

A B S T R A C T

Accurate and complete registries are an important source of knowledge about cancer. The concordance of the recording of neoplasms in the Saskatchewan cancer registry with that in hospital charts and death registrations was evaluated for 368 patients. The agreement between registry and hospital charts or death registrations was excellent (kappa: 0.93; 95% confidence interval: 0.89, 0.97), with 91.3% of those with cancer having the same neoplasm recorded in their chart or death registration as in the registry. There was only one patient whose hospital chart indicated cancer who was not in the registry and one apparent major discrepancy relating to the cancer site, which was due to the recording of the primary site in the registry and a secondary in the hospital chart. Although based on a relatively small number of patients, these results suggest a high degree of consistency between cancer registry, hospital charts and death registrations in Saskatchewan.

A B R É G É

Des registres précis et complets sont une source importante d'information sur le cancer. La concordance de l'enregistrement des néoplasmes dans le registre du cancer dans la Saskatchewan, les dossiers des hôpitaux et les certificats de décès a été évaluée pour 368 malades. Le degré d'accord entre le registre, les dossiers des hôpitaux et les certificats de décès était excellent (kappa: 0,93; intervalle de confiance à 95 %: 0,89, 0,97), avec 91,3 % de ceux avec le cancer ayant le même néoplasme enregistré dans leurs dossiers ou leurs certificats de décès comme dans le registre. Il y avait un seul malade où le dossier d'hôpital indiquait un cancer qui n'était pas dans le registre et apparemment, un désaccord important concernant le site d'un cancer où le site primaire avait été inclus dans le registre mais le site secondaire pour le dossier de l'hôpital. Bien qu'ils soient basés sur relativement un petit nombre de malades, ces résultats suggèrent un haut degré de cohérence entre le registre du cancer, les dossiers des hôpitaux et les certificats de décès dans la Saskatchewan.

Concordance on the Recording of Cancer in the Saskatchewan Cancer Agency Registry, Hospital Charts and Death Registrations

Nigel S.B. Rawson, PhD,¹ Diane L. Robson, BA²

Accurate and complete registries are an important source of information for research studies investigating a variety of issues related to disease occurrence.¹ Many countries have established cancer registries and endeavoured to make them as accurate as possible,² although studies to evaluate completeness of coverage and data accuracy have shown a considerable disparity between registries.³⁻¹² Even if the overall accuracy is good, wide variation may exist in the quality of information from different reporting sources or about individual cancers.

In Saskatchewan, cancer is a reportable disease, with information on cases being collected and maintained by the Saskatchewan Cancer Agency (SCA) in its population-based cancer registry, which is one of the oldest in the world having been started in 1932.¹³ The principal source of registrations (91% of all cases excluding non-melanoma skin cancers) is pathology reports. A further 3% of the registrations come from death certificates, and the rest are from physician service claims, hospital reports, physician referrals and an inter-provincial data exchange. All physician service claims with a cancer disease or proce-

dure code are referred to the SCA for approval before payment and any case not registered is followed up with the appropriate physician to verify the diagnosis. The SCA registry is reported to be one of the most complete in Canada,¹⁴ although the evidence is limited to the high rate of microscopically confirmed tumours² and an analysis of the recording of Wilms' tumour cases.¹⁵

Knowledge about the consistency between different data sources within a health care system is important for patient care and research purposes.^{16,17} In a study of aplastic anemia and agranulocytosis¹⁸ – conditions that can result from the disease and treatment processes of cancer – information on patients was obtained from the SCA registry, hospital chart abstracts and death registrations and an evaluation of the concordance between these data sources was possible.

METHODS

The Saskatchewan government supplies a wide range of publicly funded health care programs to provincial residents, while the SCA provides cancer therapy.^{14,19} Each eligible resident is issued with a unique health service number (HSN) that is required to obtain benefit from health care programs (>95% of the population of just over one million are eligible residents; the rest are served by federal programs). The HSN is recorded in the relevant datafile when service is provided and is the key to data linkage. The accuracy of the HSNs is maintained on a daily basis and the reliability of the linkage using the HSN is good. The accuracy and reliability of the Saskatchewan health care datafiles have been found to be acceptable for distinct

1. Professor of Pharmacoepidemiology, Division of Community Health, Faculty of Medicine, and School of Pharmacy, Memorial University of Newfoundland, St. John's, Newfoundland

2. Director of the Cancer Registry, Saskatchewan Cancer Agency, Regina, Saskatchewan

Correspondence and reprint requests: Dr. Nigel S.B. Rawson, Division of Community Health, Memorial University of Newfoundland, Health Sciences Centre, St. John's, NF, A1B 3V6, Fax: 709-737-7382, E-mail: nrawson@mun.ca

Although this study is based on data provided by Saskatchewan Health and the Saskatchewan Cancer Agency, the interpretation and conclusions are those of the authors and do not necessarily represent those of Saskatchewan Health, the Saskatchewan Cancer Agency or the Government of Saskatchewan.

conditions, but poorly defined diagnoses and those determined from laboratory results should be examined with caution.²⁰⁻²⁴ Considerable use has been made of the datafiles for research purposes.^{14,19,25,26}

In the aplastic anemia and agranulocytosis study,¹⁸ all patients with primary discharge diagnosis codes for these dyscrasias but without a secondary code for cancer between 1982 and 1991 were identified from the provincial hospitalization datafile. Information on each patient included gender, age, admission and discharge dates, and primary and secondary discharge diagnoses using four-digit International Classification of Diseases (ICD-9) codes.²⁷ Unique dummy identifiers were substituted for the HSNs before data were supplied for analysis.

Because the computer records do not contain clinical or laboratory data either to confirm the disorder or to rule out alternative causes, health record abstractors obtained this information from hospital charts for the relevant admissions. Hospital charts can be identified accurately, using the HSN and other codes recorded in both the computer record and the chart, and the retrieval rate is normally close to 100%.²⁴ The abstractors were trained to ensure a standardized approach to data collection and used a purpose-designed form to record the information. Death registrations were obtained from the provincial vital statistics records for most of the patients who had died.

The hospital chart abstraction process suggested that many dyscrasia patients with cancer reported in their chart did not have the cancer recorded as a secondary diagnosis in the hospitalization datafile, even though the dyscrasia appeared to be directly related to the neoplasm or its treatment. To maximize the identification of cancer cases, the HSNs, names and addresses of all patients in the study were supplied by Saskatchewan Health directly to the SCA for a check against its registry. The SCA provided International Classification of Diseases for Oncology (ICD-O) diagnostic codes²⁸ for the identified patients.

The concordance regarding cancer occurrence between the SCA registry, hospital charts and death registrations was

Cancer Patients	127
Died, hospital chart abstract and death registration obtained	57
Died, death registration obtained, but no hospital chart abstract	26
Alive, hospital chart abstract obtained	44
Non-cancer Patients	241
Died, hospital chart abstract and death registration obtained	88
Died, hospital chart abstract obtained, but death registration unobtainable	6
Alive, hospital chart abstract obtained	147

quantified and evaluated, using kappa²⁹ with a 95% confidence interval (CI), for those patients for whom information was available. Because there is no direct equivalence between many of the ICD-9 and ICD-O codes and diagnoses from the hospital charts and death registrations were recorded in a verbatim manner, “concordance” was defined as the same site for solid tumours and the same broad type, e.g., leukemia, for hematological cancers.

RESULTS

A total of 397 patients with a primary diagnosis code for aplastic anemia or agranulocytosis were identified from the hospitalization datafile. Twenty-seven patients (9 in the cancer registry and 18 not in it) had neither a hospital chart abstract nor a death registration because the chart was unavailable or the hospitalization occurred prior to the patient being entered in the SCA registry, and, for those who had died, no registration was found. In addition, 1 patient had thrombocytopenia and another had neurofibromatosis recorded on their respective death registrations; these conditions are considered to be benign disease by the SCA not requiring registration unless chemotherapy is needed and neither was in the cancer registry. These 29 patients were excluded from further analysis.

The availability of information about the remaining 368 patients is described in Table I. One hundred and twenty-seven (34.5%) were either recorded in the cancer registry or had a hospital chart or death registration that mentioned cancer. The median age of both these patients and the other 241 was 68 years, with a range of less than 1 year to 95 years.

The sites of the cancers recorded for the 127 patients are listed in Table II. Since

Site of Cancer	No.*
Blood	42
Lymph	19
Breast	13
Colo-rectal	11
Lung	11
Prostate	10
Bladder/kidney	8
Eye/nasopharynx/sinus	5
Lip	3
Ovary	3
Adrenal	2
Brain	2
Multiple myeloma	2
Leg (osteosarcoma)	2
Ewing's	1
Pancreas	1
Thymus	1
Other	3

* 10 patients had two sites and 1 had three sites

the study from which the patients were obtained focussed on blood dyscrasias, it is not surprising that a high proportion of the cancers (48.0%) were hematological or lymphomas. Nevertheless, over half were solid tumours with a range of sites.

Table III shows the concordance of the recording of cancer in the SCA registry, hospital charts and death registrations for all 368 patients. For 124 of the 127 cancer patients (97.6%), the SCA and one or both of the other data sources agreed that the person had cancer and, for 116 (91.3%), the other source(s) recorded the same cancer. The kappa value for Table III is 0.93 (95% CI: 0.89, 0.97), which indicates excellent agreement.³⁰ When hospital charts and death registrations were examined separately, the concordance with the SCA registry was 90.9% and 82.0%, respectively, and the kappa values were 0.93 (95% CI: 0.88, 0.97) and 0.82 (95% CI: 0.74, 0.91).

One patient, a 78-year-old male, had a hospital chart that mentioned adenocarci-

TABLE III
Concordance of Cancer Registry Data, Hospital Charts and Death Registrations

Cancer Registry	Hospital Chart	Death Registration	No.	% (n=368)
Registered	Same neoplasm	Same neoplasm	43	11.7
Registered	Same neoplasm	Different neoplasm	2*	0.5
Registered	Same neoplasm	No mention	6	1.6
Registered	Different neoplasm	Different neoplasm	2*	0.5
Registered	No mention	Same neoplasm	3	0.8
Registered	Not abstracted	Same neoplasm	22	6.0
Registered	Not abstracted	Different neoplasm	3*	0.8
Registered	Not abstracted	No mention	1	0.3
Registered	Same neoplasm	Alive	40	10.9
Registered	Different neoplasm	Alive	2*	0.5
Registered	Chemotherapy for unspecified neoplasm	Alive	1	0.3
Registered	No mention	Alive	1	0.3
Not registered	Adenocarcinoma of the lung	No mention	1	0.3
Not registered	No mention	No mention	88	23.9
Not registered	No mention	Not found	6	1.6
Not registered	No mention	Alive	147	39.9

* Details of these nine discrepant cases are provided in Table IV

TABLE IV
Details of the Nine Discrepant Cases

Case No.	Age/ Sex	Cancer Registry	Hospital Chart	Death Registration
1	68/M	Acute leukemia, NOS	Acute leukemia, NOS	Myelodysplasia
2	31/M	Erythroleukemia	Erythroleukemia	Red cell aplasia
3	83/F	Myeloproliferative disorder and "cancer of uncertain behaviour"	Acute lymphoblastic leukemia	Acute lymphoblastic leukemia
4	78/F	Cancer of the kidney	Cancer of the bladder, rectum, cervix and peritoneum	Cancer of the bladder
5	72/M	Leukemia, NOS	Not abstracted	Pre-leukemia
6	84/F	Leukemia, NOS	Not abstracted	Myelodysplasia
7	83/F	Leukemia, NOS	Not abstracted	Myelodysplasia
8	5/M	Rhabdomyosarcoma of the epididymis	Rhabdomyosarcoma of the testicle	Alive
9	67/F	Choroid melanoma	Cancer of the lung	Alive

NOS: Not otherwise specified

noma of the lung but was not in the registry, and no neoplasm was recorded in his death registration when he died two months after being discharged. There were also two patients recorded in the cancer registry who had either a death registration that did not indicate cancer and no hospital chart abstract (a 78-year-old male), or a hospital chart abstract that did not mention cancer and was alive (a 74-year-old female).

Details of the nine cases in Table III with a hospital chart and/or death registration disagreeing about the cancer type or site recorded by the SCA are provided in Table IV. The most common discrepancy was leukemia being recorded in the registry and a pre-leukemic or leukemia-associated condition reported in the death registra-

tion (cases 1, 2, 5, 6 and 7); cases 5-7 were all "death certificate only" registrations. The hospital chart for case 4 indicated widespread cancer of the "uterine endocervix involving the urinary tract, bladder, rectum and pelvic peritoneum," but the SCA had recorded cancer of the kidney, which may have been a secondary site. The discrepancy in site between registry and hospital chart in case 8 was minor and unsurprising in a 5-year-old child. The 83-year-old female (case 3) whose hospital chart and death registration both recorded acute lymphoblastic leukemia was registered by the SCA as having myeloproliferative disorder and "cancer of uncertain behaviour" because insufficient data were available to record more precisely. These differences are regarded as minor.

The only major discrepancy was the 67-year-old female (case 9) recorded by the SCA as having melanoma of the choroid (a vascular membrane covering most of the posterior of the eye between the retina and the sclera) for whom her hospital chart stated "carcinoma of lung." Further comparison of the hospital chart abstract and the SCA record for this patient indicated that the correct chart had been abstracted. The SCA had correctly recorded the primary cancer (choroid melanoma), whereas a secondary site (lung cancer) diagnosed shortly before the hospitalization in question had been entered in the hospital chart as what appeared to be another primary cancer.

DISCUSSION

This evaluation of the concordance between cancer registry data, hospital charts and death registrations in Saskatchewan is an opportunistic one and, as such, has limitations. Only a relatively small number of patients were included, all of whom had a blood dyscrasia, which may limit the generalizability of the results, although there was no other obvious influence in selecting the patients. No hospital abstract was sought for 25 patients who had a death registration mentioning cancer (Table III) because this was sufficient information to remove them from the aplastic anemia and agranulocytosis study.¹⁸ A range of cancers were included, but there were few patients with some types of cancer. Finally, the level of agreement was limited to the general cancer type or site. Nevertheless, the findings provide information about the reliability and accuracy of the Saskatchewan cancer registry data, which has been sorely lacking.

Concordance on the occurrence of cancer between the data sources was 98% and, for 91%, the registry and hospital chart or death registration agreed on the type or site of the cancer. These figures are impressive, especially considering the high proportion of hematological cancers, which can be difficult to classify.

However, high concordance should be expected with cancer being a reportable disease in Saskatchewan and the requirement for registration before physicians,

who are almost entirely remunerated on a fee-for-service basis, can receive payment for services with an associated cancer diagnosis. These conditions are also reflected in the fact that only 6 of the 127 cancer patients (4.7%) were registered posthumously by the SCA from their death records (the rate of "death certificate only" registrations is usually around 3%). These proportions are much lower than corresponding figures (>20%) reported from England,³¹⁻³³ where cancer registration is non-statutory, less complete^{10,11,16} and often inadequately funded.³⁴ In our data, there was only 1 patient with a hospital chart that recorded a neoplasm who was not registered and 2 registered patients without a chart or a death registration that mentioned cancer (Table III). The single apparent major discrepancy (Table IV), which resulted from the recording of the primary site in the registry and the imprecise reporting of a secondary site in the hospital chart, was probably due to the SCA and the hospital having different recording priorities.

Although there may be some limitations on the generalizability from results based on cases with blood dyscrasias to all patients, this evaluation suggests a high degree of concordance between the cancer registry, hospital charts and death registrations in Saskatchewan. However, because many patients with cancer do not have a secondary diagnosis recorded in the hospitalization discharge datafile (127 in this analysis), this file alone is inadequate to identify cancer patients and linkage with the more reliable SCA data is essential.

REFERENCES

- Wilson S, Prior P, Woodman CBJ. Use of cancer surveillance data for comparative analyses. *J Public Health Med* 1992;14:151-56.
- Canadian Council of Cancer Registries, Health and Welfare Canada, and Statistics Canada. The Making of the Canadian Cancer Registry: Cancer Incidence in Canada and Its Regions, 1969 to 1988. Ottawa: Minister of Supply and Services Canada, 1993.
- Brewster D, Crichton J, Muir C. How accurate are Scottish cancer registration data? *Br J Cancer* 1994;70:954-59.
- Glass S, Gray M, Eden OB, Hann I. Scottish validation study of cancer registration data childhood leukaemia 1968-1981. *Leuk Res* 1987;11:881-85.
- Gulliford MC, Bell J, Bourne HM, Petrukevitch A. The reliability of cancer registry records. *Br J Cancer* 1993;67:819-21.
- Nwene U, Smith A. Assessing completeness of cancer registration in the north-western region of England by a method of independent comparison. *Br J Cancer* 1982;46:635-39.
- Schouten LJ, Jager JJ, van den Brandt PA. Quality of cancer registry data: A comparison of data provided by clinicians with those of registration personnel. *Br J Cancer* 1993;68:974-77.
- Swerdlow AJ, Douglas AJ, Vaughan Hudson G, Vaughan Hudson B. Completeness of cancer registration in England and Wales: An assessment based on 2,145 patients with Hodgkin's disease independently registered by the British National Lymphoma Investigation. *Br J Cancer* 1993;67:326-29.
- Vickers N, Pollock A. Incompleteness and retrieval of case notes in a case note audit of colorectal cancer. *Qual Health Care* 1993;2:170-74.
- Villard-Mackintosh L, Coleman MP, Vessey MP. The completeness of cancer registration in England: An assessment from the Oxford-FPA contraceptive study. *Br J Cancer* 1988;58:507-11.
- Warnakulasuriya KAAS, Acworth P, Bell J, Johnson NW. Incompleteness of oral cancer registration in south-east England, 1971-87. *Br J Cancer* 1994;70:736-38.
- Seddon DJ, Williams EMI. Data quality in population-based cancer registrations: An assessment of the Merseyside and Cheshire Cancer Registry. *Br J Cancer* 1997;76:667-74.
- Benson DL, Robson DL. *Saskatchewan Cancer Atlas 1970-1987*. Regina: Saskatchewan Cancer Foundation, 1988.
- Strand LM, Downey W. Health databases in Saskatchewan. In: Strom BL (Ed.), *Pharmacoepidemiology* 2nd ed. Chichester: Wiley, 1994;217-29.
- Haines CS, Wang PP, Cao Y. Wilms' tumours in Saskatchewan, 1932-1990. *Chron Dis Can* 1994;15:97-101.
- Rushton L, Romaniuk H. Comparison of the diagnosis of leukaemia from death certificates, cancer registration and histological reports: Implications for occupational case-control studies. *Br J Cancer* 1997;75:1694-98.
- Rawson NSB, D'Arcy C. "Validity" and reliability: Idealism and reality in the use of computerized health care databases for pharmacoepidemiological research. *Post Market Surveill* 1991;5:31-55.
- Rawson NSB, Rutledge Harding S, Malcolm E, Lueck L. Hospitalizations for aplastic anemia and agranulocytosis in Saskatchewan: Incidence and associations with antecedent prescription drug use. *J Clin Epidemiol* 1998;51:1343-55.
- Malcolm E, Downey W, Strand LM, et al. Saskatchewan Health's linkable data bases and pharmacoepidemiology. *Post Market Surveill* 1993;6:175-264.
- Rawson NSB, Malcolm E. Validity of the Recording of Cholecystectomy and Hysterectomy in the Saskatchewan Health Care Datafiles. Pharmacoepidemiology Research Unit Technical Report #3. Saskatoon: University of Saskatchewan, 1995.
- Rawson NSB, Malcolm E. Validity of the recording of ischaemic heart disease and chronic obstructive pulmonary disease in the Saskatchewan health care datafiles. *Stat Med* 1995;14:2627-43.
- Edouard L, Rawson NSB. Reliability of the recording of hysterectomy in the Saskatchewan health care system. *Br J Obstet Gynaecol* 1996;103:891-97.
- Rawson NSB, Malcolm E, D'Arcy C. Reliability of the recording of schizophrenia and depressive disorder in the Saskatchewan health care datafiles. *Soc Psychiatry Psychiatr Epidemiol* 1997;32:191-99.
- Rawson NSB, D'Arcy C. Assessing the validity of diagnostic information in administrative health care utilization data: Experience in Saskatchewan. *Pharmacoepidemiol Drug Saf* 1998;7:389-98.
- Tennis P, Andrews E, Bombardier C, et al. Record linkage to conduct an epidemiologic study on the association of rheumatoid arthritis and lymphoma in the province of Saskatchewan, Canada. *J Clin Epidemiol* 1993;46:685-95.
- Risch HA, Howe GR. Menopausal hormone usage and breast cancer in Saskatchewan: A record-linkage cohort study. *Am J Epidemiol* 1994;139:670-83.
- World Health Organization. *Manual of the International Statistical Classification of Diseases, Injuries and Causes of Death* 9th rev. Geneva: WHO, 1977.
- Percy C, Van Holten V, Muir C (Eds.). *International Classification of Diseases for Oncology* 2nd ed. Geneva: World Health Organization, 1990.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
- Altman DG. *Practical Statistics for Medical Research*. London: Chapman & Hall, 1991;404.
- Pollock AM, Vickers N. The impact on colorectal cancer survival of cases registered by "death certificate only": Implication for national survival rates. *Br J Cancer* 1994;70:1229-31.
- Pollock AM, Vickers N. Why are a quarter of all cancer deaths in south-east England registered by death certificate only? - factors related to death certificate only registrations in the Thames Cancer Registry between 1987 and 1989. *Br J Cancer* 1995;71:637-41.
- Pollock AM, Benster R, Vickers N. Why did treatment rates for colorectal cancer in South East England fall between 1982 and 1988? - the effect of case ascertainment and registration. *J Public Health Med* 1995;17:419-28.
- Day NE, Davies TW. Cancer registration: Integrate or disintegrate? *BMJ* 1996;313:896.

Received: October 27, 1999

Accepted: February 15, 2000