

Appendix

A Summary of Studies on the Quality of Health Care Administrative Databases in Canada

Introduction

Most of the information in the Atlas is based on health care data collected for administrative reasons such as determining the eligibility of beneficiaries of public insurance programs, documenting discharges from hospital, and paying providers' claims for medical services or prescriptions filled under provincial drug plans. Researchers use the data to study the utilization of health care even though the databases were not created for these purposes. There have been extensive studies and reviews on the quality of health care data in the United States,¹ but there has not been a systematic review of studies on the quality of health care data published in Canada. In preparing for the second Atlas, ICES conducted a review of published and unpublished studies in Canada to assess the completeness of data and the levels of agreement across databases. This Appendix summarizes our findings. A more

comprehensive technical paper is available from ICES.

Researchers basically work with three levels of information from health care databases. The first level of data comprises demographic characteristics of patients, particularly age, sex, and place of residence. The second level of data includes information on diagnoses. The most responsible diagnosis defines the reason for patients receiving health services. The secondary diagnoses provide information on concomitant health problems (comorbidity) and complications arising from the disease and the management of it. Diagnosis is usually coded according to the 9th revision of the International Classification of Disease (ICD-9). The third major level of data includes information on the diagnostic, medical and surgical procedures provided by physicians. There are two classification systems for coding procedures in Canada — the Classification of Diagnostic, Therapeutic and Surgical Procedures (CCP) developed by Statistics Canada,

and the Clinical Modification of the ICD-9 (ICD-9-CM) developed for use in the United States. As hospitals in Canada may use either classification system, the Canadian Institute for Health Information (CIHI) has tables for translating ICD-9-CM codes into CCP codes for use on hospital discharge summaries. The 10th revision of the ICD will be introduced by 1998, and both procedure coding systems will be revised and updated as well.

In undertaking the review, we employed three criteria for assessing the quality of the data. The first criterion was completeness of the data. The completeness of the data was judged by the extent to which the database covered the population and the availability of information on demographic characteristics, diagnoses and procedures. The second criterion was agreement of information when data from one database was compared to the same information obtained from reabstraction of original records by a health records technologist,

another database, or clinical records. The third criterion pertained to agreement of diagnosis with expert criteria. For explicit reviews, researchers have extracted clinical information from medical records, applied explicit criteria developed by groups of experts to derive the diagnosis, and compared the “standard” diagnosis with the coded diagnosis on the record. A second approach, involving implicit review, has been to ask “expert” clinicians to review the information in the records and assess whether or not they considered the diagnosis on the abstract to be the “correct” diagnosis.

Completeness of Data

Hospital Discharge Data

Most Canadian provinces and territories submit hospital discharge data to CIHI that describes the services patients receive from a hospital on an inpatient or day surgery basis. In Manitoba, the data are submitted to the government, but hospitals may voluntarily report to CIHI as well. CIHI collects discharge data for about 34% of hospital patients in Manitoba. The government of Quebec maintains its own databases with none of the hospitals submitting data to CIHI. All provinces and territories submit

hospital discharge data to Statistics Canada, but the timeliness of the data is reduced by delays in submissions before the database is complete for a given year.

Residents who go out of province for hospital services are not included in the provincial database. Coyte and associates² estimate that services received outside of Ontario comprise less than 0.5% of all procedures performed for Ontario residents.

Researchers at ICES routinely check the databases for missing information when conducting analyses. Typically, they have found that less than 1% of the records have missing information for age, sex, and residence codes (Exhibit A.1). Similar results have been reported for other provinces³.

Physician Billings Data

It is estimated that 95% of all physicians in Canada are paid on a fee-for-service basis. Claims are submitted to provincial health insurance plans for payment. Services provided by physicians on salary or paid for by patients, Workers’ Compensation, or other third party payers are not included in the medical claims database. In Ontario, physicians whose services are covered by

Alternate Funding Plans do not file claims, so their services are excluded from the OHIP claims database as well.

Researchers at ICES and in other provinces have found that less than 1% of the demographic information on physicians’ claims is missing.^{4,7} Most provinces do not require a diagnosis on claims for payment, so it is essentially an optional field on the claim. Each province and territory has its own Schedule of Benefits that lists a fee and fee code for each service provided. The service codes are specific to the specialty of the physician and do not necessarily relate to classification systems used in other health care databases. Alberta requires an ICD-9 diagnosis code and a CCP procedure code for payment, and Nova Scotia is moving in this direction as well.

For each physician who submits claims, all provinces and territories create statistical summaries on the age and gender of their patients and the services received by patients. The statistical summaries are sent to CIHI for inclusion in the National Physician Database (NPDB). The codes in the provincial schedule of benefits are translated into slightly over 100 service codes. The NPDB can be used to provide

Exhibit A.1: Summary of Canadian Studies Reviewing Completeness of Demographic Data

Author	Database	Years Examined	Population Covered by Database	Completeness of Demographic* Data (%)
Hospital Discharge Data				
Chen ¹⁹	Ontario inpatients	1989 - 1992	100% of residents	97
Jha ³⁰	Ontario inpatients	1992 - 1993	100% of residents	99
Mustard ⁵	Manitoba Health Services	1989 - 1991	100% of residents	99
Ugnat ³¹	Ontario inpatients	1991 - 1992	100% of residents	99
Physician Billings				
Platt ⁶	Alberta Health Care Insurance	1984 - 1989	100% of residents	99
Svenson ³²	Alberta Health Care Insurance	1984 - 1989	100% of residents	99
Provincial Drug Plans				
Anderson ⁸	British Columbia Pharmacare	1981 - 1982 1986 - 1987 1988 - 1989	100% of 65+ years	96 - 97
Davidson ⁹	New Brunswick Drug Plan	1990 - 1991	100% of 65+ years	100
Guess ¹⁰	Saskatchewan Drug Plan	1983	95% of residents	100
Rawson ³³	Saskatchewan Drug Plan	1976 - 1987	95% of residents	100
Tamblyn ⁷	Régie de l'Assurance Maladie du Québec	1990	100% of 65+ years	99
Thiesson ¹²	Saskatchewan Drug Plan	1984	95% of residents	100

* Demographic information includes the patient's age, sex and place of residence

basic profiles of the age/sex-specific rates of service utilization by region of the province.

Provincial Drug Plans

All provinces and territories have plans that pay for prescription drugs for seniors, but they vary in the coverage they provide other residents. Saskatchewan and British Columbia are the only two provinces with plans that cover the entire population. In all provinces, the plans are for prescribed drugs dispensed from community pharmacies and do not include drugs provided through hospitals or other facilities that pay for drugs out of their global budgets. The provinces also vary in the drugs that they choose to insure and the types of data included in the drug plan databases. Researchers have found that less than 1% of the basic information on patients is missing in the drug plan databases.⁸⁻¹² There is no national drug claims file.

The Ontario Drug Benefit (ODB) program contains information about

the beneficiaries, the dispensers, and the prescribers of medications. Each claim contains information about the drugs prescribed, the dose, the cost of the product, and the date the prescription was filled. ICES researchers have found that there is very little missing information.

Agreement of Hospital Discharge Data with Reabstraction Studies or Data from Other Records

Reabstraction Studies

There were six reabstraction studies included in the review. The reabstraction studies by the Ontario Hospital Association, the Ministry of Health in Newfoundland and Doctors Hospital in Toronto, Ontario were reviews of records selected at random from the participating facilities. The study by the Ontario Hospital Association¹³ included 3,000 records from 43 hospitals. The reabstraction study conducted by

the Ministry of Health in Newfoundland¹⁴ included 850 records from six acute care hospitals. Doctors Hospital¹⁵ conducted its own reabstraction study of 300 records. In all three studies, the agreement on demographic and administrative data was 95% or higher. The rate of agreement on most responsible diagnosis was 81% in the OHA study, 74% in the Newfoundland study, and 96% in the Doctors Hospital review. The rates of agreement for secondary diagnoses were lower — 37% in the OHA study, 59% in the Newfoundland study, and 95% in the Doctors Hospital study. The rates of agreement for procedures were high, ranging from 88% in the OHA study to 96% at Doctors Hospital.

We reviewed four studies where researchers reabstracted charts as part of clinical studies. These are summarized in Exhibit A.2. Delfino and associates,¹⁶ and Rawson and Malcolm¹¹ found that rates of agreement on demographic data ranged from 88% to 100%. The rates of agreement

Exhibit A.2: Summary of Canadian Reabstraction* Studies of Hospital Records

Author	Province	Records Abstracted	Years Examined	Data	Agreement (%)
General Studies **					
Ontario Hospital Association ¹³	Ontario	3,000 records in 43 hospitals	1988 - 1989	Demographic Diagnosis *** Procedure	93 - 100 37 - 81 88 - 95
Newfoundland Department of Health ¹⁴	Newfoundland	850 records in six hospitals	1994	Demographic Diagnosis *** Procedure	98 - 100 59 - 74 93
Doctors Hospital, Toronto ¹⁵	Ontario	300 records in one hospital	1992	Demographic Diagnosis *** Procedure	100 95 - 96 96
Specific Studies					
Delfino ¹⁶	Quebec	1,279 records with a respiratory diagnosis in 14 hospitals	Not stated	Demographic Diagnosis	88 - 99 75 - 95
Malenka ¹⁸	Manitoba	485 records with a prostatectomy procedure in one hospital	1974 - 1980	Comorbidity	42 - 71
Rawson ¹¹	Saskatchewan	444 records with a cholecystectomy or hysterectomy procedure in 14 large hospitals	1986	Demographic Diagnosis Procedure	95 - 100 42 - 71 97 - 100
Ray ¹⁷	Saskatchewan	236 records of hip fracture in 10 large hospitals	1984 - 1985	Diagnosis	99

* Reabstraction studies of hospital records assess the agreement between the abstracted information in the database and the information in the chart that was recorded by the physician and then coded by health records staff.

** General studies are based on a random sample of all diagnoses and procedures.

*** Top end of range reflects agreement for most responsible diagnosis; bottom end of range reflects agreement for secondary diagnosis.

on diagnoses were varied. Rawson and Malcolm¹¹ reabstracted the records of patients who had a hysterectomy or cholecystectomy and found the rates of agreement for procedures were over 95%. The rates of agreement for the most responsible diagnosis were between 42% and 71%. Delfino and associates¹⁶ reabstracted the records of patients in Montreal hospitals with respiratory diseases as the most responsible diagnoses, and the rates of agreement by hospital ranged from 75% to 95%. The highest rate of agreement for most responsible diagnosis was 99% for hip fracture.¹⁷ In a study of comorbidity related to prostatectomy, Malenka and associates¹⁸ found the rates of agreement for secondary diagnoses were between 42% and 71%.

Hospital Discharge Data Compared to Data From Other Sources

We found six studies in which researchers compared information in the hospital discharge summary with data from other administrative or clinical databases. The data are summarized in Exhibit A.3. For example, in reviewing obstetrical records from three hospitals, Chen¹⁹ found the rate of agreement to be 95% or over for type of delivery. There were

generally high levels of agreement, 90% or greater, for primary procedures. There was one exception to this general finding. In comparing medical claims to hospital data for patients undergoing 11 procedures, Roos and associates²⁰ found that while the overall agreement was 90%, the rates varied from 77% (vascular surgery) to over 98% for several procedures. The rates of agreement for most responsible diagnoses ranged from 20% to 80%, with an average of 75%. Iron and associates²¹ compared OHIP records and hospital data for women and found the rates of agreement to be 94% for hysterectomy and 93% for cholecystectomy.

Researchers^{18,22} in Manitoba have undertaken a number of comparisons to examine the completeness of recording of secondary diagnoses for patients undergoing cholecystectomy, prostatectomy, and other procedures. Generally speaking, the recording of secondary diagnoses was substantially lower on the hospital discharge summary than on medical claims or clinical databases. Agreement on secondary diagnoses is important if researchers are to use comorbidities in comparisons of clinical outcomes across hospitals or jurisdictions.

Matching Diagnoses in Hospital Records with Expert Clinical Criteria

Eight studies compared the diagnoses on hospital records with expert clinical criteria. These studies are summarized in Exhibit A.4. They provide an estimate of the usefulness of hospital data for formal epidemiologic research and the clinical evaluation of medical interventions.

The rate of agreement for diagnosis was lowest for the diagnosis of Guillain-Barré Syndrome in Ontario, at 21%.²³ Using the World Health Organization's criteria for acute myocardial infarction, which are reasonably well established, the rates of agreement were 80% or better in an Ontario hospital,²⁴ and in studies conducted in Nova Scotia and Saskatchewan.²⁵ The Nova Scotia-Saskatchewan Group assessed acute myocardial infarction for the years 1977, 1981, and 1985 in both provinces. They found the rates of agreement with the World Health Organization's criteria for recorded acute myocardial infarction improved over time, from 81% in 1977 to 89% in 1985. The rates in 1985 were better for Saskatchewan (92%) than Nova Scotia (87%).

There is more ambiguity around the diagnoses of stroke and asthma,

Exhibit A.3: Summary of Canadian Studies Comparing Diagnoses and Procedures Between Hospital Discharge Data and Another Source							
Author	Province	Records Abstracted	Years Examined	Variable Examined	Data Against Which Database Being Checked	Agreement (%)	
Chen ¹⁹	Ontario	3,357 obstetrical records in three hospitals	1989 - 1992	Procedure	Clinical dataset	95 - 99	
Roos ³⁴	Manitoba	3,131 hysterectomy and cholecystectomy records in all hospitals	1974	Procedure	Physician claims	89 - 91	
Roos ²⁰	Manitoba	61,310 records for 11 procedures in all hospitals	1979 - 1983	Procedure	Physician claims	77 - 98	
Iron ²¹	Ontario	8,467 records of women who had a hysterectomy or cholecystectomy in all hospitals	1991 - 1992	Procedure	Physician claims	93 - 94	
Malenka ¹⁸	Manitoba	485 records with a prostatectomy procedure in one hospital	1974 - 1980	Comorbidity	Physician claims	35	
Roos ²²	Manitoba	112,000 records from all adult surgical procedures except obstetrics in one hospital	1979 - 1984	Comorbidity	Clinical dataset	17 - 90	

where the levels of agreement between hospital records and experts ranged from 69% to 80%.^{26,27} Using the World Health Organization's strict definition of stroke in a study at a teaching hospital in Nova Scotia, Phillips and associates²⁸ found the agreement with hospital diagnosis dropped to 35%. Rheumatoid arthritis can also be difficult to diagnose precisely. Tennis and associates²⁹ applied five criteria from the American Rheumatism Association to Saskatchewan hospital patients with a diagnosis of rheumatoid arthritis; only 45% of the charts met all five criteria. Another 16% of the charts listed three or four of the criteria.

It should be understood that these levels of disagreement reflect not only the coding process, but also the vagaries of clinical diagnosis for some conditions where physicians must work around nosological ambiguities. In any event, depending upon the disease entity under investigation, one may or may not be able to rely

solely upon the diagnostic information in the hospital discharge summary.

Comment

We draw the following conclusions from our review of Canadian studies of data quality in administrative databases.

- Demographic information on patient age, sex and residence is complete and reliable.
- Generally, there are high levels of agreement on specific surgical procedure codes found in hospital discharge data and medical claims. The use of drugs, laboratory tests, and X-rays is not routinely recorded or abstracted on hospital databases.
- Hospital data on the most responsible diagnosis vary in completeness and accuracy. Diagnoses such as acute myocardial infarction or fracture are reasonably reliable. Diagnostic data for conditions such as stroke are substantially less reliable, and the greatest disagreement with expert criteria-based reviews occurs with diagnoses such as rheumatoid arthritis where clinicians themselves may disagree.
- Clinical data on secondary diagnoses, comorbidities, and complications are less likely to be recorded accurately and comprehensively in hospital discharge abstracts, and the rates of agreement on case-mix may be accordingly low.
- The ICD-9-CM and CCP codes have not kept pace with developments in technology and clinical practice, and this limits the degree to which certain procedures are specified in the discharge database.
- Billing claims for physician services typically provide complete capture of procedure codes but these codes may not necessarily match those used in hospital records.

Exhibit A.4: Summary of Canadian Studies Comparing Diagnoses Between Hospital Discharge Data and External Criteria

Author	Province	Records Abstracted	Years Examined	Diagnosis	External Criteria	Agreement (%)
Tennis ²⁹	Saskatchewan	432 records from all hospitals	1978 - 1980	Non-arthritic Osteoarthritis Rheumatoid Arthritis	American Rheumatism Association	100 99 45
Young ³⁵	Manitoba	817 records in one diabetic program	1986	Diabetes	Expert Judgement	93
Van Walraven ²⁴	Ontario	25 records in one hospital	1987 - 1988	Acute Myocardial Infarction	World Health Organization	80
Nova Scotia - Saskatchewan Cardiovascular Disease Epidemiology Group ^{25*}	Nova Scotia and Saskatchewan	2,869 records in two provinces in all hospitals	1977 1981 1985	Acute Myocardial Infarction	World Health Organization	78 - 94
Mayo ²⁶	Quebec	96 records from five hospitals	Not stated	Stroke	Expert Judgement Claims	70 - 80
Sweet ²⁷	Prince Edward Island	423 records from five hospitals	1984 - 1988	Asthma	Expert Judgement	69
Phillips ²⁸	Nova Scotia	381 records from one hospital	1988 - 1989	Stroke	World Health Organization	35
McLean ²³	Quebec and Ontario	2,333 records from all hospitals	1983 - 1989	Guillain-Barré Syndrome	Expert Judgement	21 - 26

* Studies conducted in 1991. Authors chose three points in time respectively

Improving the quality of these databases requires attention to several issues:

- Physicians and other health professionals must be fastidious in recording relevant information, and be educated about some of the controversies in disease and procedure classification. Linkage possibilities would be expanded if both fee-for-service and non-fee-for-service clinical encounters were coded in the manner now accepted in Alberta and under consideration in Nova Scotia.
- Epidemiologists, health records technologists, and clinicians should meet to address some of the areas where current coding and classification systems are ambiguous.
- The quality of coding for procedures and diagnoses must be improved to the greatest extent possible. Continuing education and regular audits, which are already integral to the operations of health records departments of large hospitals, must be augmented and extended to institutions of all sizes.

References

1. Wray NP, Ashton CM, Kuykendall DH et al: Using administrative databases to evaluate the quality of medical care: a conceptual framework. [Review]. *Soc Sci Med* 1995;40:1707-1715
2. Coyte PC, Young W, Williams JI: Devolution of hip and knee replacement surgery? *ICES Working Paper* 1995;38:1-35
3. Roos LL, Mustard CA, Nicol JP et al: Registries and administrative data: organization and accuracy. *Med Care* 1993;31:201-212
4. Johansen H, Dutta M, Mao Y et al: An investigation of the increase in preschool-age asthma in Manitoba, Canada. *Health Reports* 1992;4:379-402
5. Mustard CA, Harman CR, Hall PF et al: Impact of a nurses' strike on the cesarean birth rate. *Am J Obstet Gynecol* 1995;172:631-637
6. Platt GH, Svenson LW, Woodhead SE: Coronary artery bypass grafting in Alberta from 1984 to 1989. *Can J Cardiol* 1993;9:621-624
7. Tamblyn R, Lavoie G, Petrella L et al: The use of prescription claims databases in pharmacoepidemiological research: the accuracy and comprehensiveness of the prescription claims database in Quebec. *J Clin Epidemiol* 1995;48:999-1009
8. Anderson GM, Kerluke KJ, Pulcins IR et al: Trends and determinants of prescription drug expenditures in the elderly: data from the British Columbia Pharmacare Program. *Inquiry* 1993;30:199-207
9. Davidson W, Molloy DW, Somers G et al: Relation between physician characteristics and prescribing for elderly people in New Brunswick. *Can Med Assoc J* 1994;150:917-921
10. Guess HA, Strand L, Helston D et al: Fatal upper gastrointestinal hemorrhage or perforation among users and nonusers of nonsteroidal anti-inflammatory drugs in Saskatchewan, Canada 1983. *J Clin Epidemiol* 1988;41(1):35-45
11. Rawson NS, Malcolm E: Validity of the recording of cholecystectomy and hysterectomy in the Saskatchewan health care datafiles. Saskatoon: Pharmacoepidemiology Research Consortium, 1995
12. Thiessen BQ, Wallace SM, Blackburn JL et al: Increased prescribing of antidepressants subsequent to beta-blocker therapy. *Arch Intern Med* 1990;150:2286-2290
13. Ontario Hospital Association, Ontario Ministry of Health, Hospital Medical Records Institute: Report of the Ontario data quality reabstracting study. Toronto: Ontario Hospital Association, 1991
14. Newfoundland Department of Health: Report of the Newfoundland reabstracting project. St. John's: Newfoundland, 1995
15. The Doctors Hospital: Data quality and utilization study. Toronto: Doctors Hospital, 1992
16. Delfino RJ, Becklake MR, Hanley JA: Reliability of hospital data for population-based studies of air pollution. *Arch Environ Health* 1993;48:140-146
17. Ray WA, Griffin MR, Downey W et al: Long-term use of thiazide diuretics and risk of hip fracture. *Lancet* 1989;April 1:687-690
18. Malenka DJ, McLerran D, Roos N et al: Using administrative data to describe casemix: a comparison with the medical record. *J Clin Epidemiol* 1994;47:1027-1032
19. Chen E. Reliability of HMRI data among selected obstetrical populations. Institute for Clinical Evaluative Sciences in Ontario unpublished manuscript 1995;1-4
20. Roos LL, Sharp SM, Wajda A: Assessing data quality: A computerized approach. *Soc Sci Med* 1989;28(2):175-182
21. Iron K, Goel V, Williams JI: Concordance of hospital discharge abstracts and physician claims for surgical procedures in Ontario. *ICES Working Paper* 1995;42:1-18

22. Roos LL, Sharp SM, Cohen MM: Comparing clinical information with claims data: some similarities and differences. *J Clin Epidemiol* 1991;44:881-888
23. McLean M, Duclos P, Jacob P et al: Incidence of Guillain-Barré syndrome in Ontario and Quebec, 1983-1989, using hospital service databases. *Epidemiology* 1994;5: 443-448
24. van Walraven C, Wang B, Ugnat AM et al: False-positive coding for acute myocardial infarction on hospital discharge records: Chart audit results from a tertiary centre. *Can J Cardiol* 1990;6(9):383-386
25. The Nova Scotia-Saskatchewan Cardiovascular Disease Epidemiology Group. Trends in incidence and mortality from acute myocardial infarction in Nova Scotia and Saskatchewan 1974 to 1985. The Nova Scotia-Saskatchewan Cardiovascular Disease Epidemiology Group. *Can J Cardiol* 1992;8:253-258
26. Mayo NE, Danys I, Carlton J et al: Accuracy of hospital discharge coding for stroke. *Can J Cardiol* 1993;9(Suppl D):121-124
27. Sweet L, Fan JSK, O'Neill J: A validation study of asthma, Prince Edward Island, 1984-1944: II. Hospitalizations. *Chronic Diseases in Canada* 1992;13(2):19-23
28. Phillips SJ, Cameron KA, Chung CR: Stroke surveillance revisited. *Can J Cardiol* 1993;9(Suppl D): 124-125
29. Tennis P, Bombardier C, Malcolm E et al: Validity of rheumatoid arthritis diagnoses listed in the Saskatchewan Hospital Separations Database. *J Clin Epidemiol* 1993;46:675-683
30. Jha P, Sykora K, DeBoer D et al: Characteristics and mortality outcomes of thrombolysis trial participants and non-participants: A population-based comparison. *J Am Coll Cardiol* 1996;27:1335-1342
31. Ugnat AM. Data quality. Institute for Clinical Evaluative Sciences in Ontario unpublished manuscript 1995;1-11
32. Svenson LW, Platt GH, Woodhead SE: Geographic variations in the prevalence rates of Parkinson's disease in Alberta. *Can J Neurol Sci* 1993;20:307-311
33. Rawson NS. An acute adverse drug reaction alerting scheme using the Saskatchewan health datafiles: Results for two nonsteroidal anti-inflammatory drugs. *Drug Invest* 1993;6(5):245-256
34. Roos LL, Roos N, Cageorge SM et al: How good are the data? Reliability of one health care data bank. *Med Care* 1982;20(3):266-276
35. Young TK, Roos NP, Hammerstrand KM: Estimated burden of diabetes mellitus in Manitoba according to health insurance claims: a pilot study. *Can Med Association J* 1991;144:318-324

