Being admitted to hospital signals an important event in an individual's health status and well-being and, concurrently, signals an important use of health care system resources. A valuable record of such events is contained in hospital discharge summary forms. In several countries throughout the world, hospital discharge summaries are compiled into computerized databases, for later use. Databases of this nature have long been used as sources of information for a vast variety of epidemiologic studies, as well as markers of health for population and public health use. Although they contain large volumes of easily accessible health data, difficulties do exist in proper interpretation and analysis of this information. Is the use of hospital discharge databases for surveillance of infectious diseases warranted? What are the current uses, advantages and disadvantages of these databases? What, if any, future direction do they have?

CURRENT USE OF HOSPITAL DISCHARGE DATABASES

A veritable throng of investigations using hospital discharge summary databases are presented in the literature. Large health organizations and influential public health bodies, such as the Centers for Disease Control and Prevention (CDC) in the United States, routinely utilize this type of data. The CDC produces annual estimates of the use of short-stay hospitals in the United States.1,2 The estimates are based on data collected through the National Hospital Discharge Survey that has been conducted annually by the National Center for Health Statistics, since 1965.1 Such indicators of health provide hospital and public health authorities with knowledge of health status of their population, and provide directions for targeted interventions to improve health.

Use of hospital discharge databases is also widespread among public health authorities and epidemiologic researchers throughout the world. No reliable statistics on the incidence of bronchiectasis in developed countries were available until Finnish researchers undertook a collection of data on over 12,500 bronchiectasis cases that occurred between 1972 and 1992, using the discharge register maintained by Finland's National Research and Development Centre for Welfare and Health.3 The finding that bronchiectasis was in marked decline between 1972 and 1992 was attributed to effective treatment of pulmonary infections and the reduced incidence of tuberculosis.3

Hospital discharge databases are not limited to large health organizations. Individual institutions of variable size also use such data. At the Wimera Base Hospital in Australia, screening of inpatient discharge summaries over a 2-year period led to the establishment of a simple, inexpensive, and effective medical quality control system in that medium-sized hospital.4

Of particular benefit in developing countries, hospital discharge databases are a readily available and inexpensive means of assessing the health of the population. In China, little was known about the disease profiles of Chinese living in an urbanized community such as Hong Kong. Accordingly, discharge summaries of 561 acute hospital medical admissions were reviewed, which indicated that cardiovascular diseases were the most important cause of acute medical admissions and mortality.5

The wide use of hospital discharge summary databases in epidemiologic studies and population surveys has revealed their importance as sources of health information. These databases have also found a niche in the field of infectious diseases.

HOSPITAL DISCHARGE DATABASES AND INFECTIOUS DISEASES

Information contained in discharge databases can, at times, be invaluable for health policy and decision making in infectious diseases. The policy of the United States Navy on immunization of high-risk personnel groups with
hepatitis B vaccine was partly based on a study conducted from 1975 to 1984. The epidemiology of viral hepatitis was reviewed using a hospital discharge summary database of all active-duty enlisted personnel admitted to a U.S. Navy treatment facility.

In addition to aiding in policy decisions, discharge databases have a role in the evaluative process of public health interventions in infectious diseases. They can also form an integral part of the intervention itself. The Birth Defects Monitoring Program uses a newborn hospital discharge database to monitor the incidence of congenital rubella syndrome in the United States. In 1989, Cochi and colleagues collected hospital discharge summary information on all cases as reported by the monitoring program from 1970 to 1985, and discovered that congenital rubella syndrome was on the verge of elimination in the United States at the end of 1985.

Owing to elevated rates of infectious diseases in developing countries, use of accessible and affordable information available in discharge databases is indispensable. In the tertiary care, United Bulawayo Hospital in Zimbabwe, a retrospective descriptive study of patients discharge summaries was used to determine the trends of the 10 most common diseases in the medical department. The study presented aggregate data that showed infections and cardiovascular diseases alone accounted for 50% of the diagnoses made in patients discharged from the medical department from 1987 to 1994.

Hospital discharge summary databases have been queried to provide significant contributions to the study of the epidemiology of disease. Although there has been strong historic reliance on these databases, they do have limitations that must be addressed.

QUALITY OF HOSPITAL DISCHARGE DATABASES

Both within and outside the field of infectious diseases, discharge databases are occasionally used with minimal regard to their limitations. Such limitations play an important role in the validity of any study that uses these databases and, therefore, should not be overlooked. In 1995, van Walwaven and Weinberg assessed the completeness of hospital discharge summaries and the efficiency of the discharge summary system in two urban teaching hospitals in Ottawa, Canada. They found considerable deficiencies, including missing discharge diagnoses and missing discharge medications. In 1998, Madsen and coworkers assessed the data quality of septicemia and sepsis registration in a Danish hospital discharge database, by comparing discharge data with data from a computerized bacteremia database at a regional department of clinical microbiology. They concluded that the hospital discharge registry revealed numerous misclassifications, and the system was found not suitable for surveillance of bacteremia at that time.

Deficiencies in discharge data can also be present in a wide range of medical pathologies, including trauma and injuries. It is uncommon to find clinical diagnoses immune to these issues. To determine the adequacy of hospital discharge data for describing the nature and severity of multiple trauma, Kramer and colleagues compared injury data coded from full hospital records with injury data coded from discharge summaries for 83 plane crash survivors admitted to 14 different hospitals in the United States. Discharge summary data missed 52% of the intracranial injuries and 30% of spinal fractures. These differences point to concerns regarding the use of discharge data for determining the nature, severity, and sequelae of injuries.

Bias resulting from incomplete discharge data can be bidirectional, at times underestimating, and at times overestimating rates of disease. For example, in 1993, a Scottish study undertook to determine the accuracy of hospital discharge data and death certificates coded as motor neuron disease (MND). Data from the Scottish Motor Neuron Disease Register (SMNDR) was compared with routinely collected Scottish Hospital In-Patient Statistics (SHIPS), using the registry as the gold standard. The sensitivity of a diagnosis of MND, as retrieved by SHIPS, was 84% and the positive predictive value was 70% overall. From these and other results, the authors concluded that coded hospital discharge data were an inaccurate record of a diagnosis of MND and, in their present form, would underestimate disease incidence in Scotland. However, in a study by Kennedy and co-workers the opposite situation arose. To test their hypotheses on population incidence trends for myocardial infarction (MI), they hoped to use hospital discharge codes for MI (ICDCM-410) as a convenient proxy for incidence trends. To evaluate the accuracy of medical records, those of patients discharged with an MI code were compared to an independent cardiology surveillance study of all patients with acute MI admitted to a large county teaching hospital. Over a 12-month period, 110 patients were coded as ICDCM-410 in medical records, but only 67 of these cases were detected by cardiology surveillance. The 43 patients not detected were miscoded as having MI even though there was no evidence of it. The authors expressed concern that this may conceal a true downward trend in the incidence of coronary artery disease.

Although some investigations have illustrated several limitations in discharge databases, numerous studies have remarked on the great potential of this resource. In 1997, Mahonen and colleagues studied the validity of the Finnish hospital discharge register data on coronary heart disease (CHD) for the purposes of epidemiologic studies and health services research. The Finnish nationwide hospital discharge register was linked with the FINMONICA acute myocardial infarction (AMI) register for the years 1983 to 1990. Data on CHD from the
discharge register gave, on average, an accurate account of changes in the occurrence of AMI in Finland, and the authors concluded that, with necessary caution, discharge register data can be used in epidemiologic studies and health services research.

In 1998, Paltiel and colleagues assessed the independent contribution of data sources used for case ascertainment in a clinical audit of the outcome of patients with lymphoma. In the absence of a unified register of patients with lymphoma diagnosed and treated at the Hadassah-Hebrew University Hospital in Jerusalem, Israel, sources for case ascertainment included two internal databases (computerized discharge data for hospitalizations and outpatient oncology database) and one external (Israeli Cancer Registry) database. All three sources contributed independently to the total, with the majority (82%) via discharge data from hospitalizations. Overlap among the sources was minor, with only 10% of the patients being common to all three sources. This indicated that use of a single internal data source may have resulted in an underestimate of the scope of lymphoma in this institution; however, hospital discharge data were capable of encompassing the majority of these patients.

Finally, Koobatian and co-workers compared the sensitivity of passive reporting of Guillain-Barré (GBS) syndrome to the Vermont Department of Health from 1980 to 1985 with that of computerized hospital discharge abstract data. In all, 51 definite and probable cases of GBS were identified from hospital discharge data during a period when only four cases had been reported to the health department through passive physician reporting. Although limited by the lack of timeliness for public health surveillance, computerized hospital discharge data were readily available and were more sensitive in detecting cases compared to passive surveillance. They are a useful tool for establishing baseline rates and examining long-term trends for selected acute diseases, such as GBS, for which there are well-established diagnostic criteria and that usually result in hospitalization.

USING HOSPITAL DISCHARGE DATABASES FOR INFECTIOUS DISEASE SURVEILLANCE IS WARRANTED, WITH CAUTION

Hospital discharge databases are affected by limitations due to miscoding, missing diagnoses, and incomplete capture of cases. Persons consulting physicians in the community and in nonhospital centers for infectious diseases are excluded from such databases. For many infectious diseases this is the predominant circumstance. Individuals seen in hospital emergency departments but not admitted, and fatalities not admitted to hospital are also excluded from hospital discharge databases. Furthermore, validity of discharge diagnoses cannot be confirmed using the databases alone. Timeliness of information may also be limited by the availability of data from such databases. Caution must be exercised when using information from these databases for epidemiologic research and health care planning and policy.

However, the large volume of information readily available makes hospital discharge databases an attractive and cost-effective source for epidemiologic investigation. Hospital discharge databases, when properly used in infectious diseases, can be efficient tools for surveillance, hypothesis generation and testing, health trend identification, and epidemiologic study. Risk assessment and intervention evaluation are possible with the well-planned and careful use of these databases. As a result, the answer to the question of whether the use of hospital discharge databases for surveillance of infectious diseases is warranted, is yes, with caution. The issue of caution revolves around our necessary commitment to the understanding of database limitations and to the accordance of proper weight to database information, taking into account these limitations. Funding agencies are often overly reliant upon these databases to make funding decisions; consequently, it is important to study and verify their value. Not doing so creates the risk of overfunding or underfunding of programs.

A greater engagement in developing approaches to resolve database limitations will help reduce funding risks. Potential underestimation and overestimation biases for rates of disease can be minimized by according greater importance to accurate coding of diagnoses in the databases. This crucial element has often been underappreciated by database users and managers alike. In Quebec, Canada, the provincial computerized database of hospital discharge summaries, known as MED-ÉCHO (Médecine et Exploitation des Données pour l’Étude de la Clientèle Hospitalière), encompasses all hospitals in the province. Discharge summary forms are completed by the attending physician of the admitted patient, to help ensure accurate coding of diagnoses. Furthermore, trained hospital archivists review the charts of each admitted patient and confirm the diagnoses presented on the discharge summary form, making alterations as necessary before the information is entered into the database. This “double-check” system is a novel approach to ensuring data quality, and has been adopted by other databases in Canada and worldwide.

The use of discharge databases differs in the developed world compared with that in the developing world, as seen in previous examples. In developing countries, databases are used as a readily available and inexpensive means of assessing the health of the population, whereas in developed countries, their usage is extended to interventions, intervention evaluation, epidemiologic research, and risk assessment. However, the importance of data quality crosses country and continent barriers. Poor data is not better than no data, but data of which
limitations are well studied and verified are always potentially useful.

As in the past, hospital discharge databases for surveillance of infectious diseases will continue to be important sources of information well into the future. With the advent of the Internet and the increasing global availability of health information, these databases may take on an even greater role in infectious disease surveillance (Koutsavlis AT. Unpublished data). Recent advances in telecommunication technology have been enormous. The National University of Singapore (NUS) is one of the first institutions to exploit the use of the Internet for the delivery of health information databases. Its World Wide Web server was established in 1993 by the NUS Biocomputing Research and User Support (BRUS) technology group, in collaboration with the Computer Resource Planning committee of the Faculty of Medicine. One of their most important achievements was the development of the first World Wide Web implementation of a health and population statistical database which contains information for Singapore and selected Asian countries, and aggregate data for world regions.

CONCLUSION

The importance of the cautious use of hospital discharge databases for surveillance of infectious diseases cannot, and should not be overlooked. Continued improvement in the quality and scope of these databases should be a surveillance priority and should take advantage of the advancing technologies of the twenty-first century.

REFERENCES