

## ORIGINAL PAPER

N.S.B. Rawson · E. Malcolm · C. D'Arcy

**Reliability of the recording of schizophrenia and depressive disorder in the Saskatchewan health care datafiles**

Accepted: 25 June 1996

**Abstract** Administrative data have long been used in psychiatric epidemiology and outcomes evaluation. This article examines the reliability of the recording of schizophrenia and depressive disorder in three Saskatchewan administrative health care utilization datafiles. Due to their comprehensive nature, these datafiles have been used in a wide range of epidemiologic studies. Close agreement was found between hospital computer data and patients' charts for personal and demographic factors ( $\geq 94.7\%$ ). Diagnostic concordance between computerized hospital data and medical charts was very good for schizophrenia (94%) but poor for depressive disorder (58%). Appropriate physician services were identified for 60% and 72% of hospital discharges for schizophrenia and depressive disorder, respectively, and *exact* diagnostic agreement between hospital and physician datafiles was 62% for schizophrenia and 66% for depressive disorder. Appropriate provincial mental health branch services were found for 83% and 38% of hospital discharges for schizophrenia and depressive disorder, respectively; *exact* diagnostic concordance between these datafiles was 75% for schizophrenia and 0% for depressive disorder. A significant number of patients with major or neurotic depression appeared to be

wrongly coded as having depressive disorder in the hospital file. The differences in diagnostic agreement may also be partly a function of how the two conditions are differentially treated in the health system. These findings suggest that more specific and severe psychiatric diagnoses are likely to be recorded accurately and consistently in the Saskatchewan datafiles. However, disorders with multiple manifestations or those for which there are several possible codes should be examined with caution and ways sought to validate them. Attention should also be paid to which service sectors are involved in the treatment of specific disorders.

---

**Introduction**

Computerized data on hospital discharges, physician payment claims and case registries have long been used in psychiatry to investigate the etiology, treatment, and outcome of various disorders. The potential biases in such sources have been appreciated by some, but more commonly these data have been uncritically accepted as valid and reliable. Recently, linked administrative health care utilization datafiles have been recognized as a major resource in epidemiologic research; in particular, the computer files of the Canadian province of Saskatchewan have been promoted [1–5] and used in a variety of studies [6–15]. Some researchers have validated their results from these datafiles [8, 11, 12, 16], but, as with most other administrative health care systems [17], little work has been done on their accuracy and reliability from a wider perspective.

This work examined aspects of the accuracy and reliability of diagnostic codes for schizophrenia and depressive disorder in the Saskatchewan datafiles. Two surgical procedures, cholecystectomy and hysterectomy, and two groups of medical diagnoses, ischemic heart disease and chronic obstructive pulmonary disease, have been considered elsewhere [18–20]; the

---

N.S.B. Rawson

College of Pharmacy and Nutrition, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

E. Malcolm

Research Services, Strategic Programs Branch, Saskatchewan Health, Regina, Saskatchewan, Canada

C. D'Arcy

Applied Research, Department of Psychiatry, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

N.S.B. Rawson (✉)

Royal University Hospital (Box 92), 103 Hospital Drive, Saskatoon, Saskatchewan, Canada S7N 0W8

procedures and specific medical diagnoses are accurately and reliably recorded in different datafiles, but less specific diagnoses are variable. Psychiatric diagnoses were included in this study because such disorders are sometimes less specific than physical ones and they lack usable biological markers for making a diagnosis. In addition, mental illnesses are a special group of diseases that are treated differently, both therapeutically and organizationally [21, 22].

With the resources available, it was only possible to include two psychiatric diagnoses in the study and they were selected to be as diverse as possible. Schizophrenia was chosen because it is a serious, chronic, and, as psychiatric illnesses go, more precisely defined disorder that would more likely involve treatment by a psychiatric specialist and/or a direct service government agency. Depressive disorder is a less distinct diagnosis, more likely to be treated in a wider range of modalities by a broader range of medical practitioners.

## Methods

The Canadian health care system covers medically necessary hospital and physicians' services for all residents regardless of their ability to pay [23]. Most hospitals are publicly funded and physicians are predominantly paid from public sources on a fee-for-service basis. Although jointly financed by federal and provincial taxation, the management of the system is almost totally a provincial responsibility, which leads to variation in the way services are provided in each of the ten provinces. For example, in some, prescription drugs are mainly provided through a provincial drug plan, whereas in others only the private market exists.

As the first jurisdiction in North America to supply both hospital and physician services to all residents [1, 2], Saskatchewan has been a pioneer in providing an extensive range of publicly funded health care programs. Currently, the provincial government's programs include hospital, physician and mental health services, prescription drugs (including psychotropic medications), and cancer therapy. The accounting systems for most of these programs have been computerized for many years [1-3]. Each eligible resident (99% of the population of just over one million) is issued with a unique health services number (HSN), which is recorded in the relevant datafile when a service is provided [2].

Saskatchewan residents can receive care for psychiatric and emotional problems from private practice psychiatrists and family physicians, as well as private practice psychologists, social workers and assorted other counsellors. As in most of Canada, access to specialist care in Saskatchewan is by referral from a general practitioner.

However, individuals can refer themselves directly to specialist services offered by the provincial mental health branch. If an admission to hospital for a psychiatric condition is required, patients may be hospitalized in a general hospital ward, a psychiatric ward of a general hospital, or a psychiatric hospital. As illustrated in Table 1, these services often result in records for psychiatric events being maintained by more than one branch of Saskatchewan Health, the provincial health department.

The starting point of our analysis was the identification of two groups of patients from information covering 1986 held in the datafile of the Hospital Services Branch of Saskatchewan Health. This file contains information about discharges from all provincial general hospitals. In 1986, there were 134 hospitals comprising 6 large teaching and research hospitals situated in the province's two main cities providing a full range of services to residents in their immediate areas and acting as the ultimate referral centers for Saskatchewan (base hospitals), 1 large hospital offering rehabilitation facilities, 7 regional hospitals in five smaller cities serving as referral centers for patients from outside their immediate vicinity and providing hospital care to residents in their locale, and 120 community or district hospitals of which 80% had 30 or fewer beds offering more limited services to their local population [2]. The year 1986 was chosen because the research design included the use of data from the provincial prescription drug plan [1-3], which changed in July 1987 for an 18-month period so that information was not recorded on an individual basis during that time.

### Patient selection criteria

The first group of patients consisted of all individuals discharged from hospital in 1986 with a primary discharge diagnosis of schizophrenic psychosis [*International Classification of Diseases* (ICD) code 295] and the second group consisted of all patients with a primary diagnosis of nonspecific depressive disorder (ICD code 311) [24]. Depressive disorder is defined as "states of depression, usually of moderate but occasionally of marked intensity, which have no specifically manic-depressive or other psychotic depressive features and which do not appear to be associated with stressful events or other features specified under neurotic depression".

Since a patient could be included in each of the two groups more than once and individuals with several hospitalizations could unduly influence the results, the first admission for each patient (designated as the index hospitalization) was selected for analysis. For each patient, their sex, age, residence code (city, town, village, or rural), primary and secondary discharge diagnoses recorded using four-digit ICD codes, procedure code recorded using the *Canadian Classification of Diagnostic, Therapeutic and Surgical Procedures* [25], and admission and discharge dates (the former being calculated from the latter and the length of stay) were extracted from the hospital discharge file. These data, together with the dates on which provincial health service coverage began and, if appropriate, finished and an indicator of whether the patient had died, were supplied for analysis after substituting a unique dummy

**Table 1** Cross-recording of psychiatric events in different provincial datafiles

Psychiatric event	Computerized data system		
	Physician service claim	Hospital discharge	Provincial mental health branch
Physician office visit	Yes	-	-
Hospitalization: general ward	Yes	Yes	-
Hospitalization: psychiatric ward	Possibly <sup>a</sup>	Yes	Yes
Hospitalization: psychiatric hospital	Possibly <sup>a</sup>	-	Yes
Provincial mental health service	-	-	Yes

<sup>a</sup> Dependent upon whether the physician was salaried or worked on a fee-for-service basis

identification number for the HSN. Three methods were used to evaluate the accuracy and reliability of the health datafile information: external consistency (comparing with hospital medical charts), internal consistency (comparing hospital, physician service, and mental health datafiles), and contextual consistency.

#### External consistency

Ideally, all the patients' medical charts would have been used in the comparison with the hospital computer data, but it was only financially possible to abstract around 200 cases from each diagnostic group. Since the base and regional hospitals were more readily accessible than small hospitals in rural communities, random samples of patients with schizophrenia or depressive disorder were selected only from individuals discharged from these hospitals.

Using a form developed by two of the authors (N.S.B.R. and E.M.), information from the charts was collected by health records abstractors specially employed by Saskatchewan Health. The abstractors were trained in the use of the form, and a small pilot exercise was performed to test it. Some minor changes were subsequently made, but these did not make the pilot forms incompatible with the main set. The form included checks about the agreement between computer data and chart information with regard to the patient's demographics and discharge date. In addition, the abstractors were required to provide an *exact* copy of the discharge diagnoses for the index hospitalization. Recoding of the diagnoses was done by the principal researcher (N.S.B.R.) from the completed forms without reference to the computer data to avoid bias. Each form included the relevant dummy patient identification to enable linkage of chart information with data from the hospital discharge file.

#### Internal consistency

The second way in which the reliability of information about the patients was evaluated involved internal comparisons of data from the hospital discharge file with data from the physician service claim file and from the files of the provincial mental health branch. Data on publicly funded, private practice medical services are recorded by the Medical Care Insurance Branch (MCIB), since this is the section of Saskatchewan Health to which physicians submit claims for their services in order to get paid. The MCIB is the custodian of the physician claims datafile, each record of which contains the patient's HSN, the physician's identification number, the service date and code, and the diagnostic code (three-digit ICD). The physician's fee is not related to the diagnosis but to the service provided. Service codes and a fee for each code have been established by the MCIB in consultation with the Saskatchewan Medical Association. Since the payment is for the service, there is no particular financial incentive to record any diagnosis more than another.

A physician's payment claim for work performed in a hospital that is recorded in the MCIB datafile may be linked to the same patient's hospital data by the HSN. Therefore, for both groups of patients, the physician claims file was searched for services with matching HSNs and dates and a diagnostic code from the ICD "Mental Disorders" chapter, excluding mental retardation codes (317–319). If more than one physician service was found for a patient, the one closest to the index admission date was selected and linked with the hospital discharge data. The agreement between diagnoses at the three-digit level from the two files was quantified.

Patients may also receive service for psychiatric disorders from clinicians directly employed by the Mental Health Services Branch (MHSB) of Saskatchewan Health. The MHSB oversees mental health services directly provided by the provincial government, and all eligible residents may receive the benefits of these services [21, 22]. Data in the computer files of the MHSB cover all inpatients treated in psychiatric facilities and general hospital psychiatric units, as well as all outpatient clients treated by MHSB staff (Table 1).

Information in the file includes patient, service, and diagnostic data, the latter being recorded using four-digit ICD codes up to 1985 and five-digit *Diagnostic and Statistical Manual of Mental Disorders, Third Revision* (DSM-III) [26] codes subsequently. For both schizophrenia and depressive disorder, there are equivalent codes in both the ICD and DSM-III systems. MHSB data covering hospitalization services can also be linked with the hospital data using the HSN and, for both patient groups, this file was searched for services with the same diagnostic codes. Again, if more than one service was found, the one closest to the index admission date was used. The agreement between diagnoses in the hospital and MHSB files was evaluated at the three-digit level.

#### Contextual consistency

Internal consistency may also be examined by exploring contextual information in the hospital discharge, physician service, and prescription drug plan datafiles associated with the two diagnoses. For the two periods between the beginning of 1984 and a patient's index admission date in 1986 and between their discharge date and the end of 1987, the hospital and physician service files were searched and counts made of ICD codes for nonorganic psychoses (295–299) for each schizophrenia patient and the depressive disorder code (311) for each depressive disorder patient. For both groups of patients, counts were also made of prescriptions for antidepressants, neuroleptics, benzodiazepines, other sedative-hypnotics, and lithium recorded in the drug plan datafile between the beginning of 1984 and admission and between discharge and the end of June 1987 (individual drug data were not available for the second half of 1987).

## Results

The numbers of patients discharged at least once from Saskatchewan hospitals in 1986 with primary diagnoses of schizophrenia or depressive disorder were 646 and 828, respectively. The calendar months of the admission dates in each group were examined to assess whether there was any bias towards the early part of the year resulting from the method of selecting the index hospitalization, but none was evident.

The proportions of males in the schizophrenia and depressive disorder groups were 53.7% and 30.4%, respectively. The age distributions of the two patient groups are shown in Figs. 1 and 2; the median ages and ranges of the male and female schizophrenia patients were 34 years (14–85) and 45 years (15–85), respectively, while for the depressive disorder patients they were 57 years (13–94) and 60 years (10–96), respectively.

The distributions of days in hospital were highly positively skewed, the medians and ranges for the schizophrenia and depressive disorder groups being 15 days (1–1371) and 9 days (1–567), respectively. The proportion of patients for whom no hospital procedure was recorded was 71.1% in the schizophrenia group and 80.6% in the depressive disorder group. Almost 3% of the schizophrenia patients had a procedure code indicating that a psychiatric commitment evaluation had been performed and just over 4% in both groups had electroconvulsive therapy. Considerably larger proportions had a chest X-ray (18.1%

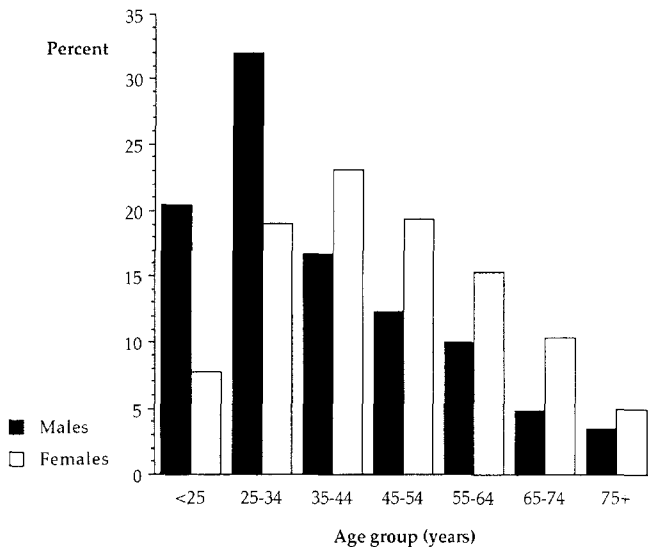


Fig. 1 Age distribution of the 646 schizophrenia patients

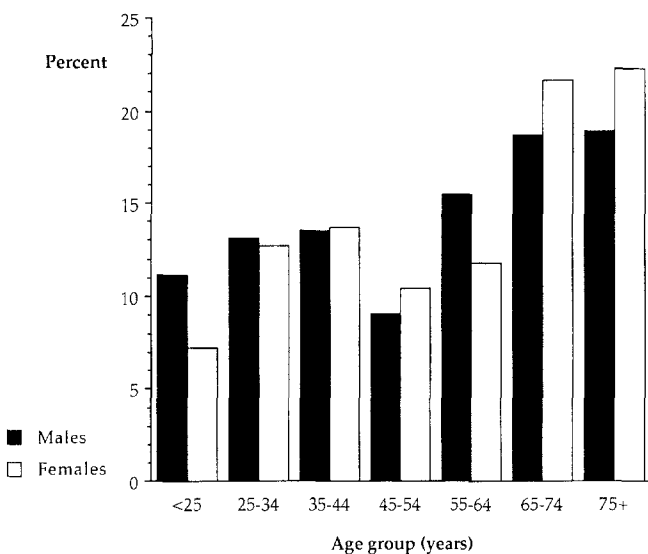


Fig. 2 Age distribution of the 828 depressive disorder patients

of the schizophrenia and 7.4% of the depressive disorder patients).

#### External consistency: hospital data and medical chart concordance

Of the 646 schizophrenia and 828 depressive disorder discharges, 565 (87.5%) and 352 (42.5%), respectively, were from base or regional hospitals. Random samples of 225 schizophrenia and 235 depressive disorder patients were chosen for abstraction, but it was not possible to obtain access to the charts of psychiatric patients admitted to three regional hospitals, which decreased the sample sizes to 135 and 156, respectively.

The charts of four other schizophrenia patients could not be accessed, further reducing the schizophrenia sample to 131.

The sex distributions of both samples were virtually identical to those of the populations from which they were drawn, the proportions of males in the two samples being 53.4% and 30.1%, respectively. With the exception of male schizophrenia patients, whose median age was 34 years, the median ages of the abstraction subsamples were lower than those of the populations from which they were taken (female schizophrenia patients: 39 years; male and female depressive disorder patients: 43 and 51 years, respectively). The rates of agreement between hospital computer data and chart information for chart number, first and last names, HSN, sex, birth month and year, residence location, and discharge date were all 94.7% or more.

There was exact agreement between primary hospital and chart discharge diagnoses for 101 (77.1%) of the 131 schizophrenia patients at the ICD four-digit level (Table 2) and for 123 (93.9%) at the three-digit level; the value of kappa [27] calculated from Table 2 was 0.76, indicating good concordance. Nevertheless, eight patients had nonschizophrenic chart diagnoses, one of whom was recorded as having multiple sclerosis.

Similar comparative data for the patients with depressive disorder are shown in Table 3. There was exact concordance for only 91 patients (58.3%), which implies only chance agreement, although 146 (93.6%) had some form of "depression." A chart diagnosis of chronic, major, or endogenous depression was recorded for 40 patients (25.6%), depressive reaction or dysthymia (neurotic depression) for 9 patients (5.8%), and depressive adjustment disorder for 4 patients (2.6%). Three patients (1.9%) had "attempted suicide" as their principal chart diagnosis.

#### Internal consistency: hospital, physician and mental health datafiles agreement

A physician service claim with one of the mental disorder diagnoses and a service date during the hospitalization was found for 385 (59.6%) of the 646 schizophrenia patients (88.6% with service on the admission date or the next day). Exact concordance between physician service and primary hospital discharge diagnoses (at the ICD three-digit level) was recorded for 238 (61.8%) of these patients (Table 4). MHSB hospitalizations with an appropriate diagnosis and date were identified for 539 (83.4%) of the schizophrenia patients (98.5% with service on the same or following day), and Table 4 shows the MHSB diagnoses for these patients; exact concordance was found for 406 (75.3%), which was regarded as good.

Physician services with a mental disorder diagnosis and a date during the hospitalization period were found for 596 (72.0%) of the 828 depressive disorder

**Table 2** Schizophrenia patients – concordance between chart and hospital primary discharge diagnoses ( $n = 131$ ). Exact concordance: four-digit level, 101 (77.1%); three-digit level, 123 (93.9%); kappa = 0.76

Chart diagnosis	Hospital primary diagnosis Schizophrenic psychosis							Total
	ICD code	Simple 295.0	Paranoid 295.3	Acute episode 295.4	Residual 295.6	Schizoaffective 295.7	Unspecified 295.9	
Schizophrenic psychosis								
Hebephrenic type	295.1	–	–	–	–	–	1	1
Paranoid type	295.3	–	33	–	1	–	2	36
Residual schizophrenia	295.6	–	1	–	17	–	2	20
Schizoaffective type	295.7	–	–	–	–	37	2	39
Unspecified	295.9	–	4	3	6	–	14	27
All schizophrenia	295	–	38	3	24	37	21	123
Other diagnoses		1	2	1	1	1	2	8

**Table 3** Depressive disorder patients – concordance between chart and hospital primary discharge diagnoses ( $n = 156$ )

Chart diagnosis <sup>a</sup>	ICD code	<i>n</i>	%
Manic-depressive psychosis, depressed type (chronic, major, or endogenous depression)	296.1	40	25.6
Manic-depressive psychosis, circular type, mixed (mixed bipolar disorder)	296.4	1	0.6
Manic-depressive psychosis, other & unspecified (manic-depressive disorder)	296.6	1	0.6
Neurotic depression (depressive reaction; dysthymia)	300.4	9	5.8
Depressive adjustment reaction (depressive adjustment disorder)	309	4	2.6
Depressive disorder (depression)	311	91	58.3
Personality disorder	301	3	1.9
“Attempted suicide”		3	1.9
Other diagnosis		4	2.6

<sup>a</sup> Terms in parentheses are alternatives used in the charts

**Table 4** Schizophrenia patients – physician service claim and provincial mental health branch diagnoses

Diagnostic category	ICD code	Physician service claim		Provincial mental health branch	
		<i>n</i>	%	<i>n</i>	%
Schizophrenic psychoses	295	238	61.8	406	75.3
Affective psychoses	296	29	7.5	32	5.9
Paranoid states	297	12	3.1	7	1.3
Other nonorganic psychoses	298	25	6.5	17	3.2
Neurotic disorders	300	30	7.8	13	2.4
Personality disorders	301	9	2.3	3	0.6
Depressive disorder	311	25	6.5	–	–
Alcohol and drug problems	303, 304	13	3.4	9	1.7
Other mental health diagnoses		4	1.0	9	1.7
Observation and evaluation for suspected conditions	V71	–	–	36	6.7
Other V-codes		–	–	7	1.3
Total		385	100.0	539	100.0

patients (93.1% with service on admission or the next day). Exact concordance between physician service and primary hospital discharge diagnoses was obtained for 395 (66.3%) of the patients (Table 5). MHSB hospitaliz-

ations with the selected diagnoses and appropriate dates were identified for 312 (37.7%) of the depressive disorder patients (91.3% with service on the same or following day), and the diagnoses for these patients are

**Table 5** Depressive disorder patients – physician service claim and provincial mental health branch diagnoses

Diagnostic category	ICD code	Physician service claim		Provincial mental health branch	
		<i>n</i>	%	<i>n</i>	%
Schizophrenic psychoses	295	13	2.2	5	1.6
Affective psychoses	296	59	9.9	224	71.8
Other nonorganic psychoses	297, 298	9	1.5	6	1.9
Neurotic disorders	300	80	13.4	40	12.8
Personality disorders	301	11	1.8	–	–
Adjustment reaction	309	15	2.5	4	1.3
Depressive disorder	311	395	66.3	–	–
Alcohol and drug problems	303, 304	11	1.8	13	4.2
Other mental health diagnoses		3	0.5	7	2.2
Observation and evaluation for suspected conditions	V71	–	–	9	2.9
Other V-codes		–	–	4	1.3
Total		596	100.0	312	100.0

**Table 6** Schizophrenia patients – psychiatric diagnoses and prescriptions between January 1984 and admission and between discharge and December 1987<sup>a</sup>

Diagnoses and prescriptions	Before admission		After discharge	
	<i>n</i>	%	<i>n</i>	%
<i>Physician service datafile:</i>				
Schizophrenic psychoses (ICD 295)	424	65.6	429	66.4
Affective psychoses (ICD 296)	160	24.8	127	19.7
Paranoid states (ICD 297)	73	11.3	52	8.0
Other nonorganic psychoses (ICD 298)	160	24.8	127	19.7
<i>Hospital discharge datafile:</i>				
Schizophrenic psychoses (ICD 295)	263	40.7	312	48.3
Affective psychoses (ICD 296)	74	11.5	61	9.4
Paranoid states (ICD 297)	17	2.6	11	1.7
Other nonorganic psychoses (ICD 298)	27	4.2	14	2.2
<i>Prescription drug plan datafile:</i>				
Antidepressants	215	33.3	162	25.1
Neuroleptics	495	76.6	502	77.7
Lithium	81	12.5	63	9.8
Benzodiazepines	272	42.1	182	28.2
Other sedative-hypnotics	110	17.0	98	15.2

<sup>a</sup>The “after discharge” period for the prescription drug plan data is to the end of June 1987

presented in Table 5. None of the MHSB records had a diagnostic code of 311.

#### Contextual consistency of hospital, physician, and prescription drug datafiles

Information about events occurring between the start of 1984 and the patient's admission and between discharge and the end of 1987 (end of June 1987 for prescription data) in the schizophrenia group is shown in Table 6. In both hospital discharge and physician service data, the proportions of patients receiving services for affective psychoses, paranoid states, or other nonorganic psychoses were up to five percentage points lower in the posthospitalization

period, whereas the proportion with schizophrenic psychosis was higher in this period. The proportion of individuals receiving neuroleptics was marginally greater in the posthospitalization period, but those for the other types of drugs, especially benzodiazepines, declined.

The proportions of depressive disorder patients receiving antidepressants and lithium were higher after hospitalization, 72.9% and 7.5% respectively, than in the preadmission period, 66.2% and 6.6%, respectively (data not shown). However, the proportions receiving drugs in the other three groups were lower in the posthospitalization period, with the largest difference being in benzodiazepines (neuroleptics: 30.0% to 24.2%; benzodiazepines: 69.8% to 50.4%; other sedative-hypnotics: 18.5% to 16.8%).

## Discussion

The validity of administrative datafiles like those of the Saskatchewan health system may be investigated in more than one way. We started with hospitalizations because data about them were expected to be more accurately recorded since there is greater rigor in the capture and recording of data in hospitals than in physicians' offices [17]. Due to limited resources, we were unable to proceed from medical charts to the hospital data, which would have allowed us to calculate estimates of sensitivity and specificity.

It is important to understand that this work was not an evaluation of the accuracy and reliability of the diagnostic abilities of physicians but of the information recorded in the provincial health care utilization datafiles. There has been increased use of the Saskatchewan health care datafiles for research purposes, often without external validation, which makes an evaluation of their accuracy and reliability critical. In retrospect, the choice of depressive disorder may not have been the most useful one, since we would not expect anyone to rely solely on health care utilization datafiles to study the epidemiology of nonspecific psychiatric disorders. Nevertheless, our investigation did allow us to compare the reliability of the recording of a more specific psychiatric disorder with that of a nonspecific condition, which is of value.

In any health care utilization data system, many individuals with a variety of qualifications and training are involved in the recording of information, and it is not surprising that errors occur, especially with transpositions of numbers; it is imperative that they be kept to a minimum. However, the value of a data system as a resource for epidemiologic research depends not only on the accuracy of the recording of information, but also on the internal consistency of the data [17]. Therefore, this study included comparisons within and between datafiles and comparisons with external information.

The examination of the medical charts of reasonably representative samples of patients discharged from base or regional hospitals showed that, for both diagnostic groups, there was close agreement between charts and hospital computer data on personal and demographic factors ( $\geq 94.7\%$ ). This concordance is comparable with those from random samples of Medicaid and Medicare patients [28], although that study was not limited to teaching centers and larger receiving hospitals.

In comparisons with hospital discharge computer data, appropriate physician service claims were identified for 59.6% of the schizophrenia patients, and exact concordance between hospital and physician service diagnoses occurred in 61.8% of these record pairs, which is little more than chance agreement (Table 4). However, appropriate MHSB hospitalization records

were identified for a higher proportion of patients (83.4%), with exact diagnostic concordance being recorded for 75.3%. From 1985 onwards, MHSB diagnoses have been coded using DSM-III, but this should not cause a problem because there is a direct correspondence between the three-digit codes of ICD and DSM-III for both of the study diagnoses. The higher rate of matching records between hospital discharge and MHSB files and the higher diagnostic concordance are not surprising given the large role played by the MHSB in treating schizophrenia, as demonstrated by the fact that 61.6% of the schizophrenia patients were registered MHSB clients in 1986.

A comparison of hospital data with external chart information (Table 2) showed a high degree of accuracy in the diagnostic code, with 77.1% and 93.9% being correctly coded at the ICD four- and three-digit levels, respectively; the latter figure is comparable with that of 89.7% for psychosis (codes 295 to 298) obtained in an assessment of the reliability of psychiatric diagnoses recorded in the New York City Medicaid system [29]. These results suggested that the coding of chronic psychiatric illnesses is accurate. They also illustrated the benefit, if not the necessity, of including MHSB data in linkage studies of these conditions, as others have found [7].

On the other hand, the addition of MHSB data in the analysis of the depressive disorder patients appears to be more confusing than clarifying. Appropriate physician service claims were identified for 72.0% of these patients and exact diagnostic concordance was found for 66.3%, which again is little more than chance agreement (Table 5), whereas appropriate MHSB hospitalization records were located for only 37.7% of the patients and none had a diagnostic code of 311, even though the meaning of this code is the same in both ICD and DSM-III. However, 54.2% of the patients with a matched MHSB record had a diagnostic code of 296.2, which in DSM-III is a single episode of major depression, 11.2% had a code of 296.3 (recurrent major depression), and 11.5%, a code of 300.4 (neurotic depression or dysthymia). These DSM-III codes do not adequately correspond with ICD codes; the closest match for DSM-III 296.2 and 296.3 is ICD 296.1 (manic-depressive psychosis, depressed type), which includes unipolar depression. This suggests that many patients may have been wrongly coded as having a depressive disorder or nonspecific depression when they really had major, chronic, or neurotic depression. The results of the chart abstraction (Table 3), where 25.6% had a chart diagnosis of major, chronic, or endogenous depression and 5.8%, one of neurotic depression, strongly supported this conclusion.

While comparisons between datafiles are vital to substantiate their internal consistency and comparisons with medical charts are important to establish external consistency, context analyses are also relevant. These can take a variety of forms including the

identification of “logically time-sequenced relationships” [30] between events that should or should not occur before or after a hospitalization (e.g., physician services for gynecologic disorders should precede an admission for a hysterectomy but should rarely occur afterwards), which have been used with success previously [18,20,31,32]. Our results confirmed that schizophrenia patients consume a high level of health care services and, due to the chronic nature of their disease, do so both before and after their index hospitalization (Table 6). The proportions of schizophrenia patients receiving neuroleptics and depressive disorder patients receiving antidepressants in the posthospitalization period were marginally greater than the corresponding figures in the preadmission period. In addition, both diagnostic groups exhibited a sharp decline in the use of benzodiazepines between the pre- and post-hospitalization periods, which suggests that the guidelines for more appropriate prescribing of these drugs may be being followed.

Little use has been made of the Saskatchewan health care datafiles for the study of psychiatric disorders, in spite of work in the 1970s advocating their usefulness for service delivery evaluation [4,5]. In 1987, Babiker [7] used the MHSB and prescription drug plan datafiles to evaluate the efficacy of neuroleptics in preventing the readmission to hospital of patients with schizophrenia during a 2-year period. By using the MHSB datafile to identify patients hospitalized for schizophrenia, the study focused on more severely ill patients. In our data, patients with a diagnosis of schizophrenia in both the hospital and MHSB files had the same readmission rate as those with a diagnosis of schizophrenia in the hospital file only (48%), but the rate of use of neuroleptics in the 2 years after hospitalization in the former group (80.3%) was significantly greater ( $P < 0.05$ ) than that in the latter group (73.3%). A different picture of the efficacy of these drugs may have emerged if the patients had been identified from the hospital datafile.

Much of the evidence for and against an association between  $\beta$ -blocker therapy and depression is based on studies using drug database information alone (with antidepressant use as a marker for depression) [10,33,34] or in combination with physician service claim diagnoses [35,36], including data from Saskatchewan [10,35,36]. The work of Thiessen et al. [10] in 1990 showed that the incidence of concurrent antidepressant use was significantly higher in  $\beta$ -blocker users when compared with a reference group receiving no  $\beta$ -blockers (relative risk 2.6; 95% confidence interval 2.3–3.0), but the study was based entirely on data from the prescription drug plan and, therefore, no evaluation of diagnoses was performed. Diagnostic data from the physician claims datafile have been included in more recent work [35,36]. In cohorts of patients dispensed  $\beta$ -blockers, anti-hypertensives, or diuretics concurrently with antidepressants, the proportion with a phys-

ician service claim with a diagnosis of “depression” ranged from only 18% to 31% [35]. In our data, 66% of patients with depressive disorder severe enough to require hospitalization were dispensed antidepressants before admission and 73% received them after discharge and, in the 72% of hospitalizations for which a physician claim was identified, there was diagnostic agreement in only 66% of the cases. These results indicated the inadequacy of using health care utilization datafiles alone to identify patients with “depression.”

In conclusion, this analysis showed that, for schizophrenia (a chronic psychiatric illness), the concordance between hospital data and chart information was good; that there was little more than chance diagnostic agreement between hospital discharge and physician service data (where appropriate claims were identified), although the consistency was much better between hospital and MHSB data; that the contextual setting was clinically realistic. However, for a more varied psychiatric illness (depressive disorder), the results were, not surprisingly, much poorer, possibly because there are at least four three-digit ICD codes (296, 298, 300 and 311) under which “depression” may be coded, most with several fourth-digit subcodes. In addition, the availability of more specific codes in DSM-III in the MHSB data may have added to the problem. We emphasize again that this is an evaluation of the reliability of information in the Saskatchewan administrative health care utilization datafiles, which represent how patients were *treated*, irrespective of whether the diagnosis was accurate and met “gold standard” criteria, which is of importance to applied researchers [29].

Although the reliability of data relating to schizophrenia and depressive disorder is not necessarily generalizable to other psychiatric diagnoses, these findings suggest that researchers considering the use of the Saskatchewan datafiles for epidemiologic studies of psychiatric disorders are likely to obtain valid results for more specific, chronic illnesses as long as they include the use of the MHSB datafiles in their analysis. However, results for disorders, whether psychiatric or not, for which there are multiple, potential diagnostic codes should be examined with caution and ways to validate them, either by cross reference with other data or external information, must be sought. Attention should also be paid to which service sectors are involved in the treatment. Within these limitations, greater use should be made of the Saskatchewan datafiles for the epidemiologic study of psychiatric disorders in a large, defined, representative population.

**Acknowledgements** The authors thank Yvonne Byers-Selinger, Nora Earl, Darleen Gelowitz, Lori Moskal, Rosemary Samson, and Grace Yam for the abstraction of the charts. Dr. Rawson acknowledges the financial support of a Socio-Health Evaluation Grant from the Saskatchewan Health Research Board (now Health Services Utilization and Research Commission) and career awards from Merck Frosst Canada Inc. and the Medical Research Council of



Canada. Dr. D'Arcy acknowledges research program support from Saskatchewan Health. Although this study was based on data provided by Saskatchewan Health, the interpretation and conclusions are those of the authors and do not necessarily represent those of Saskatchewan Health or the Government of Saskatchewan.

## References

1. Rawson NSB, D'Arcy C, Blackburn JL, Bowen RC, Wallace SM, Downey W, Senft D (1992) Epidemiologic research using linked computerized health care datafiles in Saskatchewan, Canada. *Psychiatric Pharmacoepidemiology Research Consortium Technical Report no. 2*, Saskatoon, Saskatchewan
2. Malcolm E, Downey W, Strand LM, McNutt M, West R (1993) Saskatchewan Health's linkable data bases and pharmacoepidemiology. *Post Market Surveill* 6: 175-264
3. Strand LM, Downey W (1994) Health databases in Saskatchewan. In: Strom BL (ed) *Pharmacoepidemiology*, 2nd edn. Wiley, Chichester, pp 217-229
4. Cassell WA, Smith CM, Rankin MG (1970) Psychiatric services in Saskatchewan under Medicare. *Can Psychiatr Assoc J* 15: 63-71
5. D'Arcy C, Vandem Ham M, Goldie S (1976) The development of a comprehensive psychiatric service utilization research datafile. *Can J Public Health* 67: 237-248
6. Beck RG, Horne JM (1980) Utilization of publicly insured health services in Saskatchewan before, during and after copayment. *Med Care* 18: 787-806
7. Babiker IE (1987) Comparative efficacy of long-acting depot and oral neuroleptic medications in preventing schizophrenic recidivism. *J Clin Psychiatry* 48: 94-97
8. Guess HA, West R, Strand LM, Helston D, Lydick EG, Bergman U, Wolski K (1988) Fatal upper gastrointestinal hemorrhage or perforation among users and non-users of non-steroidal anti-inflammatory drugs in Saskatchewan, Canada 1983. *J Clin Epidemiol* 41: 35-45
9. Ray WA, Griffin MR, Downey W (1989) Benzodiazepines of long and short elimination half-life and the risk of hip fracture. *JAMA* 262: 3303-3307
10. Thiessen BQ, Wallace SM, Blackburn JL, Wilson TW, Bergman U (1990) Increased prescribing of antidepressants subsequent to  $\beta$ -blocker therapy. *Arch Intern Med* 150: 2286-2290
11. Spitzer WO, Suissa S, Ernst P, Horwitz RL, Habbick B, Cockcroft D, Boivin J-F, McNutt M, Buist AS, Rebeck AS (1992) The use of  $\beta$ -agonists and the risk of death and near death from asthma. *N Engl J Med* 326: 501-506
12. Ray WA, Griffin MR, West R, Strand L, Melton LJ (1990) Incidence of hip fracture in Saskatchewan, Canada, 1976-1985. *Am J Epidemiol* 131: 502-509
13. Tennis P, Andrews E, Bombardier C, Wang Y, Strand L, West R, Tilson H, Doi P (1993) Record linkage to conduct an epidemiologic study on the association of rheumatoid arthritis and lymphoma in the province of Saskatchewan, Canada. *J Clin Epidemiol* 46: 685-695
14. Rawson NSB (1993) An acute adverse drug reaction alerting system using the Saskatchewan health datafiles. *Drug Invest* 6: 245-256
15. Risch HA, Howe GR (1994) Menopausal hormone usage and breast cancer in Saskatchewan: a record-linkage cohort study. *Am J Epidemiol* 139: 670-683
16. Tennis P, Bombardier C, Malcolm E, Downey W (1993) Validity of rheumatoid arthritis diagnoses listed in the Saskatchewan hospital separations database. *J Clin Epidemiol* 46: 675-683
17. Rawson NSB, D'Arcy C (1991) "Validity" and reliability: idealism and reality in the use of computerized health care databases for pharmacoepidemiological research. *Post Market Surveill* 5: 31-55
18. Rawson NSB, Malcolm E (1995) Validity of the recording of cholecystectomy and hysterectomy in the Saskatchewan health care datafiles. *Pharmacoepidemiology Research Unit Technical Report no. 3*, Saskatoon, Saskatchewan
19. Edouard L, Rawson NSB (1996) Reliability of the recording of hysterectomy in the Saskatchewan health care system. *Br J Obstet Gynaecol* 103: 891-897
20. Rawson NSB, Malcolm E (1995) Validity of the recording of ischaemic heart disease and chronic obstructive pulmonary disease in the Saskatchewan health care datafiles. *Stat Med* 14: 2627-2643
21. D'Arcy C (1983) The Saskatchewan Mental Health Information System: exposition and comment. In: Bennett EM, Trute B (eds) *Mental health information systems: problems and prospects*. Edward Mellen Press, New York, pp 131-150
22. Health and Welfare Canada (1990) *Mental health services in Canada, 1990*. Minister of Supply and Services, Ottawa
23. Nair C, Karim R (1993) An overview of health care systems: Canada and selected OECD countries. *Health Rep* 5: 259-279
24. World Health Organization (1977) *Manual of the international statistical classification of diseases, injuries and causes of death*. 9th rev. World Health Organization, Geneva
25. Statistics Canada (1986) *Canadian classification of diagnostic, therapeutic and surgical procedures*. Minister of Supply and Services, Ottawa
26. American Psychiatric Association (1980) *Diagnostic and statistical manual of mental disorders*, 3rd edn. American Psychiatric Association, Washington, DC
27. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33: 159-174
28. Demlo LK, Campbell PM, Brown SS (1978) Reliability of information abstracted from patients' medical records. *Med Care* 16: 995-1005
29. Schwartz AH, Perlman BB, Paris M, Schmidt K, Thornton JC (1980) Psychiatric diagnoses as reported to Medicaid and as recorded in patient charts. *Am J Public Health* 70: 406-408
30. Roos LL, Roos NP, Cageorge SM, Nicol JP (1982) How good are the data?-reliability of one health care data bank. *Med Care* 20: 266-276
31. Roos NP, Henteleff PD, Roos LL (1977) A new audit procedure applied to an old question: is the frequency of T&A justified? *Med Care* 15: 1-18
32. Roos NP (1984) Hysterectomies in one Canadian province: a new look at risks and benefits. *Am J Public Health* 74: 39-46
33. Avorn J, Everitt DE, Weiss S (1986) Increased antidepressant use in patients prescribed  $\beta$ -blockers. *JAMA* 255: 357-360
34. Bright RA, Everitt DE (1992)  $\beta$ -blockers and depression: evidence against an association. *JAMA* 267: 1783-1787
35. Johnson JA (1994) *New  $\beta$ -blocker use and concurrent antidepressant prescribing: a pharmacoepidemiologic study*. MSc thesis, University of Saskatchewan, Saskatoon, Saskatchewan
36. Johnson JA, Wallace SM (1995) Linkage of the Saskatchewan health databases to investigate the relationship between  $\beta$ -blocker and antidepressant use. *Pharmacoepidemiol Drug Safety* 4 [Suppl 1]: S30