC and by other researchers to analyze data from the High School and Beyond Survey. NORC's experience with the program indicates that it produces accurate results at reasonable cost. Detailed documentation of the program is available through NCES.

Estimating NHIS Standard Errors

The relative standard error of an estimate is obtained by dividing the standard error of the estimate by the estimate itself and is expressed as a percentage of the estimate. A set of relative standard error curves was drawn using the empirical relationship between the size of the estimate and the relative standard error of the estimate. Included in this appendix are charts from which the relative standard errors can be determined for estimates shown in the report. The charts provide an estimate of the approximate relative standard error rather than the precise error for any specific aggregate or percentage. Although they take the size of the denominator into account, they do not take the differential effects of the clustered design into account; that is, two population groups of the same size may have different design effects.

Three classes of statistics for the health survey are identified for purposes of estimating variances.

1. *Narrow range*—This class consists of (1) statistics which estimate a population attribute; that is, the number of persons in a particular condition group, and (2) statistics for which the measure for a single individual during the reference period used in data collection is usually either 0 or 1 and, on occasion, may take on the value 2 or very rarely 3.
2. *Medium range*—This class consists of other statistics for which the measure for a single individual during the reference period used in data collection will rarely lie outside the range 0 to 5.
3. *Wide range*—This class consists of statistics for which the measure for a single individual during the reference period used in data collection can range from 0 to a number in excess of 5; for example, the number of conditions.

In addition to classifying variables according to whether they are narrow-, medium-, or wide-range, statistics in the survey are further classified as to whether they are based on a reference period of 2 weeks, 6 months, or 12 months.

General Rules for Determining Relative Standard Errors

The following rules will enable the reader to determine approximate relative standard errors from the charts for estimates presented in this report. These charts represent standard errors of NHIS data.

Rule 1. Estimates of aggregates—Approximate relative standard errors for estimates of aggregates such as the number of persons with a given characteristic are obtained from appropriate curves, figures I–V. The number of persons in the total U.S. population or in an age-sex-color class of the total population is adjusted to official Bureau of the Census figures and is not subject to sampling error.

Rule 2. Estimates of percentages in a percent distribution—Relative standard errors for percentages in a percent distribution of a total are obtained from appropriate curves, figures VI–VII. For values which do not fall on one of the curves presented in the chart, visual interpolation will provide a satisfactory approximation.

Rule 3. Estimates of rates where the numerator is a subclass of the denominator—This rule applies for prevalence rates or where a unit of the numerator occurs, with few exceptions, only once in the year for any one unit in the denominator. For example, in computing the rate of visual impairments per 1,000 population, the numerator consisting of persons with the impairment is a subclass of the denominator, which includes all persons in the population. If converted to rates per 100, such rates may be treated as though they were percentages and the relative standard errors obtained from the percentage charts for population estimates. Rates per 1,000, or on any other base, first must be converted to rates per 100; then the percentage chart will provide the relative standard error per 100.

Rule 4. Estimates of rates where the numerator is not a subclass of the denominator—This rule applies where a unit of the numerator often occurs more than once for any one unit in the denominator. For example, in the computation of the number of persons injured per 100 currently employed persons per year, it is possible that a person in the denominator could have sustained more than one of the injuries included in the numerator. Approximate relative standard errors for rates of this kind may be computed as follows.

(a) Where the denominator is the total U.S. population or includes all persons in one or more of the age-sex-color groups of the total population, the relative error of the rate is equivalent to the relative error of the numerator, which can be obtained directly from the appropriate chart.

(b) In other cases the relative standard error of the numerator and of the denominator can be obtained from the appropriate curve. By squaring each of these relative errors, adding the resulting values, and extracting the square root of the sum, an upper bound on the standard error can be attained which will overstate the error to the extent that the correlation between numerator and denominator is greater than zero.

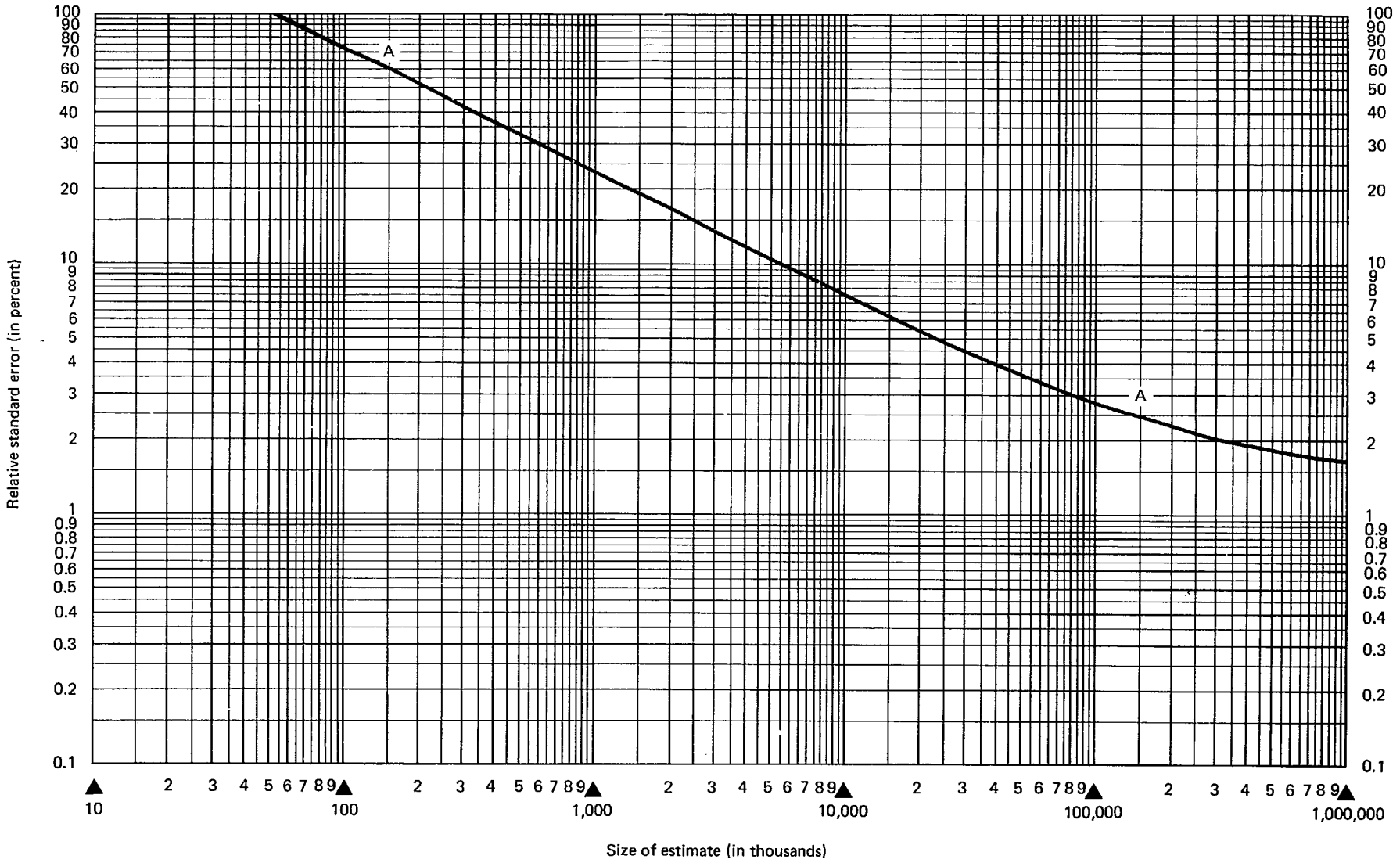
Rule 5. Estimates of the difference between two statistics (for example, mean, rate, total)—The standard error of a difference is approximately the square root of the sum of the squares of each standard error considered separately. A formula for the standard error of the difference

$$d = X_1 - X_2$$

is

$$\sigma_d = \sqrt{(X_1 V_{X_1})^2 + (X_2 V_{X_2})^2}$$

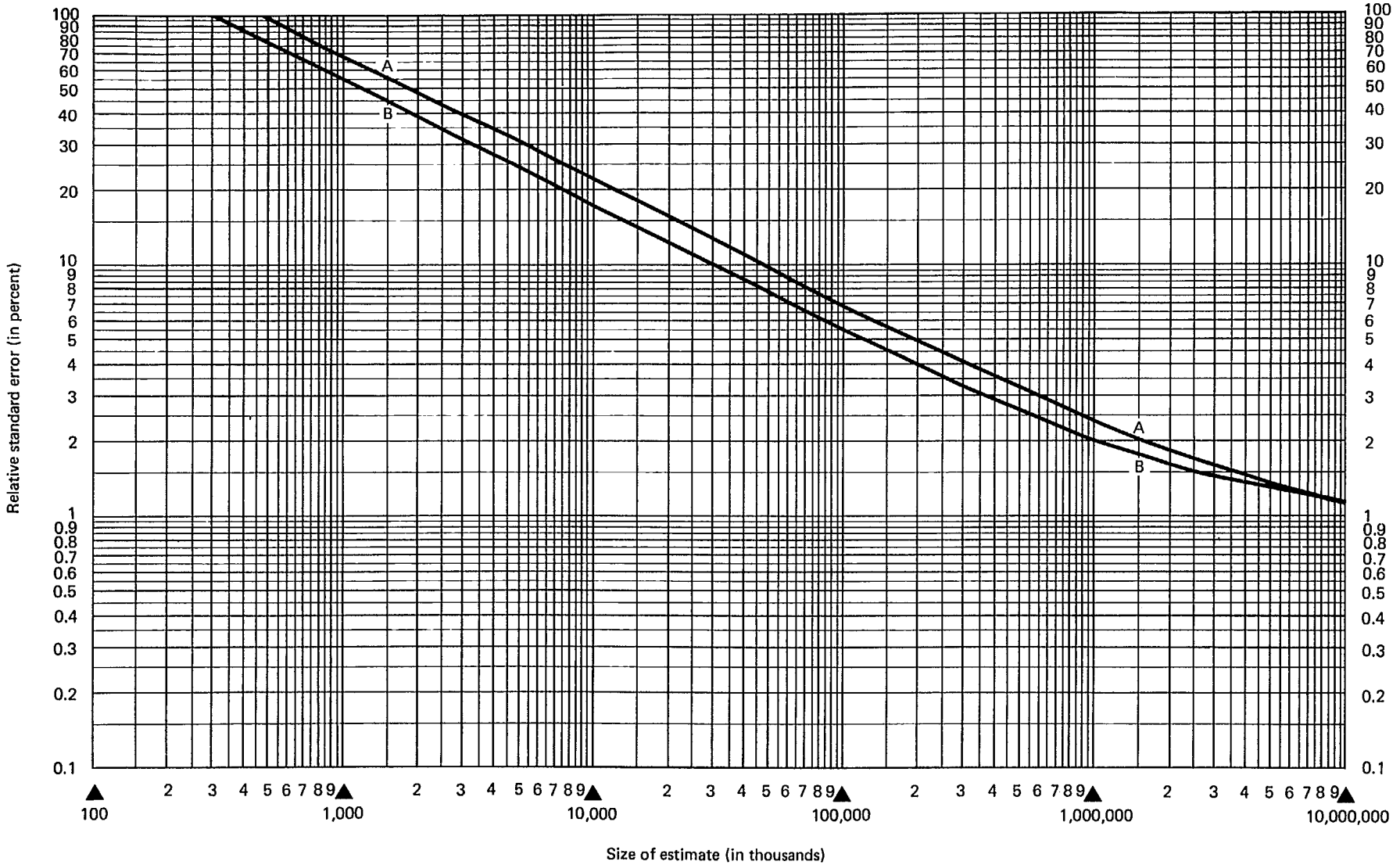
where X_1 is the estimate for class 1, X_2 is the estimate for class 2, and V_{X_1} and V_{X_2} are the relative standard errors of X_1 and X_2, respectively. This formula will represent the actual standard error quite accurately for the difference between separate and uncorrelated characteristics, although it is only a rough approximation in most other cases. The relative standard error of each estimate involved in such a difference can be determined by one of the four previous rules, whichever is appropriate.



NOTE: This curve represents estimates of relative standard errors based on 1 to 4 quarters of data collection for narrow range estimates of aggregates using a 2-week reference period.

Example of use of chart: An estimate of 1,000,000 acute respiratory conditions (on scale at bottom of chart) has a relative standard error of 23 percent (read from scale at left side of chart), or a standard error of 230,000 (23 percent of 1,000,000).

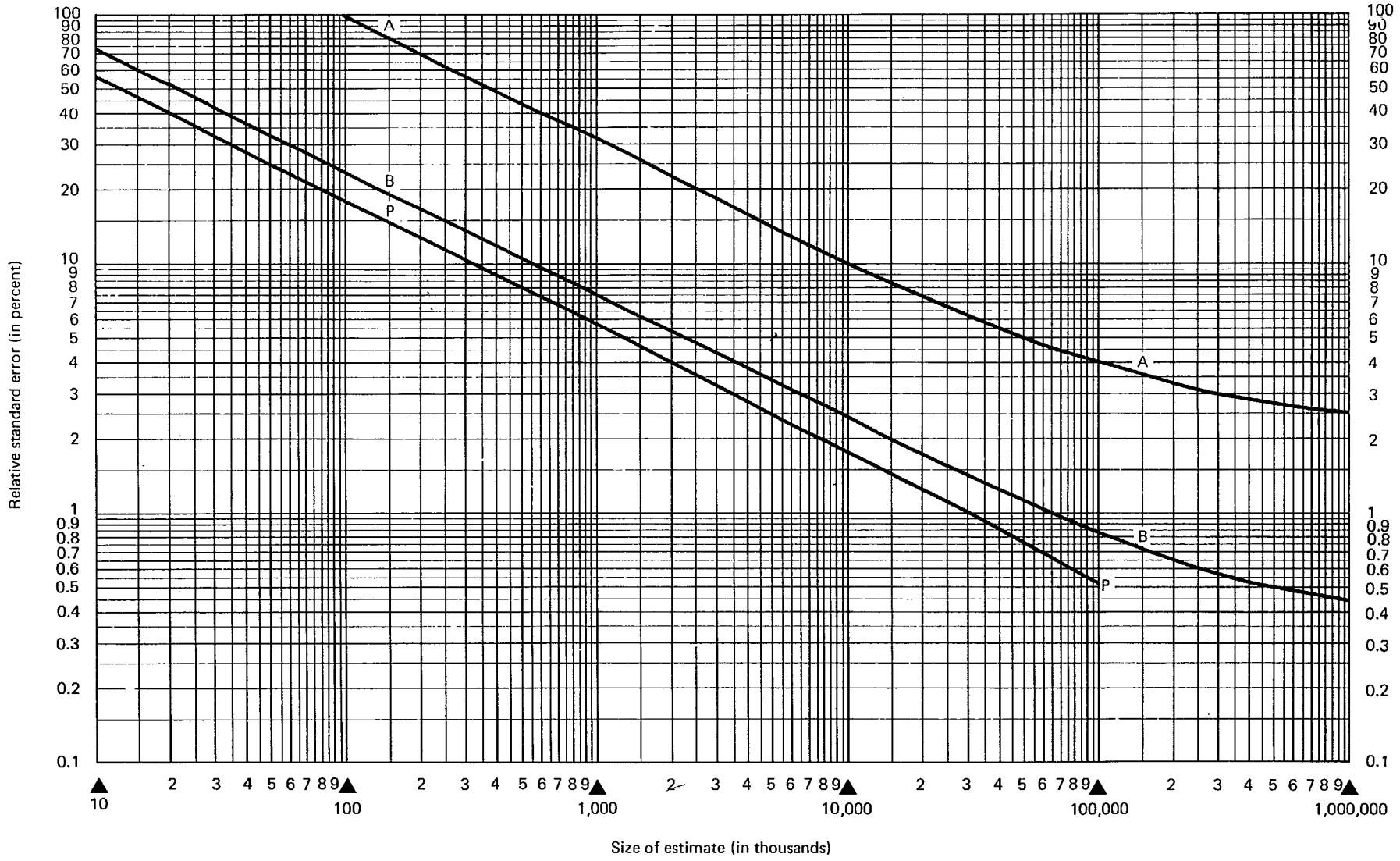
Figure I. Relative standard errors for number of acute conditions or persons injured



NOTE: These curves represent estimates of relative standard errors based on 1 to 4 quarters of data collection for wide range estimates of aggregates using a 2-week reference period.

Example of use of chart: An estimate of 10,000,000 days of restricted activity (on scale at bottom of chart) has a relative standard error of 22 percent (read from curve A on scale at left side of chart), or a standard error of 2,200,000 (22 percent of 10,000,000).

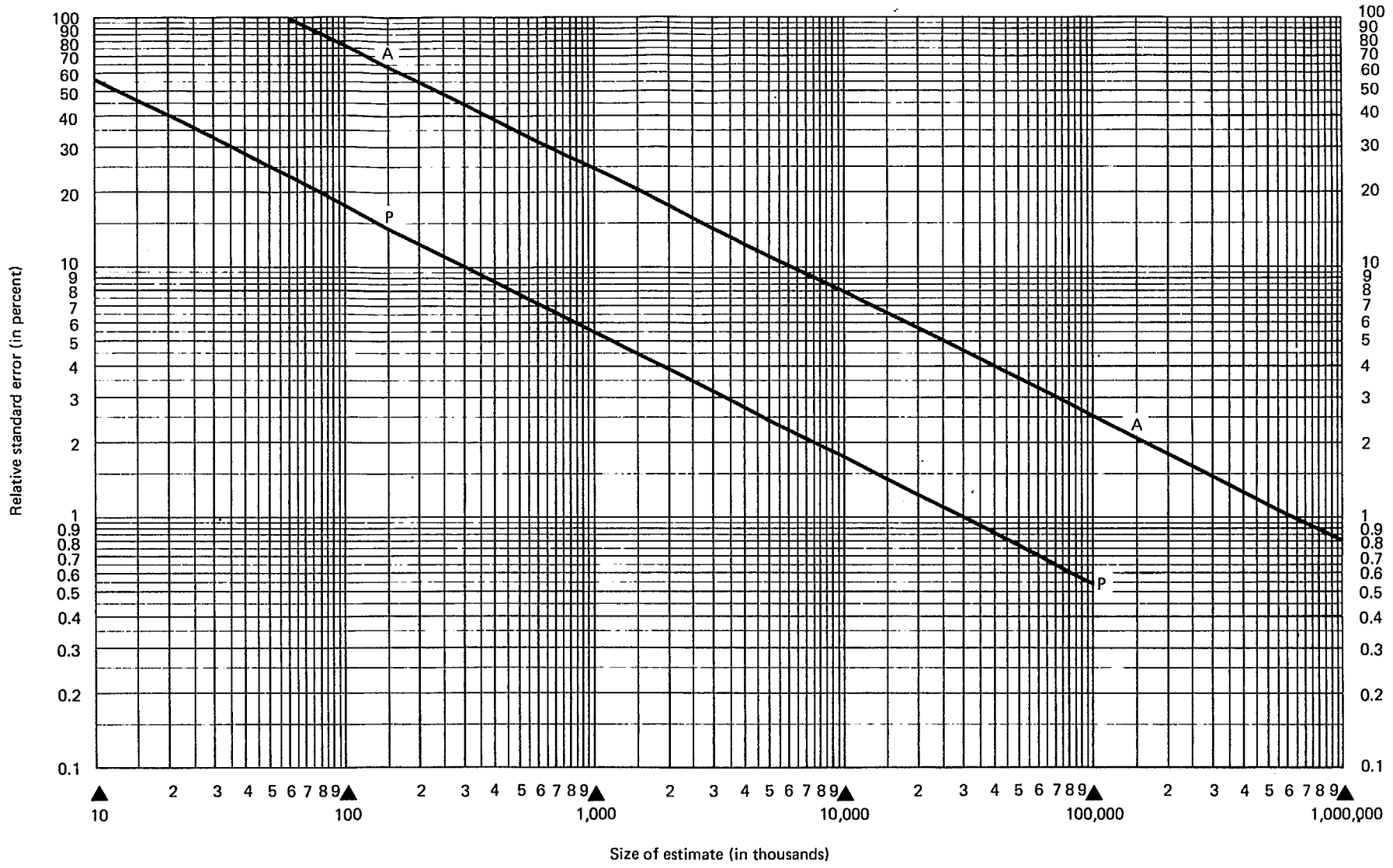
Figure II. Relative standard errors for days of restricted activity or bed disability (A) and for days lost from work or school (B)



NOTE: The curves related to short-stay hospital days and discharges are based on 4 quarters of data collection for wide and narrow range estimates of aggregates using a 6-month reference period; the curve for population characteristics is based on 4 quarters of data collection for narrow range estimates of aggregates.

Example of use of chart: An estimate of 10,000,000 hospital days (on scale at bottom of chart) has a relative standard error of 10.2 percent (read from curve A on scale at left side of chart), or a standard error of 1,020,000 (10.2 percent of 10,000,000). An estimate of 1,000,000 discharges from short-stay hospitals (curve B) has a relative standard error of 7.4 percent. An estimate of 1,000,000 persons in the Northeast Region (curve P) has a relative standard error of 5.7 percent.

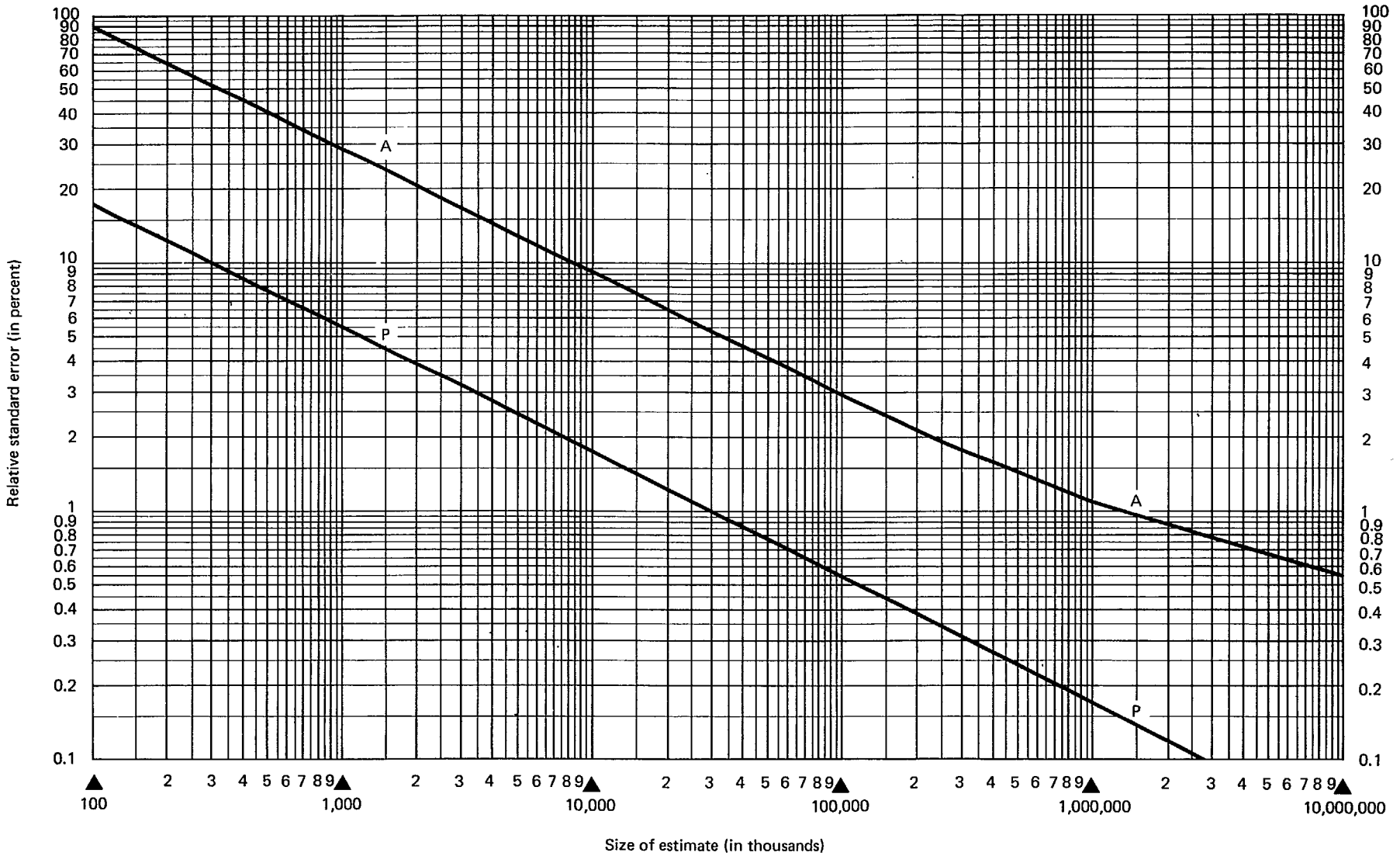
Figure III. Relative standard errors for number of short-stay hospital days (A), short-stay hospital discharges (B), and population characteristics (P)



NOTE: The curve related to hospital days is based on 4 quarters of data collection for wide range estimates of aggregates using a 12-month reference period; the curve for population characteristics is based on 4 quarters of data collection for narrow range estimates of aggregates.

Example of use of chart: An estimate of 10,000,000 days of hospitalization in the past year (on scale at bottom of chart) has a relative standard error of 7.8 percent (read from curve A on scale at left side of chart), or a standard error of 780,000 (7.8 percent of 10,000,000). An estimate of 1,000,000 persons with 1 hospital episode or more (curve P) has a relative standard error of 5.7 percent.

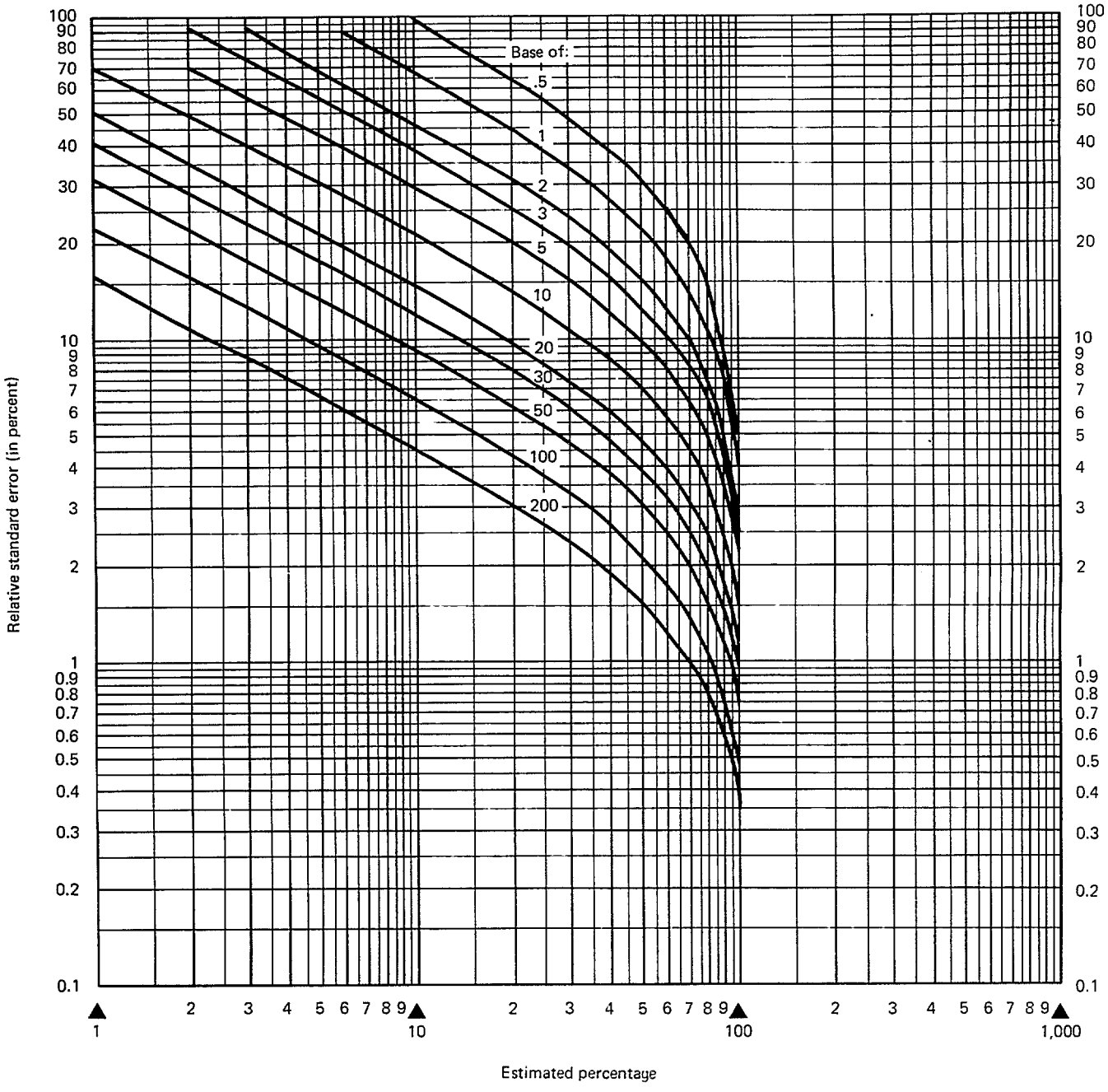
Figure IV. Relative standard errors for short-stay hospital days based on a 12-month reference period (A) and population characteristics (P)



NOTE: The curve related to physician or dental visits is based on 1 to 4 quarters of data collection for medium range estimates of aggregates using a 2-week reference period; the curve for population characteristics is based on 4 quarters of data collection for narrow range estimates of aggregates.

Example of use of chart: An estimate of 10,000,000 dental visits (on scale at bottom of chart) has a relative standard error of 9.2 percent (read from curve A on scale at left side of chart), or a standard error of 920,000 (9.2 percent of 10,000,000). An estimate of 1,000,000 persons in the Northeast Region (curve P) has a relative standard error of 5.7 percent.

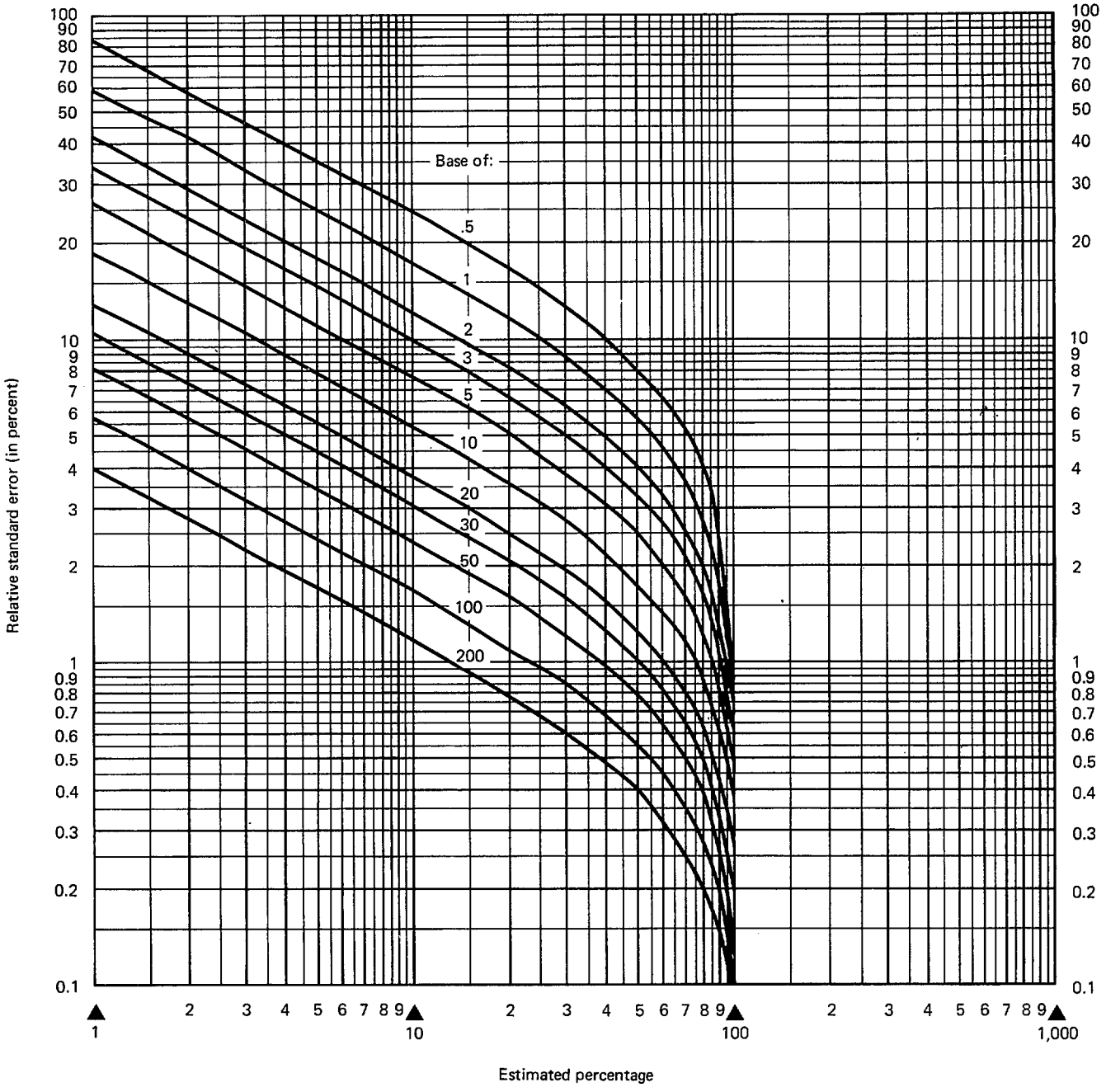
Figure V. Relative standard errors for number of physician or dental visits based on a 2-week reference period (A) and population characteristics (P)



NOTE: These curves represent estimates of relative standard errors of percentages of acute conditions or persons injured based on 1 to 4 quarters of data collection for narrow range data using a 2-week reference period.

Example of use of chart: An estimate of 20 percent (on scale at bottom of chart) based on an estimate of 10,000,000 has a relative standard error of 14.5 percent (read from the scale at the left side of chart), the point at which the curve for a base of 10,000,000 intersects the vertical line for 20 percent. The standard error in percentage points is equal to 20 percent ×14.5 percent, or 2.9 percentage points.

Figure VI. Relative standard errors of percentages of acute conditions or persons injured (Base of percentage shown on curves in millions)



NOTE: These curves represent estimates of relative standard errors of percentages of population characteristics based on 4 quarters of data collection for narrow range estimates.

Example of use of chart: An estimate of 20 percent (on scale at bottom of chart) based on an estimate of 10,000,000 has a relative standard error of 3.6 percent (read from the scale at the left side of chart), the point at which the curve for a base of 10,000,000 intersects the vertical line for 20 percent. The standard error in percentage points is equal to 20 percent ×3.6 percent, or 0.72 percentage points.

Figure VII. Relative standard errors of percentages of population characteristics (Base of percentage shown on curves in millions)

Department of Health and Human Services
Otis R. Bowen, M.D., Secretary

Health Care Financing Administration
William L. Roper, M.D., Administrator

Office of Research and Demonstrations
Judith B. Willis, Director

Office of Research
J. Michael Fitzmaurice, Ph.D., Director

Division of Program Studies
Carl Josephson, Director

Survey Studies Branch
Herbert A. Silverman, Ph.D., Chief

Public Health Service
Robert E. Windom, M.D.,
Assistant Secretary for Health

National Center for Health Statistics
Manning Feinleib, M.D., Dr.P.H., Director

Office of Vital and Health Statistics Systems
Peter L. Hurley, Associate Director

Division of Health Interview Statistics
Owen T. Thornberry, Jr., Ph.D., Director

Utilization and Expenditure Statistics Branch
Robert A. Wright, Chief