NATIONAL CENTER For HEALTH STATISTICS | Series 4 Number 11

VITAL and HEALTH STATISTICS
DOCUMENTS AND COMMITTEE REPORTS

## use of hospital data for

# Epidemiologic and Medical-Care Research

## A Report of the United States National Committee on Vital and Health Statistics

How hospital data may be used for epidemiologic and medical-care research is discussed with examples of past applications.

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service
Health Services and Mental Health Administration

Washington, D.C. June 1969


Public Health Service Publication No. 1000-Series 4-No. 11

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C., 20402 - Price 25 cents

# NATIONAL CENTER FOR HEALTH STATISTICS



## OFFICE OF HEALTH STATISTICS ANALYSIS



Public Health Service Publication No. 1000-Series 4-No. 11

*Library of Congress Catalog Card Number 70-600272*

# FOREWORD

This report, prepared by a Subcommittee of the U.S. National Committee on Vital and Health Statistics, views the development of hospital data in the United States and discusses how they may be used for epidemiologic studies of chronic disease, for disease surveillance purposes, and for medical-care research. The full potential of morbidity and other data in hospital records needs to be exploited for epidemiologic and medical-care research.

Robert L. Berg, M.D., Chairman
U.S. National Committee on
Vital and Health Statistics

## U.S. NATIONAL COMMITTEE ON VITAL AND HEALTH STATISTICS

Robert L. Berg, M.D., Chairman
Professor and Chairman, Department of Preventive Medicine and Community Health
The Medical Center
The University of Rochester
Rochester, New York

Iwao M. Moriyama, Ph.D., Executive Secretary
Director, Office of Health Statistics Analysis
National Center for Health Statistics
Public Health Service*
Washington, D.C.

Donald J. Davids
Chief, Records and Statistics Section
Colorado State Department of Public Health
Denver, Colorado

Clyde V. Kiser, Ph.D.
Senior Member, Technical Staff
Milbank Memorial Fund
New York, New York

Herbert E. Klarman, Ph.D.
Department of Public Health Administration
School of Hygiene and Public Health
The Johns Hopkins University
Baltimore, Maryland

Charles E. Lewis, M.D.
Professor and Chairman, Department of Preventive Medicine and Community Health
University of Kansas Medical Center
Kansas City, Kansas

Forrest E. Linder, Ph.D.
Department of Biostatistics
School of Public Health
University of North Carolina
Chapel Hill, North Carolina

Robert Parke, Jr.
Program Planning Officer
Bureau of the Census
U.S. Department of Commerce
Washington, D.C.

John R. Philp, M.D.
Health Officer
County of Orange Health Department
Santa Ana, California

Donovan J. Thompson, Ph.D.
Department of Preventive Medicine
School of Medicine
University of Washington
Seattle, Washington

Theodore D. Woolsey, ex officio
Director, National Center for Health Statistics
Public Health Service*
Washington, D.C.

## SUBCOMMITTEE ON EPIDEMIOLOGIC USE OF HOSPITAL DATA

Paul M. Densen, Sc.D., Chairman
Director, Center for Community Health and Medical Care
School of Public Health and Medical School
Harvard University
Boston, Massachusetts

Warren Winkelstein, Jr., M.D., Secretary
Professor, Department of Epidemiology
School of Public Health
University of California
Berkeley, California

Jacob E. Bearman, Ph.D.
Professor, Biometry Division
School of Public Health
College of Medical Sciences
University of Minnesota
Minneapolis, Minnesota

Alexander D. Langmuir, M.D,
Chief, Epidemiology Program
National Communicable Disease Center
Health Services and Mental Health Administration, Public Health Service*
Atlanta, Georgia

Alfonse T. Masi, M.D., Dr. P.H.
Department of Medicine
The University of Tennessee
Memphis, Tennessee

Robert W. Miller, M.D.
Chief, Epidemiology Branch
National Cancer Institute
National Institutes of Health
Public Health Service*
Bethesda, Maryland

Robert M. Sigmond
Executive Director
Hospital Planning Association of Allegheny County
Pittsburgh, Pennsylvania

Vergil N. Slee, M.D.
Director, Commission on Professional and Hospital Activities
Ann Arbor, Michigan

Paul F. Wehrle, M.D.
Chief Physician
Pediatrics and Communicable Disease Services
Los Angeles County General Hospital
Los Angeles, California

*Department of Health, Education, and Welfare.

# CONTENTS

*IN THIS REPORT a Subcommittee of the U.S. National Committee on Vital and Health Statistics considers the potential value, the available methods, and the problems of using hospital data for epidemiologic studies.*

*Increasing standardization of diagnostic procedures improves the reliability of hospital diagnoses. Increasing utilization of hospitals reduces the difference between the medical experience of the hospitalized and the general populations. Both trends enhance the potential of hospital data for epidemiologic study. Hospital records provide access to large numbers of cases of specific diseases which would be difficult to identify in general population surveys.*

*Techniques are described, problems are identified, and examples are given for several types of hospital studies: retrospective (usually case-control) studies of specific diseases or conditions; prospective (cohort) studies with either analytical or clinical-trial objectives; disease surveillance intended to identify changes in levels of incidence; and medical-care research which is concerned with relationships between utilization and the characteristics of physicians, hospitals, and case management as well as the specific diagnosis.*

*Opportunities for hypothesis testing are pointed out, along with requirements, in terms of procedures for recording data and the selection of control groups.*

# USE OF HOSPITAL DATA FOR EPIDEMIOLOGIC AND MEDICAL-CARE RESEARCH

## INTRODUCTION

At the suggestion of the Epidemiology Section of the American Public Health Association made during its 1964 meeting in New York City, the Subcommittee on Epidemiologic Use of Hospital Data was constituted by the U.S. National Committee on Vital and Health Statistics. The task of this Subcommittee was to make a study of the uses of diagnostic and other data on hospital patients, including statistics needed for epidemiologic and medical-care research, studies of current therapeutic practices, and for health surveillance.

With the increased standardization of diagnostic tests and criteria, hospital data have become more reliable in recent years. The fact that the general hospital is being regarded more and more as the hub of a complex of facilities providing medical care—inpatient and outpatient, preventive and curative—increases the possibility that hospital data will provide a useful indicator of the medical experience of the general population. This opens up the possibility of utilizing diagnostic information in hospital records for purposes other than the treatment of patients. Indeed the subcommittee believes that the full potential of information in hospital records has yet to be exploited.

The major advantage presented by such data lies in the ready accessibility of large numbers of cases of specific illnesses. To exploit this advantage to the full, it is essential that rigorous scientific procedures be followed.

In this report hospital refers to the short-term general hospital. It is recognized that additional studies of data from chronic disease and psychiatric hospitals and from nursing homes will sometimes be necessary because short-term general hospitals are utilized primarily by patients with acute illnesses and with surgical problems.

The three areas in which the information in hospital records can contribute are epidemiologic research, disease surveillance, and medical-care research. Each of these areas will be briefly reviewed with respect to the role of the hospital record.

## EPIDEMIOLOGIC RESEARCH

While data on hospital patients have, in the past, provided a resource for testing hypotheses on disease causation, it is anticipated that this resource will be increasingly useful in the future. This is due to the increasing availability of hospital care through the extension of private and governmental support as well as to the increased recognition of the hospital as a technological center. Masi[1] has pointed out the potential uses of hospital data for epidemiologic research as well as their limitations. While some problems attend the use of hospital data for epidemiologic purposes, the availability of these data as well as their comprehensive nature suggest that their use be fully exploited.

The two approaches to the epidemiologic testing of hypotheses are (1) the retrospective or case-control method and (2) the prospective or cohort method. In general, findings of retrospective studies will need to be extended by the appli-

cation of more rigorous designs based on either community population samples or on prospective studies of carefully defined hospital patients and controls. Prospective studies may be regarded as more definitive testing of hypotheses.

### Retrospective (case-control) Studies

Retrospective (case-control) studies can proceed along either of two lines—namely, one in which cases and controls are both drawn from the hospital population or one in which cases are drawn from a hospital population and controls from nonhospital populations or the community at large.

*Cases and controls from hospital populations.*—Several studies in which both cases and controls were derived from hospital populations are available for comment. These studies usually involve diseases which cause the majority of patients with the diseases to be hospitalized and which can be diagnosed using fairly standard and reproducible criteria, e.g., lung cancer,[2] sarcoidosis,[3] and Hashimoto's disease.[4]

The usual method of conducting a retrospective study involves identifying a sample of patients with the diagnosis and subsequently obtaining relevant information regarding the characteristics of these patients. A control sample is obtained from patients with other diagnoses, and the frequency of the characteristics under study is compared in the two groups. A classic example of such a study was reported by Levin and others[2] and was concerned with the relationship between cancer of certain sites and tobacco smoking. The investigators routinely obtained a history of tobacco usage from all patients admitted to the Roswell Park Memorial Institute, a cancer research hospital, beginning in 1938 and showed an association between cigarette smoking and cancer of the lung and between pipe smoking and cancer of the lip. This study clearly demonstrated the following advantages of using data on hospitalized patients for retrospective epidemiologic study:

a. Availability of a large number of cases of a particular disease and of suitable controls for comparison analysis.

b. Possibility of collecting information on independent variables of interest such as smoking from cases and controls.

c. High degree of diagnostic accuracy in defining case and control populations.

On the other hand, this study also illustrated some of the problems which have been cited in connection with the use of hospital populations.

a. Hospital detected cases may not be inclusive and are subject to selection according to factors such as socioeconomic status, severity of disease, and psychological and culturally determined attitudes of the hospital patient.

b. Because of the selection referred to, the population base furnishing the cases cannot be precisely defined.

c. Most hospital records are not designed with research uses in mind. The records are frequently incomplete and the approaches to data collection are not standardized. Further, there is considerable variability between hospitals in diagnostic accuracy and reproducibility. However, such difficulties may be minimized in a research institution like the Roswell Park Memorial Institute, where the record system and the clinical activities may be subjected to the control of the research staff. In other situations, research uses may be expected to promote standardization and improvement in quality.

The most serious criticism of the type of study outlined above is that the hospital admission rates for each of the designated disease groups (cases and controls) may be different and may be related to the independent variable under study. Berkson[5] has discussed this problem and has delineated the conditions under which valid comparisons may be made. If unequal admission rates are unrelated to the independent variable under study, then comparisons of the diagnostic groups can be made without bias. The problem is to determine whether or not the independent variable affects the admission rate. This determination is not always possible, but it must always be attempted.

*Cases from hospital populations and controls from nonhospitalized populations.*—There are several ways of guarding against the possibility of error from unrepresentativeness of controls.

Among them is the selection of hospital case controls from a wide variety of diagnoses so that possible hidden factors of bias related to certain diagnostic categories will be diluted. Another approach is to draw the control sample from the general population usually by one of two methods. The first is to obtain a matched sample by controlling on a number of standard variables such as age, race, and economic status. The second, and preferable approach, is to obtain a random sample of the population and then to match it to the sample. The advantage of the latter is that one can get frequency estimates of the independent variables of interest in the population at large and compare them with both the matched control sample and with the case group. Both of these methods were used in a study of pregnancy wastage and coronary artery disease in females by Winkelstein and others.[6] These investigators obtained a case population of women who had survived a myocardial infarction and compared their pregnancy patterns with a matched sample of women drawn by canvassing houses adjacent to the residences of the cases and matching on age, race, and marital status. A second control was obtained from a random sample of the population surveyed for other purposes, which also provided complete information on pregnancy experience. These findings indicated that there was about twice as much pregnancy wastage among the cases of myocardial infarction as among either matched controls or controls drawn from the population sample. A difficult problem to overcome in this type of study is the difference in recall for the characteristic under consideration between hospitalized patients and population controls. Furthermore, it is difficult to hide from the interviewer the identity of the cases as contrasted with the controls.

The advantage of this type of retrospective design is apparent when dealing with diseases of low prevalence in the population. Under such circumstances it is usually difficult, if not impossible, to identify cases from a population sample; therefore, reliance must be on hospital data for the initial recruitment of the case population.

## Prospective (cohort) Studies

*Analytic studies.*—A prospective hospital-based analytic study differs from the retrospective study in that data are collected according to a defined protocol on a defined cohort. In the prospective study the investigator begins with a hypothesis and designs a protocol to collect systematically the history, physical examination data, and laboratory data necessary to test the hypothesis. Thus, the fullest control can be exercised in collecting the required information. However, there may be selection in the hospitalized segment of the total cases in the community depending upon the hospital utilization pattern for any disease. Examples of some prospective controlled epidemiologic studies conducted in hospitals will illustrate some of these points.

A study of viral hepatitis in a group of Boston hospitals highlights the transition from a retrospective to a prospective analysis of hospital data as well as the increased power of the prospective study design. After a retrospective review of hospitalized hepatitis patients, Grady and Chalmers[7,8] designed a prospective study in 1963, involving 10 Boston hospitals, to determine whether hepatitis patients had different types of exposures than matched controls. Biweekly visits were made to each of the participating hospitals. Included in the study were patients over 15 years of age with a working diagnosis of hepatitis. These patients were given a standard questionnaire which included questions about personal contact with other hepatitis patients, possible anicteric cases, travel, eating habits and places, parenteral exposures, and transfusions. The same questionnaire was given to paired controls. The controls were matched by age, race, sex, marital status, hospital pay status, and date of admission. They were previously healthy persons admitted for incidental surgery, e.g., fractures and appendicitis.

The findings from the study showed that hepatitis patients who had not recently been transfused had a significantly larger consumption of raw clams and oysters than the controls. These cases had also received more injections from physicians. Some other factors such as dental procedures and exposure to certain insects showed no differences. Thus, some of the exposures previously suspected of leading to hepatitis were confirmed in this study, and others such as exposure to cockroaches and dental procedures, which were also suspected on clinical inference, were clearly excluded.

As a byproduct of the study, it was found that there were quite distinctive clinical features for cases of hepatitis transmitted through the parenteral and oral routes. Rigorously controlled analytic prospective epidemiologic studies offer opportunities such as this to describe more precisely the pattern of disease because it is necessary to define and categorize cases prior to the study.

In another analytic prospective hospital-based study, the National Center for Radiological Health of the Public Health Service[9] is conducting a cooperative multihospital followup of some 40,000 adult patients with a history of thyrotoxicosis. About three-fourths of these patients had radioactive iodine treatment for thyrotoxicosis, and the remainder had surgery. Both groups are being followed biannually according to a specified protocol. The main objective was to determine if there were more complications (e.g., blood dyscrasias or thyroid cancer among the radioactive iodine treated patients) than would be expected from the experience among surgically treated controls which might be attributable to radiation exposure. Since no treatments were imposed on either the case or control groups, this was an observational analytic study. A systematic followup is being made to determine any differences in outcome.

As with retrospective studies, concern in prospective studies must be with the selection of patients and the possibilities of spurious associations. Generalization of the conclusions is, of course, strictly limited by the sample of patients coming to the hospital in terms of the severity of disease and other selective factors.

The difference between an epidemiologic study and a clinical study sometimes becomes hazy. If the major purpose of an investigation is to test a hypothesis using case and control populations or to derive comparative prevalence and incidence rates in a defined population, it may be said that the study is epidemiologic. On the other hand, a descriptive study of diseases not involving any hypotheses or a control population is a clinical study. With increasing collaboration between clinicians and epidemiologists, the difference becomes unimportant. What is important is to employ epidemiologic methods that are effectively consistent with statistical theory and clinical technology for the conduct of studies that have the greatest potential for improving our knowledge for reducing illness, disability, and mortality, and for improving the prognosis of the patients.

*Clinical trials.*—A particular class of prospective study is that which deals with the testing of the relative efficacy of certain forms of therapy, including drugs and vaccines. As with retrospective studies, concern must be with the problems of selecting patients and the possibilities of spurious association. As in many experimental studies, however, the risk of shortcomings using hospital patients is diminished wherever treatments can be randomly assigned. The generalization of conclusions is, of course, limited to the nature of the hospital population in terms of diseases, their severity, and other selective factors. However, for the purposes of certain treatments, there may be little reason to suspect that there will be a significant difference in effect between nonhospital cases and hospital patients. For example, if aspirin can be shown to be effective in treating rheumatoid arthritis in hospital patients, is there any reason a priori for thinking that the nonhospital sample would respond differently? This is analogous to the dilemma of using volunteers in a vaccine field trial. Volunteers of certain age, race, and sex with a particular socioeconomic or educational background may possibly react differently to a vaccine than would members of the entire community. Nevertheless, if the vaccine shows a high degree of efficacy in the test group, it should also prove to be effective when applied to the whole community.

An example of a study in this area is the one conducted by Chalmers and others[10] on veterans of the Korean war, where hepatitis patients were randomly assigned to one of four hospital wards identical in all aspects of management except for a single factor of treatment. Systematic evaluation of the patients over the short term of several months revealed no differences in the outcome, regardless of the treatment regimen. A 10-year followup of the majority of these patients showed no adverse effects, which could be attributed either to type of exercise or to dietary regimen.[11]

More recently, Garceau and others[12] in a coordinated study of 10 Boston teaching hospitals comprising the Boston Inter-hospital Liver Group tested the efficacy of surgery versus conserva-

tive management of portal hypertension. Patients with liver cirrhosis who fulfilled certain physical and laboratory criteria were invited to participate in the study knowing that they would have a random chance of being selected for surgery or for medical management. This study was important because the criteria for eligibility for surgery were defined prior to the study, and eligible volunteers were randomly assigned to this category. Although patients who were randomly picked for surgery did better than those who were not eligible for surgery because of medical contraindications, they did *no* better than the comparison group—the patients randomly assigned to conservative management. If a rigorous study method can be applied to a procedure as major as portacaval shunt surgery in a disease as variable and potentially life threatening as advanced liver cirrhosis, then this approach can also be applied to most therapeutic questions in medicine. The most important factor is a dedication to discover the true answer to a therapeutic dilemma. As long as the dilemma exists and the patient is frankly informed of his or her risks and the alternatives, then there is little or no infringement of personal or professional ethics in conducting such studies. Rather the investigator should try to discover the answer to such questions as soon as possible in order to provide all patients with the preferred therapy and to spare many a needless procedure.

Currently a long-term, collaborative interhospital drug trial is being started to test the efficacy of various medications on the treatment and survival of patients with coronary heart disease.[13] More than 8,400 patients in 50 clinics across the country will be involved.

Certain coronary patients will receive one of four drugs or a placebo to evaluate their ability to prevent recurrence of myocardial infarction. This prototype of a drug-trial study has unlimited potential for determining the effects of drugs, and it represents a fertile field of mutual endeavor for epidemiologists, statisticians, and clinicians.

Mainland and Sutcliffe[14,15] have provided magnificent examples of experimental studies of arthritis among hospital outpatients. In these studies the Cooperating Clinics Committee of the American Rheumatism Association is engaged in testing the efficacy of various medications on the course of rheumatoid arthritis. The effect of aspirin was studied in a systematically controlled fashion, and, in turn, the effect of other medications will be compared with it.

The same theory that applies to inpatients would apply to outpatient populations whether they be individuals attending a clinic or occupational groups enrolled in hospitalization programs. Another possibility not developed in this report because of the paucity of published reports is the potential of using Blue Cross records and similar hospital insurance programs as a source of cases and controls. It is hoped that eventually these records can be used prospectively to determine comparative morbidity in a study of disease associations or possibly for confirming high-risk morbidity groups identified through independent sources.

## EPIDEMIOLOGIC SURVEILLANCE

The existing system of epidemiologic surveillance based primarily on the reporting of notifiable diseases by physicians is not satisfactory because of the gross incompleteness of reporting and the limited nature of coverage of disease problems. The diagnostic data in hospital records should provide a particularly useful source of information for detection of sudden changes in morbidity.

The key to this problem is to develop case-finding methods which would provide the needed data and at the same time overcome the problem of time lag inherent in discharge data. In some instances admission diagnoses could alert the health official to the need for thorough investigation. Another possibility would be to assign the responsibility to an epidemiologic officer[16,17] within the hospital who could experiment with the most effective means of finding cases of public health concern.

In the use of hospital records, care must be taken to avoid erroneous conclusions due to errors or incompleteness in the observations recorded. In surveillance as in the hypothesis-testing use of hospital records, the observer must not confound differences of etiologic significance with results arising from the fact that patients choose hospitals or are referred to hospitals in nonrandom fashion. Undue concern with this bias has, in recent years, resulted in the underutilization of hos-

pital records as a source of information of great potential value.

Problems are also to be expected in the estimation of the incidence and prevalence of various diseases. Estimates based on hospital admissions will be less valid for chronic diseases than for illnesses characterized by a single acute attack that requires hospitalization and results in immunity of long duration. Furthermore, the use of in-hospital data will cause the loss of large numbers of acute and minor diseases which are diagnosed and treated in physicians' offices or in outpatient departments.

It is recognized that the reporting of diseases and preliminary diagnoses by one hospital or by a few hospitals in a large metropolitan area would be of limited value because patients choose or are referred to hospitals in nonrandom fashion. Bias introduced in this way could lead to differences which would suggest spurious etiologic relationships. However, undue concern with such possibilities in recent years has led to hospital records being underutilized as a valuable source of information. These risks would be greatly reduced if all hospitals in a metropolitan area or other defined geographic area would join in a common computer service for broad administrative and fiscal purposes; this would make possible the systematic collection of admission and discharge diagnoses that would have maximal value for epidemic surveillance.

## MEDICAL-CARE RESEARCH

The use of hospital data for purposes of medical-care research is comparable to its use for epidemiologic purposes. However, rather than to identify the characteristics of patients with specific diseases, the goal in this case is to identify the characteristics of patients who seek various care and to determine the effect of organizational and professional characteristics of the hospital on receipt of care. Such information suggests hypotheses as to why certain patients do or do not receive care and why such care may be costly. In this case the hospital also provides a resource for testing hypotheses.

Obtaining medical care is greatly influenced by the disease process, including symptoms and severity, as well as by such demographic characteristics of the patient as age and sex. Therefore, interpretation of utilization patterns must also take into account diagnoses and associated disabilities. Such diagnostic information is usually more accessible and reliable when obtained from hospital records than from other sources. The uses of these data are discussed below under four headings, each of which represents a major factor influencing patterns of hospital utilization.

### Influence of Diagnoses Per Se

For this purpose discharge data by diagnostic categories are useful. Data for a specific hospital may be of value to the hospital itself in evaluating its functions and in servicing the community. However, these data are usually of limited value for any one hospital in the community since the population served by a given hospital is ordinarily difficult to delineate. If all hospitals in the community are included, discharge data for residents of that community and the population estimates for the total community may be used to establish discharge rates.

Diagnostic data for the country as a whole are available from the Hospital Discharge Survey of the National Center for Health Statistics, which is a probability sample of U.S. hospitals.[18,19] For smaller areas such as States and metropolitan areas, corresponding data are not routinely tabulated. There are, however, collection mechanisms in existence which do have the potential for providing such data for a sizable and definable population subgroup. Among such mechanisms are the Blue Cross and/or commercial hospital insurance programs and certain group practice programs such as the Health Insurance Plan of Greater New York[20-22] and the Kaiser-Permanente Program in California.[23] In addition, the Professional Activity Study at Ann Arbor Michigan, insofar as it serves hospitals in several areas, might provide similar information. Data potentially available from social programs such as Medicare and Medicaid may also be used for this purpose. In Allegheny County, Pennsylvania, the Blue Cross has a program which provides a continuous "patient register" and related statistical tabulations for each hospital.

In many of these medical-care programs, the diagnostic data become available as part of a system designed to pay claims or for some other administrative reason. In others it will be necessary to code a sample of the records. In such cases it is desirable that the diagnostic tabulations have some meaning in terms of the organization's own purposes in order to justify the sizable expense represented by such procedures.[24]

### Influence of Organizational Structure for Providing Medical Care and of Type of Physician

It has been shown that hospitalization rates for members of group practice programs are lower than those for comparable groups serviced by other programs or for groups seeking care independently.[20-22] The differences, however, are not uniform for all diagnostic categories. For example, the difference in hospitalization rates between individuals served by group and by nongroup programs is greater for patients with tonsillectomies and adenoidectomies than for those with neoplasms. Such knowledge is of considerable importance in planning a framework by which medical services are made available to the population. Moreover, since the population is given a choice of several different kinds of medical-care programs (as is the case in the Medicaid program in New York City), it is of more than passing interest to examine the utilization patterns by diagnosis among the several plans. Differences in patterns after appropriate adjustments for population characteristics may yield clues as to differences in medical practice which may, in turn, suggest ways of examining the quality of medical care in different settings.

Discharge rates by diagnoses are required for the study of this problem, and they need to be classified as follows:

1. Type of practice of physician discharging the patient
    a. Solo
    b. Partnership
    c. Group practice
    d. Hospital-based
2. Type of hospital from which patient was discharged
    a. AHA accredited or nonaccredited
    b. Teaching or nonteaching or university or nonuniversity
    c. Auspices (voluntary, proprietary, governmental)
    d. Type of outpatient department organization (speciality clinics only, general medical or speciality clinic, and other)

Various cross-classifications of these and other variables may also be desired. Such classifications require that the necessary information be available on the discharge report and appropriately coded.

### Influence of the Physician's Characteristics

The following classification of the physician's professional characteristics is suggested:

1. Physician specialty (general practitioner, internist, pediatrician)
2. Years since graduation
3. Type of staff appointment (none, courtesy, active, consultant)

Again some degree of cross-classification may be desired.

In this connection, it would be helpful if each physician in a community was assigned a code number, as was done in Allegheny County, Pennsylvania. The code number would identify his type of practice, specialty, board certification, age, and other factors. It would permit grouping of data from various locations—either by physician or by physician characteristics. It would also permit comparisons of the behavior of the same physician in different settings.

### Influence of Various Types of Case Management

The concept of progressive patient care has given rise to a variety of different methods of case management. Since the diagnosis is clearly one of the determinants of case management, one may raise the question of whether subsequent experience of a patient with a particular diagnosis

varies with different methods of case management. Subsequent experience may be measured in many ways, such as readmission rates during a given period after initial discharge, length of time required to return to work, and subsequent use of physician's or nurse's services. These data will not necessarily be available from a routine reporting system. It may be necessary to do a followup on the patient and to obtain information by direct interview. These data may then be correlated with the presence or absence and use of:

1. Formal intensive care unit in the hospital
2. Home care program
3. Nursing home affiliation by type of affiliation
4. Formal discharge planning program.

It will be essential in carrying out such studies to have detailed knowledge of the criteria for admission to a particular kind of case-management program. A control or comparison group is also needed. In general, the requirements for more definitive study here are similar to those in the discussion of hypotheses testing under Prospective Studies.

## CONFIDENTIALITY

A problem which has been raised recently in some of the epidemiologic studies utilizing hospital data has been that of confidentiality and privileged communications. This issue involves the relationship between the hospital staff physician and the patient. In some States much of the information, particularly the diagnosis, obtained and recorded in the hospital records is of a privileged and confidential nature.[25] Making such information available for epidemiologic purposes is considered by some to be a violation of trust. Thus, it is sometimes necessary to obtain the permission of every patient whose record is to be used. It is obvious that this can be a cumbersome and difficult process.

In some instances, special legislation has been passed permitting the use of hospital data for such purposes as medical audits. New York State has recently enacted legislation which gives the State health officer access to individual records.

Edgar S. Dunn[26] has stated that it is necessary to distinguish clearly between intelligence data about individuals as individuals, such as medical records a doctor keeps to trace changes in the wellbeing of his patients, and statistical data that relate to groups of individuals. Statistical data are concerned with aggregates, averages, percentages, and so forth that describe relationships characteristic of groups of, for example, hospital cases. No personal information about the individual needs to be available to anyone other than those engaged in the research. Respect for confidentiality of medical information has been characteristic of epidemiologic studies. The need for the protection of an individual's privacy is well appreciated and understood by individual researchers, and specific steps are usually taken to provide the necessary protection.

# REFERENCES


[1]Masi, A. T.: Potential uses and limitations of hospital data in epidemiologic research. *Am.J.Pub.Health* 55(5):658-668, May 1965.

[2]Levin, M. L., Goldstein, H., and Gerhardt, P. R.: Cancer and tobacco smoking. *J.A.M.A.* 143(4):336-338, May 1950.

[3]Buck, A. A.: Epidemiologic investigations of sarcoidosis. *Am.J.Hyg.* 74(2):137-202, Sept. 1961.

[4]Masi, A. T., and others: Hashimoto's disease, a clinicopathological study with matched controls. *Lancet* 1(7377):123-126, Jan. 1965.

[5]Berkson, J.: Limitations of the application of four-fold table analysis to hospital data. *Biometrics* 2:47-53, 1946.

[6]Winkelstein, W., Jr., Stenchever, M., and Lilienfeld, A. M.: Occurrence of pregnancy, abortion and artificial menopause among women with coronary artery disease, a preliminary study. *J.Chron.Dis.* 7(4):273-286, Apr. 1958.

[7]Grady, G. F., Chalmers, T. C., and the Boston Inter-Hospital Liver Group: Viral hepatitis in a group of Boston hospitals-I. A retrospective study of 1675 patients. *New England J. Med.* 272(13):657-661, Apr. 1965.

[8]Grady, G. F., Chalmers, T. C., and the Boston Inter-Hospital Liver Group: Viral hepatitis in a group of Boston hospitals-II. A prospective controlled epidemiologic study. *New England J.Med.* 272(13):662-666, Apr. 1965.

[9]Saenger, E. L., Thoma, G. E., and Tompkins, E. A.: Incidence of leukemia following treatment of hyperthyroidism, a preliminary report of the Cooperative Thyrotoxicosis Therapy Follow-up Study. *J.A.M.A.* 205(12):855-862, Sept. 1968.

[10]Chalmers, T. C., and others: The treatment of acute infectious hepatitis, controlled studies of the effects of diet, rest, and physical reconditioning on the acute course of the disease and on the incidence of relapses and residual abnormalities. *J.Clin.Invest.* 34(7):1163-1235, July 1955. Part II.

[11]Chalmers, T. C., and Nefzinger, M. D.: The treatment of acute infectious hepatitis, 10-year followup study. *Gastroenterology* 42:472, 1962.(abstract)

[12]Garceau, A. J., the Boston Inter-Hospital Liver Group, and others: A controlled trial of prophylactic portacaval-shunt surgery. *New England J.Med.* 270(10):496-500, Mar. 1964.

[13]National Heart Institute: *The Coronary Drug Project.* PHS Pub. No. 1695. National Institutes of Health. Washington. U.S. Government Printing Office, 1967.

[14]Mainland, D., and Sutcliffe, M. I.: Aspirin in rheumatoid arthritis, a seven-day double-blind trial—preliminary report. *Bull.Rheumat.Dis.* 16:388-391, Nov. 1965.

[15]Cooperating Clinics Committee of the American Rheumatism Association: A seven-day variability study of 499 patients with peripheral rheumatoid arthritis. *Arthritis Rheumat.* VIII(2):302-335, Apr. 1965.

[16]Fuerst, H. T., and Lichtman, H. S.: Directions in hospital epidemiology. *New York J.Med.* 66(10):1202-1205, May 1966.

[17]Lichtman, H. S., and Stuart, G. M.: The Community Oriented Hospital Epidemiologist. Paper prepared for presentation at the 93d annual meeting of the Epidemiology Section of the American Public Health Association, Oct. 21, 1965, Chicago, Ill.

[18]Sirken, M. G.: The hospital discharge survey. *Pub.Health Rep.* 82(1):9-16, Jan. 1967.

[19]National Center for Health Statistics: Patients discharged from short-stay hospitals, United States, October-December 1964. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 13-No. 1. Public Health Service. Washington. U.S. Government Printing Office, Oct. 1966.

[20]Densen, P. M., and others: Prepaid medical care and hospital utilization. American Hospital Association. Chicago Hospital Monograph Series No. 3, 1958. p. 55.

[21]Densen, P. M., Jones, E. W., Balamuth, E., and Shapiro, S.: Prepaid medical care and hospital utilization in a dual choice situation. *Am.J.Pub.Health* 50(11):1710-1726, Nov. 1960.

[22]Densen, P. M., Shapiro, S., Jones, E. W., and Baldinger, I.: Prepaid medical care and hospital utilization, Comparison of a group practice and a self-insurance situation. *Hospitals* 36(22):62, Nov. 1962.

[23]Weissman, A.: A morbidity study of the Permanente Health Plan population. *Permanente Found.M.Bull.* X(1-4):12-26, Aug. 1952.

[24]Densen, P. M., and Shapiro, S.: Methodological problems in the study of special population groups, health insurance and industrial groups. *Ann.New York Acad.Sc.* 107(2):490-505, May 1963.

[25]Hayt. E., and Hayt, J.: *Legal Aspects of Medical Records.* Berwyn, Ill. Physicians' Record Company, 1964. pp. 74, 75, 93.

[26]U.S. Congress, House Committee on Government Operations: The computer and invasion of privacy. Hearings before a subcommittee of the Committee on Government Operations, House of Representatives, 89th Congress, July 26-28, 1966. Washington. U.S. Government Printing Office, 1966. p. 92.


———○○○———


☆ U. S. GOVERNMENT PRINTING OFFICE: 1969—342049/54

## OUTLINE OF REPORT SERIES FOR VITAL AND HEALTH STATISTICS

**Public Health Service Publication No. 1000**

*Series 1.* *Programs and collection procedures.*—Reports which describe the general programs of the National Center for Health Statistics and its offices and divisions, data collection methods used, definitions, and other material necessary for understanding the data.

*Series 2.* *Data evaluation and methods research.*—Studies of new statistical methodology including: experimental tests of new survey methods, studies of vital statistics collection methods, new analytical techniques, objective evaluations of reliability of collected data, contributions to statistical theory.

*Series 3.* *Analytical studies.*—Reports presenting analytical or interpretive studies based on vital and health statistics, carrying the analysis further than the expository types of reports in the other series.

*Series 4.* *Documents and committee reports.*—Final reports of major committees concerned with vital and health statistics, and documents such as recommended model vital registration laws and revised birth and death certificates.

*Series 10.* *Data from the Health Interview Survey.*—Statistics on illness, accidental injuries, disability, use of hospital, medical, dental, and other services, and other health-related topics, based on data collected in a continuing national household interview survey.

*Series 11.* *Data from the Health Examination Survey.*—Data from direct examination, testing, and measurement of national samples of the population provide the basis for two types of reports: (1) estimates of the medically defined prevalence of specific diseases in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics; and (2) analysis of relationships among the various measurements without reference to an explicit finite universe of persons.

*Series 12.* *Data from the Institutional Population Surveys.*—Statistics relating to the health characteristics of persons in institutions, and on medical, nursing, and personal care received, based on national samples of establishments providing these services and samples of the residents or patients.

*Series 13.* *Data from the Hospital Discharge Survey.*—Statistics relating to discharged patients in short-stay hospitals, based on a sample of patient records in a national sample of hospitals.

*Series 14.* *Data on health resources: manpower and facilities.*—Statistics on the numbers, geographic distribution, and characteristics of health resources including physicians, dentists, nurses, other health manpower occupations, hospitals, nursing homes, and outpatient and other inpatient facilities.

*Series 20.* *Data on mortality.*—Various statistics on mortality other than as included in annual or monthly reports—special analyses by cause of death, age, and other demographic variables, also geographic and time series analyses.

*Series 21.* *Data on natality, marriage, and divorce.*—Various statistics on natality, marriage, and divorce other than as included in annual or monthly reports—special analyses by demographic variables, also geographic and time series analyses, studies of fertility.

*Series 22.* *Data from the National Natality and Mortality Surveys.*—Statistics on characteristics of births and deaths not available from the vital records, based on sample surveys stemming from these records, including such topics as mortality by socioeconomic class, medical experience in the last year of life, characteristics of pregnancy, etc.

For a list of titles of reports published in these series, write to:

Office of Information
National Center for Health Statistics
U.S. Public Health Service
Washington, D.C. 20201